ISNN 3030-3907

2026/3
VOLUME 5

DEVELOPMENT OF SCIENCE

ILMIY JURNAL



Development of science

Ilmiy jurnal

2026/3 VOLUME 5

ISSN 3030 -3907

Ilmiy jurnal OAK rayosatining 2024 yil 27 sentabrdagi 361-son qaroriga asosan 02.00.00 kimyo fanlari hamda 2025 yil 12 fevral 367- son qaroriga asosan 05.00.00 texnika fanlari bo'yicha falsafa doktori (PhD) va fan doktori (DSc) ilmiy darajasiga talabgorlarning dissertatsiya ishlari yuzasidan dissertatsiyalari asosiy ilmiy natijalarini chop etish tavsiya etilgan milliy ilmiy nashrlar ro'yxatiga kiritilgan.

BOSH MUHARRIR:
Temirov Alisher Hoshim o‘g‘li – texnika fanlari falsafa doktori (PhD),Dotsent.
TAHRIRIYAT KENGASH RAISI:
Fozilov Sadriddin Fayzullayevich – texnika fanlari doktori, professor
TAHRIRIYAT KENGASH A’ZOLARI

TEXNIKA FANLARI/ TECHNICAL SCIENCES/ ТЕХНИЧЕСКИЕ НАУКИ

Bozorov G‘ayrat Rashidovich
Buxoro Davlat Texnika Universiteti, professor
Abdurahmonov Olim Rustamovich
Buxoro Davlat Texnika Universiteti, professor
Adizov Bobur Zamirovich
O‘zbekiston Respublikasi Fanlar akademiyasi Umumiy va noorganik kimyo instituti, professor
Hayitov Ruslan Rustamjonovich
Buxoro Davlat Texnika Universiteti, professor
Jumayev Qayum Karimovich
Buxoro Davlat Texnika Universiteti, professor.
Berdiyev Dorob Murodovich
Islom Karimov nomidagi Toshkent Davlat Texnika Universiteti, professor
Maxmudov Rafiq Amonovich
Buxoro Davlat Texnika Universiteti, professor.
Tilloyev Lochin Ismatilloyevich
Buxoro Davlat Texnika Universiteti, dotsent.
Raxmonov Qahramon Sanakulovich
Buxoro Davlat Texnika Universiteti, dotsent.
Barakayev Nusratullo Rajabovich
Islom Karimov nomidagi Toshkent Davlat Texnika Universiteti, professor.
Jo‘rayev Xayrullo Fayziyevich
Buxoro Davlat Texnika Universiteti, professor.
Panoyev Erali Rajabboyevich
Buxoro Davlat Texnika Universiteti, dotsent.
Ataullayev Sherzod Nabiullyevich
Buxoro Davlat Texnika Universiteti, dotsent.
Murodov Malik Negmurodovich
Buxoro Davlat Texnika Universiteti, dotsent.
Axmedova Ozoda Bahronovna
Islom Karimov nomidagi Toshkent Davlat Texnika Universiteti.
Safarov Bahri Jumayevich
Buxoro Davlat Texnika Universiteti, dotsent.
Xaydarov Axtam Amonovich
Buxoro Davlat Texnika Universiteti, dotsent.
Safarov Jasur Alijon o‘g‘li
Buxoro Davlat Texnika Universiteti, dotsent.
Асадова Раъно Дилмуратовна
Тошкент Давлат техника университети доценти

KIMYO FANLARI/ CHEMICAL SCIENCES/ ХИМИЧЕСКИЕ НАУКИ

Axmedov Voxid Nizomovich
Buxoro Davlat Texnika Universiteti, professor
Umarov Baqo Bafoyevich
Buxoro Davlat Universiteti, professor
Kodirov Abduaxad Abduraximovich
Qarshi davlat Universiteti, professor
Qurbonov Mingniqul Jumagulovich
Qarshi davlat Universiteti, dotsent
Turemuratov Sharibay Naurizbaevich
O‘zbekiston Respublikasi Fanlar akademiyasi Qoraqalpog‘iston bo‘limi Bosh ilmiy kotibi, professor
Tursunov Kaxor Shonazarovich
Qarshi davlat Universiteti, professor
Do‘stov Hamro
Buxoro Davlat Texnika Universiteti, professor
Panoyev Nodir Shavkatovich
Buxoro Davlat Texnika Universiteti, dotsent.
Olimov Bobur Bahodir o‘g‘li
Buxoro Davlat Texnika Universiteti, dotsent.
Kasimov Sherzod Abduazimovich
Termiz Davlat Universiteti, professor.
Aliqulov Rustam Valiyevich
Termiz Davlat Universiteti, professor
Tropov Xamza Tursunovich
Samarqand Davlat Universiteti, professor

IQTISODIYOT FANLARI/ ECONOMICS/ ЭКОНОМИЧЕСКИЕ НАУКИ

Nizamov Aslitdin Badreddinovich
Buxoro Davlat Texnika Universiteti, professor.
Boboyev Akmal Choriyevich
Buxoro Davlat Texnika Universiteti, professor.
Ergashxojayeva Shaxnoza Djasurovna
Toshkent davlat iqtisodiyot universiteti, professor
Saidaxmedova Nodira Ilxomovna
Toshkent davlat iqtisodiyot universiteti, professor
Muxsinov Bekzod Toxirovich
Buxoro Davlat Texnika Universiteti, dotsent.
Shakarov Qulmat Ashirovich
Renessans ta'lim universiteti,dotsent.
Jumaeva Zulfiya Qayumovna
Buxoro Davlat Texnika Universiteti, dotsent.
Norova Salomat Yusupovna
Buxoro Davlat Texnika Universiteti, dotsent.
Avezova Shaxnoza Maxmudjonovna
Buxoro Davlat Texnika Universiteti, dotsent.
Jalilov Jamshid G‘anijonovich
Toshkent davlat iqtisodiyot universiteti, dotsent
Xasanova Gulruh Djumanazarovna
Buxoro Davlat Texnika Universiteti, dotsent.

