© 2026 International Conference on Multidisciplinary Sciences and Innovations . All Rights Reserved.

FORENSICLLM: RAQAMLI KRIMINALISTIKA UCHUN MAHALLIY (LOKAL) KATTA TIL MODELI — ARXITEKTURA, METODOLOGIYA VA SAMARADORLIK TAHLILI

Yusupova Surayyo Muyasar qizi
suray31052002@gmail.com

Annotatsiya

Ushbu maqolada raqamli kriminalistika sohasida ishlab chiqilgan mahalliy (lokal) katta til modeli — ForensicLLM tizimining arxitekturaviy, metodologik va empirik jihatlari chuqur ilmiy tahlil qilinadi. AQShning Luiziana shtat universiteti tadqiqotchilari tomonidan 2025-yilda taqdim etilgan ushbu model LLaMA-3.1-8B asosida Retrieval Augmented Fine-tuning (RAFT) usuli yordamida 1082 ta ilmiy maqola va 1390 ta kriminalistik artefakt ma'lumotlari asosida sozlangan. Maqolada modelning texnik asoslari — tokenizatsiya, joylashtirish (embedding), diqqat mexanizmi (attention), kvantizatsiya hamda LoRA/QLoRA moslashtirish usullari — shuningdek uning baholash natijalari (BERTScore, BGE-M3, G-Eval metrikalari bo'yicha bazaviy modeldan ustunligi va manba ko'rsatishdagi 86,6 foizlik aniqligi) tahlil qilinadi. Alohida e'tibor foydalanuvchi tadqiqoti natijalariga, model cheklovlariga hamda AQSh Daubert mezoni asosidagi dalil qabul qilinishi talablariga qaratiladi. Xalqaro (Yevropa, Rossiya, Singapur) va milliy (O'zbekiston) kontekst bilan solishtirilgan holda, ForensicLLM'ning DFIR sohasidagi ahamiyati va istiqbollari baholanadi.

Kalit so'zlar: ForensicLLM, raqamli kriminalistika, LLaMA, RAFT, RAG, fine-tuning, LoRA, Daubert mezoni, dalil qabul qilinishi, katta til modeli.

Kirish

Raqamli kriminalistika (Digital Forensics, DF) raqamli dalillarni qonuniy va ilmiy jihatdan asoslangan tarzda olish, autentifikatsiya qilish va tahlil qilishni o'z ichiga oladi, bunda tergovchilar uchun eng muhim vazifalardan biri — olingan dalilning sud jarayonida qabul qilinishini ta'minlashdir [2]. AQShda 1993-yilgi Daubert v. Merrell Dow Pharmaceuticals ishi asosida shakllangan Daubert mezoni ilmiy metodologiyaning haqqoniyligi va ishonchliligini baholash uchun to'rtta asosiy mezonni belgilaydi: metodologiyani sinovdan o'tkazish imkoniyati va uning oldingi sinovlari, ma'lum xatolik darajasi, taqrizdan o'tganlik va nashr etilganlik holati, hamda tegishli ilmiy hamjamiyat tomonidan tan olinganlik [2]. Shu sababli, DF sohasida qo'llaniladigan har qanday vosita va uslub taqrizdan o'tgan va ilmiy jihatdan tan olingan bo'lishi shart. Katta til modellari (LLM) ko'plab tabiiy til vazifalarida yuqori natija ko'rsatsa-da, ularning o'quv ma'lumotlari va ichki mexanizmlari haqidagi shaffoflik yetishmasligi natijalarning ilmiy asosini tekshirishni murakkablashtiradi; bundan tashqari, DF sohasida ma'lumotlar maxfiyligi hal qiluvchi ahamiyatga ega bo'lib, bulutli LLM'lardan foydalanish nozik ma'lumotlarning uchinchi tomonlarga oshkor bo'lish xavfini keltirib chiqaradi [2]. ChatGPT kabi umumiy maqsadli bulutli modellar internetga ulanishni talab qiladi, bu esa internetga kirish cheklangan xavfsiz muhitlarda ularni amaliy jihatdan foydasiz qiladi, shuningdek bunday modellar keng qamrovli qo'llanilishga ega bo'lsa-da, ko'pincha sohaga xos mutaxassislikka ega emas [2; 3].