PEDAGOGIKA FANLARI/ PEDAGOGICAL SCIENCES / ПЕДАГОГИЧЕСКИЕ НАУКИ

Siddiqova Sadoqat G‘afforovna
Buxoro Davlat Texnika Universiteti, dotsent.
Zoyirov Erkin Xalilovich
Buxoro Davlat Texnika Universiteti, DSc, professor
Gaffarov Feruz Hasanovich
Puchon Universiteti, *professor*
Ruziyev Davron Yuldashevich
Buxoro Davlat Universiteti, professor
Muradova Firuza Rashidovna
Buxoro Davlat Texnika Universiteti, professor.
Yunusova Gulandon Samiyevna
Buxoro Davlat Texnika Universiteti, professor.
To‘xtayeva Zebo Sharifovna
Buxoro Davlat Texnika Universiteti, professor.
Navro‘zova Gulchixra Nigmatovna
Buxoro Davlat Texnika Universiteti, dotsent.
Uzoqova Laylo Polvonovna
Buxoro muhandislik – texnologiya instituti, dotsent.
To‘raqulova Marjona Qiyom qizi
Buxoro Davlat Texnika Universiteti, dotsent.
Muradova Zarina Rashidovna
Buxoro Davlat Texnika Universiteti, dotsent.
Murodov Sanjar Aslonovich
Buxoro Davlat Texnika Universiteti, dotsent.
Shodiyev Jahongir Jo‘raqulovich
Buxoro Davlat Texnika Universiteti, dotsent
Gaffarov Laziz Hasanovich
Buxoro muhandislik – texnologiya instituti, dotsent.

Tahririyat a’zolarning to‘liq ro‘yxati jurnalning www.devos.uz saytida keltirilgan.

Tahririyat hay’ati a’zolarining tarkibini o‘zgartirish hamda qo‘shimcha kiritish yoki tahririyat a’zoligidan chiqarish bosh muharrir tomonidan amalga oshiriladi.

ISSN 3030-3907

Development of science ilmiy jurnali 2024 yil 15 aprel kuni №256445 sonli guvohnoma bilan Davlat ro‘yxatidan o‘tgan.

To‘plam elektron ko‘rinishda (pdf formatida) mualliflarga taqdim etiladi. To‘plamga kiritiladigan maqolalarning mazmuni, ulardagi statistik ma’lumotlar va me’yoriy hujjatlarning to‘g‘riligiga hamda har qanday fikr - mulohazalarga mualliflarning o‘zlari mas’ul hisoblanadilar. Talablarga javob bermaydigan maqolalar to‘plamga kiritilmaydi. Tashkiliy qo‘mita maqola matnini qisqartirish, qisman tuzatish kiritish, seksiyalarga joylashtirish huquqiga ega.

ISSN 3030-3907

SOG'LIQNI SAQLASHDA SINTETIK MA'LUMOTLARNI YARATISH USULLARI

Quzibayev Xudayshukur Shavkat o‘g‘li
Texnika fanlari falsafa doktori, Urganch RANCH texnologiyalar universiteti
Email: khudayshukur.kuzibaev@ranch.university
Otaboyeva Aziza Azamat qizi
Urganch RANCH texnologiyalar universiteti magistri
Email:otaboyevaaziza8gmail.com

Annotatsiya. Ushbu tadqiqot ishida sog‘liqni saqlash sohasida sintetik ma’lumotlarni yaratish usullari, ularning qo‘llanilishi va samaradorligi tahlil qilinadi. Zamonaviy raqamli tibbiyot tizimlarida ma’lumotlar hajmining ortib borishi bilan birga, maxfiylik va ma’lumotlar yetishmovchiligi muammolari dolzarb ahamiyat kasb etmoqda. Shu nuqtai nazardan, sintetik ma’lumotlar sun’iy intellekt modellarini o‘qitish uchun muqobil va xavfsiz yechim sifatida qaraladi. Tadqiqotda turli xil ma’lumot turlari jadval, tasvir, vaqtli qatorlar va omiks ma’lumotlari asosida sintetik ma’lumotlarni generatsiya qilish metodlari ko‘rib chiqilgan. Ular statistik, mashinali o‘rganish va chuqur o‘rganish yondashuvlari asosida tahlil qilingan. Natijalar shuni ko‘rsatadiki, chuqur o‘rganishga asoslangan usullar yuqori samaradorlikka ega bo‘lib, ayniqsa GAN modellar keng qo‘llanilmoqda. Shuningdek, sintetik ma’lumotlar yordamida bemor maxfiyligini saqlagan holda, tibbiy tadqiqotlar samaradorligini oshirish, sun’iy intellekt modellarining aniqligini yaxshilash va sog‘liqni saqlash tizimlarida innovatsion yechimlarni joriy etish imkoniyati mavjud.

Kalit so‘zlar: sintetik ma’lumotlar, sog‘liqni saqlash, sun’iy intellekt, chuqur o‘rganish, ma’lumotlar maxfiyligi, GAN modellari

Аннотация. В данной работе рассматриваются методы генерации синтетических данных в сфере здравоохранения, а также анализируется их эффективность и области применения. В условиях стремительного роста объёмов медицинских данных особую актуальность приобретают вопросы конфиденциальности и ограниченного доступа к данным. В этом контексте синтетические данные выступают в качестве безопасной альтернативы для обучения моделей искусственного интеллекта. В исследовании проанализированы методы генерации синтетических данных для различных типов данных: табличных, изображений, временных рядов и омикс-данных. Методы классифицированы на статистические, машинного обучения и глубокого обучения. Полученные результаты показывают, что методы глубокого обучения, особенно основанные на GAN, обладают высокой эффективностью. Использование синтетических данных позволяет повысить

ISSN 3030-3907

точность моделей, обеспечить конфиденциальность пациентов и ускорить развитие медицинских исследований.

Ключевые слова: синтетические данные, здравоохранение, искусственный интеллект, глубокое обучение, конфиденциальность данных, GAN модели

Abstract. This study explores methods for generating synthetic data in the healthcare domain, along with their applications and effectiveness. With the rapid growth of medical data, issues such as data privacy and limited accessibility have become increasingly important. In this context, synthetic data emerges as a secure alternative for training artificial intelligence models. The research examines various data types, including tabular, image, time-series, and omics data, and analyzes generation methods based on statistical, machine learning, and deep learning approaches. The results indicate that deep learning-based methods, particularly Generative Adversarial Networks (GANs), demonstrate high performance. The use of synthetic data enables improved model accuracy, preservation of patient privacy, and acceleration of innovation in healthcare systems.

Keywords: synthetic data, healthcare, artificial intelligence, deep learning, data privacy, GAN models

1. KIRISH

Elektron sog‘liqni saqlash qaydlari (EHR), taqiladigan sog‘liqni saqlash qurilmalari, genomik ketma-ketlik, tibbiy tasvirlash, mobil sog‘liqni saqlash ilovalari va teletibbiyot kabi raqamli sog‘liqni saqlash texnologiyalarining jadal o‘sishi har kuni ulkan hajmdagi ma’lumotlar ishlab chiqarilishiga olib kelmoqda. Ushbu ma’lumotlar ilg‘or tahlil va sun’iy intellekt (AI) yordamida sog‘liqni saqlash natijalarini sezilarli darajada yaxshilashi mumkin [1,2]. Biroq, bemor ma’lumotlarining nozik (maxfiy) tabiati ulardan foydalanish imkoniyatini cheklaydi va tadqiqot hamda ishlanmalar jarayonida jiddiy to‘siqlarni keltirib chiqaradi [3–5].

Sintetik ma’lumotlar — bu shaxslarning kimligini xavf ostiga qo‘ymasdan, real dunyo ma’lumotlariga taqlid qila oladigan, sun’iy ravishda yaratilgan ma’lumotlardir. Shunday qilib, sintetik ma’lumotlar AQShdagi "Sog‘liqni saqlash sug‘urtasining harakatchanligi va hisobdorligi to‘g‘risida"gi qonun (HIPAA) yoki Yevropadagi "Ma’lumotlarni himoya qilish umumiy reglamenti" (GDPR) kabi qoidalarga muvofiq, bemor maxfiyligini saqlagan holda sog‘liqni saqlash ma’lumotlarining boy imkoniyatlaridan foydalanishning noyob usulini taklif etadi.

Sog‘liqni saqlashda sintetik ma’lumotlarning qiymati juda katta ahamiyatga ega. Sintetik ma’lumotlardan sun’iy intellekt modellarining samaradorligini oshirish, simulyatsiya qilingan klinik sinovlar orqali dori-darmonlarni yaratishni

ISSN 3030-3907

tezlashtirish, mavjud ma’lumotlarni to‘ldirish va ma’lumotlar hajmini oshirish orqali ulardan foydalanish imkoniyatini yaxshilash hamda shaxsni aniqlash mumkin bo‘lgan har qanday ma’lumotdan (PPI) qochgan holda asl ma’lumotlarni takrorlash orqali maxfiylikni himoya qilishda foydalanish mumkin [6–10].

Shunday qilib, sintetik ma’lumotlar nafaqat bemor anonimligini ta’minlaydi, balki tadqiqotchilarga ma’lumotlar mavjudligi bilan bog‘liq to‘siqlarni yengib o‘tishga imkon beradi, bu esa ularga bemorlarning shaxsini oshkor qilish xavfisiz keng ko‘lamli tajribalar va simulyatsiyalar o‘tkazish imkonini beradi. Bundan tashqari, sintetik ma’lumotlar turli populyatsiyalar bo‘ylab AI modellarining umumlashish qobiliyatini (generalizability) yaxshilash uchun yanada xilma-xil va ochiq ma’lumotlarni ishlab chiqishga yordam beradi. Bu, ayniqsa, ma’lumotlar zararli tarafkashliklar (masalan, yosh, irq, jins) kontekstida bir tomonlama yoki yetarli darajada ifodalanmagan holatlarda juda muhimdir; bunda sintetik ma’lumotlarni yaratishdan ushbu muammolarni yumshatish metodologiyasi sifatida foydalanish mumkin.

Sintetik ma’lumotlar AI modellarini o‘qitishda real ma’lumotlarning o‘rnini bosuvchi vosita bo‘lib xizmat qilishi mumkin. Ammo biz sintetik ma’lumotlarni qanday yaratishimiz mumkin? Sintetik ma’lumotlar o‘xshash xususiyatlarga ega yangi ma’lumot nuqtalarini yaratish uchun real ma’lumotlarning statistik xususiyatlarini aniqlash orqali yaratilishi mumkin. Adabiyotlarga ko‘ra, yuqori sifatli sintetik jadval, tasvir, radiomika, vaqtli qatorlar va omiks ma’lumotlarini yaratish uchun turli xil usullar taklif qilingan bo‘lib, ular quyidagicha toifalanadi:

Biroq, maxfiylik muhim omil bo‘lib, u faqat ma’lumotlarning haqqoniyligi (fidelity) bilan kafolatlanmaydi. Shu maqsadda, ma’lumotlarni himoya qilish bo‘yicha eng yaxshi amaliyotlar, shaxsni aniqlash imkoniyatini (identifiability) baholashning aniqroq standartlari va maxfiylikni ta’minlagan holda innovatsiyalarni rag‘batlantirish uchun mutanosib tartibga solish yondashuvlari qabul qilinishi kerak. Shunday qilib, yuqori sifatli sintetik ma’lumotlarning mavjudligi tadqiqotchilarga ko‘p modalli AI modellarini ishlab chiqish imkonini beradi. Bundan tashqari, sintetik ma’lumotlar real ma’lumotlar to‘plamlarida tez-tez uchramaydigan murakkab bemor ssenariylarini simulyatsiya qilish imkonini beradi, shu orqali sog‘liqni saqlash tizimlarining kam uchraydigan, ammo o‘ta muhim holatlarga qarshi mustahkamligini oshiradi. Bundan tashqari, sintetik ma’lumotlardan foydalanish orqali tadqiqotchilar ko‘p modalli ma’lumotlarni birlashtirish va tahlil qilish jarayonida yuzaga keladigan ko‘plab logistik va etik to‘siqlarni chetlab o‘tishlari mumkin, bu esa tadqiqot sur’atini tezlashtiradi. Oxir-oqibat, sintetik ma’lumotlardan foydalanish shaxsiy ma’lumotlar maxfiyligining qat’iy standartlariga rioya qilgan holda, davolash samaradorligini va bemor

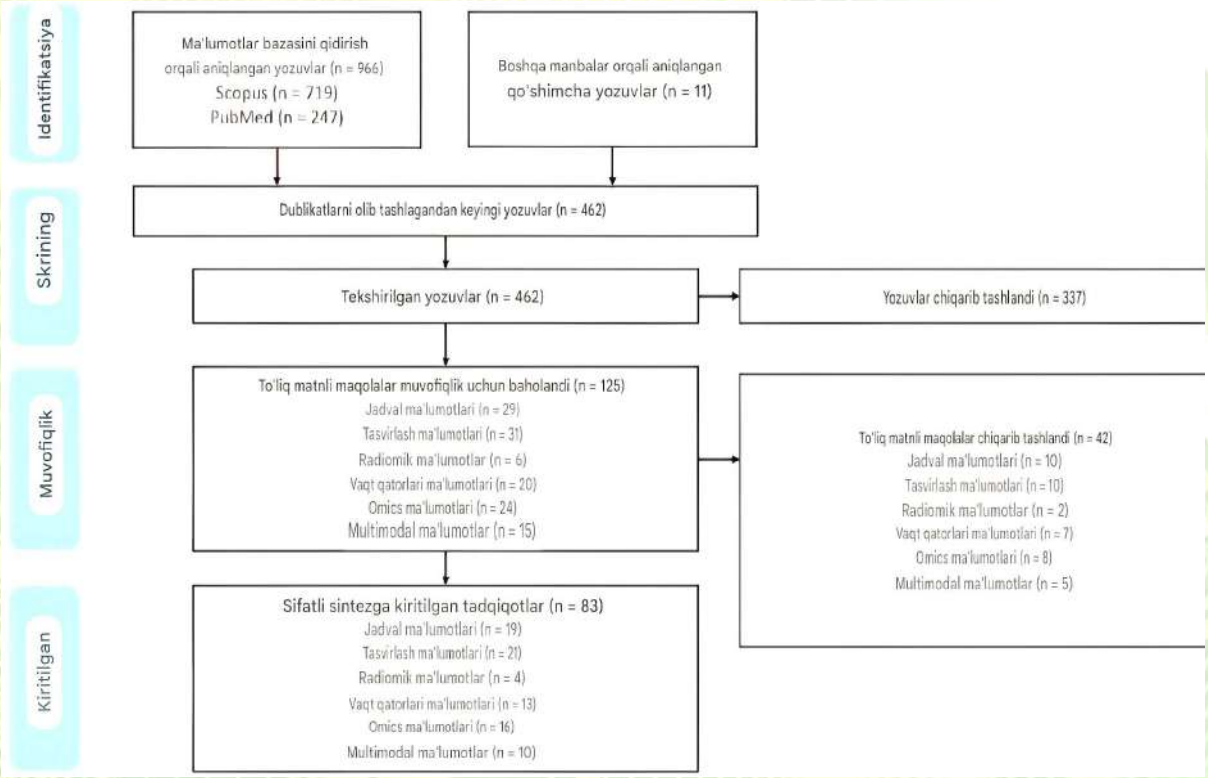
ISSN 3030-3907

natijalarini yaxshilab, shaxsiylashtirilgan tibbiyotni sezilarli darajada rivojlantirishi mumkin.

Ushbu sharhning maqsadi — innovatsiyalarni rag‘batlantirish va turli sog‘liqni saqlash sohalaridagi umumiy muammolarni yanada samarali hal qilish, shuningdek, maqsadli tibbiy tadqiqotlar va amaliyotda sintetik ma’lumotlarning ta’sirini yaxshilash uchun sintetik ma’lumotlarni yaratish metodologiyalari, kodlar va sintetik ma’lumotlarga ega ochiq manbali repozitoriylarni har tomonlama tahlil qilishdan iborat. Ushbu sharhning asosiy maqsadlari quyidagilardan iborat:

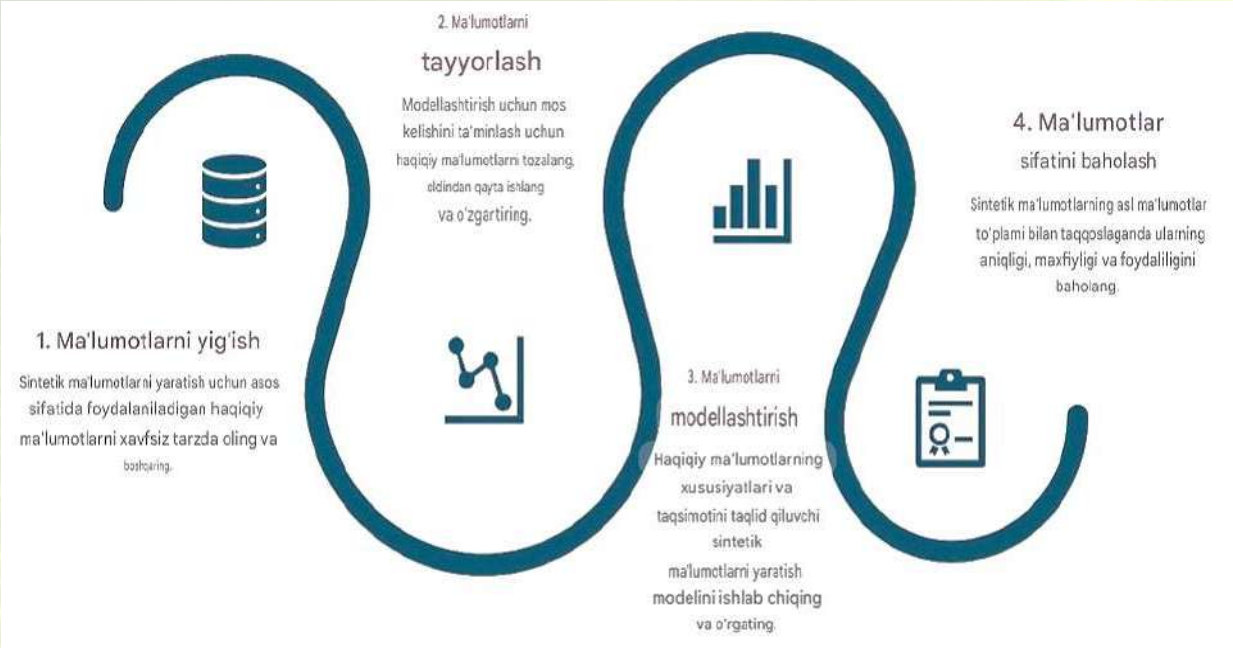
2. METODLAR

2.1. Ko‘rib chiqish jarayoni Biz sog‘liqni saqlashda sintetik ma’lumotlarni yaratish texnologiyalarining turli xil qo‘llanilish holatlarini mustahkam mavzuviy tahlil qilishni ta’minlash maqsadida, yuqori sifatli taqriz qilinadigan jurnallar va xalqaro konferensiyalarga asoslangan PubMed va Scopus ma’lumotlar bazalari bo‘yicha mavjud adabiyotlarning tizimli sharhini o‘tkazdik. Bizning tahlilimiz beshta asosiy ma’lumot turiga qaratilgan: jadval ma’lumotlari, tasviriy ma’lumotlar, radiomika ma’lumotlari (tasvirga asoslangan miqdoriy xususiyatlar), vaqtli qatorlar ma’lumotlari va omiks ma’lumotlari. Shuningdek, ko‘p modalli (multimodal) sintetik ma’lumotlarni yaratishning maxsus holatlari ham o‘rganildi.



1-rasm. Tizimli ko‘rib chiqish uchun PRISMA oqim diagrammasi; bunga ma’lumotlar bazalarini qidirish, saralangan annotatsiyalar soni va olingan to‘liq matnlar kiradi.

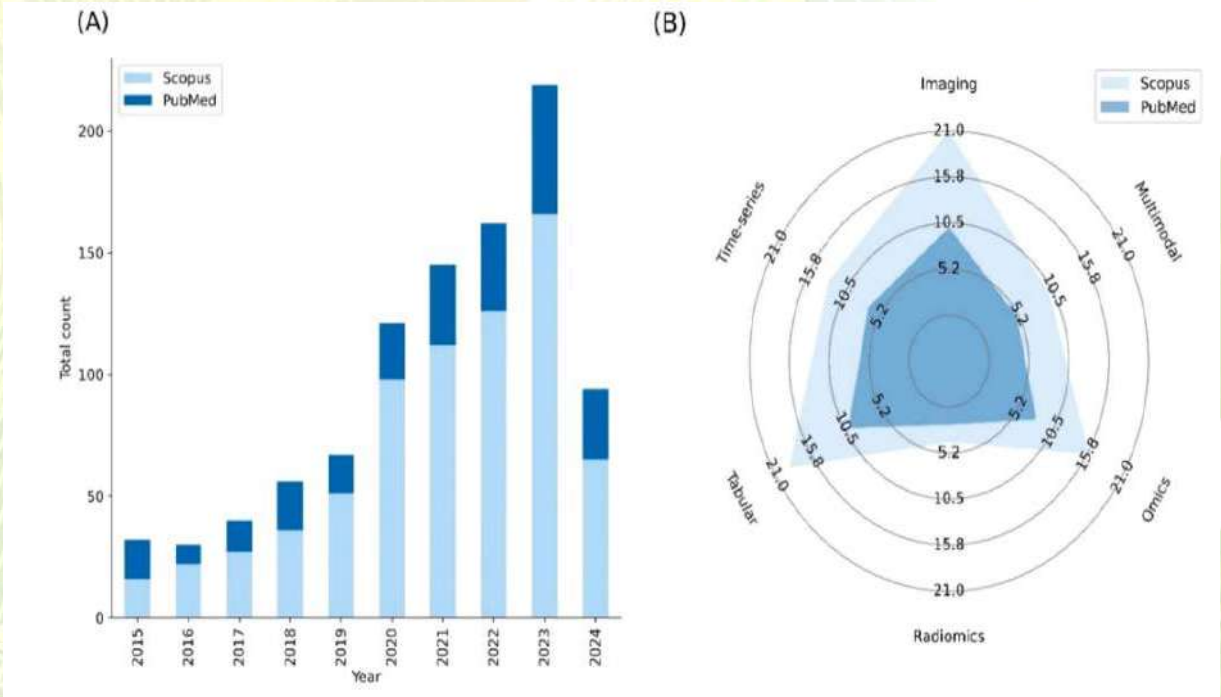
ISSN 3030-3907



2-rasm. Sintetik ma’lumotlarni yaratish ish jarayonining to‘rtta bosqichi

3. NATIJALAR

3.1. Sohadagi tendensiyalarning qisqacha mazmuni



3-rasm. Quyidagilarning umumiy tavsifi: (A) PubMed va Scopus ma'lumotlar bazalari bo'yicha sog'liqni saqlash sohasida sun'iy ma'lumotlarni yaratishga oid tadqiqotlarning yillik umumiy soni; (B) PubMed va Scopus ma'lumotlariga ko'ra, turli xil ma'lumot turlari (beshta asosiy ma'lumot turi va multimodal ma'lumot holatlari) bo'yicha tadqiqotlarning yakuniy soni.

ISSN 3030-3907

1-Jadval

Yo'nalish	Eng yuqori ko'rsatkich	Ulushi
Asosiy usul	Deep Learning (Chuqur o'rganish)	**72.6%**
Asosiy til	Python	**75.3%**

Shuningdek, bir nechta DL yondashuvlarining kombinatsiyasidan iborat bo‘lgan Konvolyutsion neyron tarmoqlari ansambli (Ensemble of CNNs) tasvirlardagi "taqsimotdan tashqari" (out-of-distribution) obyektlarni aniqlash imkoniyatini oshiradi, bu esa ishonchli tibbiy tashxis qo‘yish uchun juda muhimdir. Xuddi shunday, Normallashtiruvchi oqimlar (Normalizing Flows) kompyuter tomografiyasi (KT) ma’lumotlarini olish va rekonstruksiya qilishdagi anomaliyalar ta’sirini kamaytirish, aniqroq va barqaror tasvirlarni ta’minlash uchun qo‘llanilgan.

Nihoyat, tadqiqotchilarga turli xil tibbiy tasvirlar (jumladan, mammografiya, bosh miya MRT, endoskopiya, ko‘krak qafasi rentgenogrammasi, yurak MRT va ko‘krak bezi DCE-MRT) bilan ishlash, ularni yaratish va ulardan foydalanish imkonini beruvchi **medigan** deb nomlangan Python kutubxonasi ishlab chiqilgan. Ushbu kutubxona tarkibida bir nechta oldindan o‘qitilgan (pre-trained) GAN-ga asoslangan modellar (CT-GAN, WGAN, SinGAN, PGGAN, FastGAN, pix2pix) mavjamud.

Asosiy tushunchalar:

- **PET reconstruction:** PET skaneridan olingan xom ma'lumotlarni tushunarli tasvirga aylantirish jarayoni.
- **Vision Transformers:** Tasvirlarni tahlil qilishda "diqqat" (attention) mexanizmidan foydalanuvchi yangi avlod modellari.
- **Out-of-distribution:** Model o'rgatilgan ma'lumotlar to'plamiga kirmaydigan, kutilmagan yoki noodatiy holatlar.
- **Medigan:** Tibbiy sun'iy tasvirlar bilan ishlash uchun maxsus yaratilgan ochiq manbali vosita.

Sun’iy tasviriy ma’lumotlarning aniqligi (fidelity) va sifatini baholashda keng qo‘llaniladigan metrikalar qatoriga Frechet Inception Distance (FID) va Inception Score (IS) kiradi; bular belgilar taqsimotini solishtirish hamda tasvir klassifikatorlarining ish samaradorligini baholash orqali sun’iy ma’lumotlarning real ma’lumotlarga o‘xshashlik darajasini aniqlaydi.

Shuningdek, sun’iy va real tasvirlar o‘rtasidagi vizual va statistik o‘xshashlikni miqdoriy jihatdan ifodalash uchun Strukturaviy o‘xshashlik indeksi (SSIM), Signal-shovqinning cho‘qqi nisbati (PSNR), Normallashtirilgan o‘rtacha kvadratik xato (NMSE) va O‘rtacha mutloq xato (MAE) kabi ko‘rsatkichlardan foydalaniladi.

Mashinali o‘rganishga (ML) asoslangan usullar

ISSN 3030-3907

Bu usullar oddiy nazorat ostidagi o‘rganish (supervised learning) algoritmlaridan ilg‘or sun’iy intellekt modellarigacha farqlanadi. Masalan, regressiya o‘rgatuvchilari bilan jihozlangan ikki darajali Yashirin Markov Modellari (HMM) [22]. Bunda birinchi darajali HMM faoliyatlarning realistik ketma-ketligini yaratsa, ikkinchi darajali HMM ushbu faoliyatlarni aks ettiruvchi sensor hodisalarini shakllantiradi. Regressiya o‘rgatuvchilari vaqt oralig‘i va davomiyligini aniq aks ettirish uchun statistik regressiyani qo‘llaydi. ML usullari statistik usullarga qaraganda ancha moslashuvchan bo‘lib, ko‘pincha ichma-ich joylashgan ketma-ketliklardan iborat sintetik ma’lumotlarni yaratishda qo‘llaniladi.

Chuqur o‘rganishga (DL) asoslangan usullar

DL usullari sog‘liqni saqlashga oid vaqtli qatorlarni sintetik generatsiya qilishning o‘zagini tashkil etadi. Ular ko‘pincha **GAN (Generative Adversarial Networks)** modellariga tayanadi.

- **WGAN-GP [70]:** Treningni barqarorlashtirish va yaqinlashishni (convergence) yaxshilash orqali an’anaviy GANlarni takomillashtiradi.
- **DoppelGANger (DGAN) [70]:** Multi-Layer Perceptron (MLP) yordamida metama’lumotlarni generatsiya qilishning noyob yondashuvini taklif etadi.
- **Time Series GAN (TS-GAN) [71]:** Vaqtinchalik bog‘liqlikni (temporal dependencies) saqlab qolish uchun **LSTM** tarmoqlariga tayanadi.
- **TTS-GAN [29]:** Transformer modelining o‘z-o‘ziga e’tibor (self-attention) mexanizmidan foydalanadi.

Omiks ma’lumotlari (Omics data)

Ilmiy adabiyotlarga ko‘ra, mavjud sintetik omiks generatsiyasi yondashuvlarining aksariyati ko‘p qatlamli omiks ma’lumotlarini (masalan, transkriptomika, metabolomika, proteomika, gen ekspressiyasi) simulyatsiya qilishda o‘rnatilgan statistik prinsiplar va modellarga tayanadi. Bunday usullarga (5-jadval) misol qilib quyidagilarni keltirish mumkin: tasodifiy tanlangan va permutatsiya qilingan boyitilgan yo‘llar [76], kaufzal belgi klasterlari [77], tasodifiy kovariatsiya usuli (RCM) va Kaskad usuli [78], ehtimollik modellashtirish [79], bir tekis taqsimotdan tasodifiy generatsiya [80], MVND [81], darajali qonun taqsimoti [76], tasodifiy perturbatsiyalar (g‘alayonlar) [82], simulyatsiya qilingan chiziqli test (s-test) [83], stoxastik blok modellari (SBM) [14] va stoxastik differentsial tenglamalarga asoslangan metastabillikka ega vaqt bo‘yicha rivojlanuvchi grafiklar [15].

5-jadvalga muvofiq, ushbu usullar quyidagi maqsadlarda qo‘llanilgan:

1. Asosiy taqsimotlarni va yo‘llar/klasterlar soniga asoslangan statistik farazlarni saqlagan holda yarim-sintetik metabolomika ma’lumotlarini ishlab chiqarish;

ISSN 3030-3907

2. Sintetik omiks ma’lumotlarida stratifikatsiyalangan kauzal aniqlash yondashuvlarini validatsiya qilish;
3. Qo‘shimcha tizimli xatoliklarni (biases) hisobga olgan holda gen ekspressiyasi ma’lumotlarini simulyatsiya qilish;
4. Ko-ekspressiya tarmoqlarida o‘sma va normal hujayralar uchun tarmoq topologiyalarini yaratish;
5. Ko‘p omilli geterogen ma’lumotlar tahlilidagi realistik murakkabliklarni modellashtirish;
6. Inson mikrobiomasi dinamikasini o‘rganish va saraton bilan bog‘liq murakkab ko‘p qatlamli omiks ma’lumotlarini modellashtirish.

Chuqur o‘rganishga (DL) asoslangan usullar, jumladan, **WGAN-GP [84]**, **omicsGAN [85]**, **Virtual Autoencoders (VAEs)** va **Deep Boltzmann Machines (DBMs) [86]** ham qo‘llanilgan (5-jadval), ammo nisbatan kamroq darajada. Ular asosan yuqori o‘lchamli mikroarray va lipidomika ma’lumotlaridagi sinflar nomutanosibligi muammolarini hal qilish, kasallik fenotipini bashorat qilishni yaxshilash hamda murakkab omiks ma’lumotlari qonuniyatlari va o‘zaro ta’sirlarining interpretatsiyasini (tushunarliligini) oshirish uchun xizmat qiladi.

Sifat va aniqlikni baholash metrikalari

Sintetik omiks ma’lumotlarining real ma’lumotlarni qanchalik aniq aks ettirishini ta’minlash uchun samaradorlikka asoslangan, statistik va vizual usullarning keng doirasi qo‘llaniladi:

- **Samaradorlik metrikalari:** Tasniflash samaradorligini, klasterlanish o‘xshashligini va belgilar ahamiyatini baholash uchun **Recall**, **Precision**, **AUC**, **Adjusted Rand Index (ARI)**, **Overlap Coefficient (OC)** va **Variable Importance in Projection (VIP)** kabi ko‘rsatkichlardan foydalaniladi.
- **Statistik testlar:** Sintetik va real ma’lumotlar taqsimotini solishtirish uchun **Welch t-testi**, **Student t-testi**, standart og‘ish, **Kaplan-Meier omon qolish grafiklari** va **Log-Rank testi (P-qiymatlari)** ishlatiladi.
- **Vizual tahlil:** Empirik korrelyatsiyalar va normalizatsiya samaradorligini solishtirish uchun ko‘pincha **issiqlik xaritalari (heatmaps)** va **gistogrammalar (bar graphs)** qo‘llaniladi.
- **Ilg‘or metrikalar:** Tarkibiy va vaqtinchalik aniqlikni baholash uchun korrelyatsiya matritsalari orasidagi masofa, **Frobenius normasi**, **diffuziya vaqti** va **entropiya** kabi parametrlar qo‘llaniladi. Shuningdek, klasterlash sifatini o‘lchash uchun **Davies-Bouldin indeksi (DBI)** va modulni aniqlash skori (MDS) kabi maxsus vositalar ishlatiladi.

Omiks va mashinali o‘rganish usullari

ISSN 3030-3907

Omiks sohasida SBM (Stochastic Block Model) murakkab munosabatlar va jamoaviy tuzilmalarni samarali modellashtiradi, biroq u parametrlarni aniq sozlashni talab qiladi va hisoblash jihatidan og‘ir bo‘lishi mumkin. Shunga o‘xshash, RCM va Kaskad usuli realistik gen ekspressiyasi ma’lumotlarini simulyatsiya qilishga qaratilgan, ammo ular biologik murakkabliklarni haddan tashqari soddalashtirib yuborishi xavfi bor.

Bayes modellari va **Daraxt ansambllari (Tree Ensembles)** chiziqli bo‘lmagan ma’lumotlar naqshlari bilan ishlashda moslashuvchanligi bilan ajralib turadi. Ular qaror qabul qilish jarayonlari uchun muhim bo‘lgan noaniqliklarni bashorat qilishda juda samarali. Biroq, ularning samaradorligi ma’lumotlar miqyosi bilan cheklangan va real vaqtdagi muhitlar uchun hisoblash resurslarini ko‘p talab qiladi. **RBF (Radial Basis Function)** asosidagi sun’iy neyron tarmoqlari murakkab chiziqli bo‘lmagan o‘zaro ta’sirlarni boshqarish uchun mo‘ljallangan bo‘lib, naqshlarni aniqlashda kuchli mexanizm taklif etadi, ammo o‘qitish bosqichida katta resurslarni talab qiladi.

Chuqur o‘rganish (DL) asosidagi generatorlar

7-jadvalda, shuningdek, muayyan muammolarni hal qilish uchun mo‘ljallangan turli DL generatorlari taqdim etilgan:

- **ADS-GAN va CTGAN:** Maxfiylikni saqlagan holda sintetik tabulyar (jadval) ma’lumotlarni yaratishda samarali. Ular yuqori o‘lchamli ma’lumotlar bilan yaxshi ishlasa-da, "mode collapse" (rejimning buzilishi) kabi muammolarning oldini olish uchun giperparametrlarni ehtiyotkorlik bilan sozlashni talab qiladi.
- **Enhanced Balancing GAN va Attention-GAN:** Tasvirlarni sintez qilishda sinflar nomutanosibligi va kontrastni yaxshilash kabi muammolarni hal qiladi. Ular murakkab detallarni yaxshi qamrab oladi, biroq kichik ma’lumotlar to‘plamlarida "overfitting" (qayta o‘qitish) muammosiga moyil bo‘lishi mumkin.
- **Contrastive Diffusion Model va Flow-Based Network:** Yuqori aniqlikdagi va nozik detallarga ega tasvirlarni yaratishda yetakchi hisoblanadi. Ular super-rezolyutsiya kabi vazifalarda samarali, ammo o‘ta katta hisoblash resurslarini talab qiladi.

Tibbiy vizuallashtirishda **TGAN** va **CMSG-Net** kabi modellar annotatsiya qilingan (belgilangan) tasvirlarning tanqisligini hal qilish va AI modellarini o‘qitish uchun ma’lumotlarning foydaliligini oshirishda keng qo‘llaniladi. Omiks sohasida esa **omicsGAN** va ehtimollik modellashtirish usullari murakkab biologik naqshlarni simulyatsiya qilish uchun ilg‘or GAN texnologiyalaridan foydalanadi. Bu fenotipni bashorat qilishni yaxshilasa-da, qayta o‘qitish (**overfitting**) muammosidan qochish uchun katta ma’lumotlar to‘plamini talab qiladi.

ISSN 3030-3907

Generator turlarining umumiy tavsifi

- **DL-generatorlar (Chuqur o‘rganish):** Murakkab va yuqori o‘lchamli ma’lumotlar bilan ishlash qobiliyatiga ega, ammo bu yuqori hisoblash resurslari, ma’lumotlar hajmi va modelni sinchkovlik bilan sozlash evaziga erishiladi.
- **ML-generatorlar (Mashinali o‘rganish):** Murakkab, chiziqli bo‘lmagan munosabatlarni modellashtirishda mustahkam va hisoblash jihatidan nisbatan samaralidir.
- **Ehtimollik modellari:** Modellashtirish jarayoniga noaniqlikni (uncertainty) kiritish xususiyati bilan ajralib turadi. Biroq, ular ehtiyotkorlik bilan dizayn qilishni va parametrlarni sozlashni talab qiladi, shuningdek, murakkab ma’lumotlar bilan ishlashda kengayish (scalability) imkoniyati cheklangan bo‘lishi mumkin.

Ishonchli Sun’iy Intellekt (TwAI) va Sintetik ma’lumotlar

Sintetik ma’lumotlar maxfiylik, adolat va mustahkamlikni oshirish orqali **Ishonchli AI (Trustworthy AI)** tamoyillariga sezilarli hissa qo‘shadi:

1. **Maxfiylik (Privacy):** Real ma’lumotlarga "taqlid qiluvchi", ammo shaxsiy yoki nozik ma’lumotlarni o‘z ichiga olmaydigan sintetik ma’lumotlarni yaratish shaxsiy daxlsizlikni himoya qiladi va ma’lumotlarning sizib chiqish xavfini kamaytiradi.

2. **Adolatlilik (Fairness):** Sintetik ma’lumotlar real ma’lumotlardagi tizimli xatoliklarni (biases) tuzatish uchun ishlatilishi mumkin (masalan, demografik tafovutlarni kamaytirish uchun kam vakillik qilingan guruhlar sonini sun’iy oshirish). Bu AI modellarida kamsitishni (discrimination) kamaytiradi.

3. **Mustahkamlik (Robustness):** Ma’lumotlarni ko‘paytirish (augmentation) uchun turli va yuqori sifatli ma’lumotlarni taqdim etish orqali mustahkam AI modellarini yaratishga imkon beradi, shuningdek, modelning adversarial (tajovuzkor) hujumlarga nisbatan zaifligini kamaytiradi.

5. XULOSA VA KELAJAKDAGI YO‘NALISHLAR

Ushbu sharh sog‘liqni saqlash sohasida tabulyar, tasviriy, radiomika, vaqtli qatorlar va omiks kabi turli ma’lumotlar modalliklari bo‘yicha sintetik ma’lumotlar generatsiyasi texnologiyalarini ishlab chiqish va joriy etishga qaratilgan tadqiqotlar soni eksponentsial ravishda o‘sib borayotganini ko‘rsatadi. Ushbu tadqiqotlar sintetik ma’lumotlardan nafaqat maxfiylik muammolarini hal qilish, balki real ma’lumotlarning mavjudligi va xilma-xilligini oshirish uchun ham foydalanadi. Bu esa bemorlar natijalarini yaxshilash va tibbiy tadqiqotlarni qo‘llab-quvvatlash uchun AI-ga asoslangan diagnostika va bashorat qilish modellarini o‘qitishda o‘ta muhimdir.

Turli modalliklar bo‘yicha asosiy xulosalar:

ISSN 3030-3907

• Tabulyar ma’lumotlar: Klinik sinovlar simulyatsiyasi va kasallik rivojlanishini modellashtirish uchun statistik (MVND, bootstrapping) va ehtimollik (Bayes modellari) usullari keng qo‘llaniladi. GMM va Tree Ensembles kabi mashinali o‘rganish usullari esa *in silico* klinik sinovlar uchun keng ko‘lamli virtual populyatsiyalarni yaratishda samarali.

• Tasviriy va Radiomika ma’lumotlari: CycleGAN va Enhanced Balancing GAN kam sonli sinflar (minority classes) muammosini hal qilish va uslublarni o‘tkazishda yetakchi o‘rin tutadi. Radiomikada WGAN-GP va CTGAN kabi modellar turli tibbiy holatlarni farqlashga yordam beruvchi sintetik belgilarni yaratishda muhim rol o‘ynaydi.

• Vaqtli qatorlar (Time-series): Temporal (vaqtinchalik) dinamika muhim bo‘lgan sohalarda WGAN-GP va MTGAN mustahkam yechimlarni taklif etadi. Transformer-Based GANs va CR-VAEs kabi murakkab arxitekturalar vaqtli bog‘liqliklarni saqlash va ma’lumotlar aniqligini oshirish yo‘nalishidagi evolyutsiyani ko‘rsatadi.

• Omiks va Multimodal ma’lumotlar: OmicsGAN va DBM kabi chuqur o‘rganish usullari murakkab biologik jarayonlarni talqin qilishni osonlashtiradi. Multimodal sohada esa MultiNODEs va TC-MultiGAN kabi ilg‘or neyron tarmoqlar turli ma’lumotlar turlarini (masalan, EHR va MRT tasvirlari) yagona tizimga integratsiya qilish imkonini beradi

ADABIYOTLAR RO'YXATI

1. Shilo S, Rossman H, Segal E. Inqilob o‘qlari: sog‘liqni saqlashda katta ma’lumotlarning muammolari va istiqbollari. *Nat Med* 2020; vol. 26(1):29–38. https://doi.org/10.1038/s41591-019-0727-5.

2. Agrawal R, Prabakaran S. Raqamli sog‘liqni saqlashda katta ma’lumotlar: umumiy amaliyot uchun olingan saboqlar va tavsiyalar. *Heredity* 2020; vol. 124(4):525–34. https://doi.org/10.1038/s41437-020-0303-2.

3. Appenzeller A, Leitner M, Philipp P, Krempel E, Beyerer J. Tibbiy ma’lumotlar tahlili uchun shaxsiy sintetik ma’lumotlarning maxfiyligi va foydaliligi. *Appl Sci* 2022; vol. 12(23):12320. https://doi.org/10.3390/app122312320.

4. S.M. Bellovin, P.K. Dutta, N. Reitinger, Maxfiylik va sintetik ma’lumotlar to‘plamlari, vol. 22.

5. Yale A, Dash S, Dutta R, Guyon I, Pavao A, Bennett KP. Maxfiylikni saqlaydigan sintetik salomatlik ma’lumotlarini yaratish va baholash. *Neurocomputing* 2020; vol. 416:244–55. https://doi.org/10.1016/j.neucom.2019.12.136.

ISSN 3030-3907

6. Gonzales A, Guruswamy G, Smith SR. Sog‘liqni saqlashda sintetik ma’lumotlar: narrativ sharh. *PLOS Digit Health* 2023; vol. 2(1):e0000082. https://doi.org/10.1371/journal.pdig.0000082.
7. Murtaza H, Ahmed M, Khan NF, Murtaza G, Zafar S, Bano A. Sintetik ma’lumotlar generatsiyasi: sog‘liqni saqlash sohasidagi zamonaviy holat. *Comput Sci Rev* 2023; vol. 48:100546. https://doi.org/10.1016/j.cosrev.2023.100546.
8. J. Jordon va boshqalar, “Sintetik ma’lumotlar – nima, nima uchun va qanday?” arXiv, 2022-yil 6-may. Kirish sanasi: 2024-yil 28-may. [Onlayn]. Mavjud: 〈 http://arxiv.org/abs/2205.03257〉 .
9. Figueira A, Vaz B. Sintetik ma’lumotlar generatsiyasi, baholash usullari va GANlar bo‘yicha tadqiqot. *Mathematics* 2022; vol. 10(15):2733. https://doi.org/10.3390/math10152733.
10. O. Mendelevitch, “Metodlar va eksperimental natijalar sharhi”.
11. Cheng V, Suriyakumar VM, Dullerud N, Joshi S, Ghassemi M. Haqiqiy bo‘lguncha soxtalashtira olasizmi?: Differensial maxfiylikdagi sintetik ma’lumotlarning tasniflash adolatliligiga ta’siri. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, Virtual tadbir. Kanada: ACM; 2021. p. 149–60. https://doi.org/10.1145/3442188.3445879.
12. Ferrara E. Sun’iy intellektdagi adolatlilik va tarafkashlik: manbalar, ta’sirlar va yumshatish strategiyalari bo‘yicha qisqacha sharh. *Sci* 2023; vol. 6(1):3. https://doi.org/10.3390/sci6010003.
13. Smania G, Jonsson EN. Kovariatlar simulyatsiyasi uchun muqobil usul sifatida shartli taqsimotni modellashtirish: birgalikdagi ko'p o'lchovli normal va bootstrap usullari bilan taqqoslash. *CPT Pharmacomet Syst Pharmacol* 2021; vol. 10(4):330–9. https://doi.org/10.1002/psp4.12613.

ISSN 3030-3907