LLaMA va Mistral kabi kichikroq ochiq manbali modellar muqobil yechim sifatida, domenga xos ma'lumotlar to'plamida moslashtirilganda kattaroq modellarga qiyoslanadigan natijalarga erishishi mumkin, bunda kvantizatsiya va Low-Rank Adaptation (LoRA) kabi usullar hisoblash talablarini sezilarli darajada kamaytiradi [2]. Mavjud tadqiqotlarning aksariyati kattaroq bulutli modellarni baholashga qaratilgan bo'lib, kichikroq, ixtisoslashtirilgan modellarni o'rganish deyarli amalga oshirilmagan [2].

ForensicLLM gacha bo'lgan tadqiqotlar asosan mavjud umumiy maqsadli modellarning DF sohasidagi qo'llanilishini o'rganishga qaratilgan edi: Scanlon va boshqalar (2023) ChatGPT'ni

© 2026 International Conference on Multidisciplinary Sciences and Innovations . All Rights Reserved.

artefaktlarni tahlil qilish, dalil izlash va hodisaga javob berish kabi DF vazifalarida baholab, model muntazam ifodalar generatsiya qilish, kalit so'z qidiruvini kuchaytirish, hujjatlarni qisqartirish va kodlashda yordam berishda salohiyat ko'rsatganini, biroq o'quv ma'lumotlaridagi noaniqlik, domenga xos mutaxassislikning yo'qligi, natijalarning nodeterministik xarakteri va gallyutsinatsiyaga moyillik kabi jiddiy cheklovlarga ega ekanligini aniqlagan [3]; Michelet va Breitinger (2023) ChatGPT-3.5 va lokal LLaMA-2-13B modellarini DF hisobotlarini yaratishda taqqoslab, ChatGPT'ning aniqroq va to'liqroq matn generatsiya qilishda ustun kelganini, biroq ikkala model ham inson tomonidan sezilarli darajada tahrirlashni talab qilganini ko'rsatgan; Henseler va van Beek (2023) esa ChatGPT'ning yuridik mutaxassislarga raqamli tergovda yordam berish salohiyatini o'rganib, samaradorlikni oshirsa-da, inson nazorati muhim bo'lib qolishini xulosa qilgan. Bundan farqli o'laroq, Silalahi va boshqalar (2023) mavjud modelni tahlil qilish o'rniga, dron parvoz jurnallarida anomaliyalarni aniqlash uchun BERT modelini salbiy sentimentni tanib olishga maxsus moslashtirgan, bu esa hodisani ko'rsatuvchi bilvosita usul edi. Ushbu tadqiqotlar birgalikda shuni ko'rsatadiki, ForensicLLM'gacha bo'lgan davrda DF sohasidagi LLM tadqiqotlari asosan mavjud, umumiy maqsadli modellarning imkoniyat va cheklovlarini hujjatlashtirish bilan cheklangan bo'lib, domenga maxsus moslashtirilgan, mahalliy ishlaydigan va tekshiriluvchi manba asosidagi modelni yaratishga qaratilgan tizimli urinish deyarli mavjud emas edi, bu esa aynan ForensicLLM tadqiqotining ilmiy yangiligini belgilaydi [2].

Mazkur maqolaning maqsadi — aynan shu bo'shliqni to'ldirish uchun ishlab chiqilgan ForensicLLM tizimini, uning ilmiy manbasi bo'lgan Sharma va boshqalar tomonidan 2025-yilda Forensic Science International: Digital Investigation jurnalida chop etilgan tadqiqot asosida, chuqur nazariy-texnik jihatdan tahlil qilishdir [2]. Ushbu tadqiqot uchta tadqiqot savolini o'rtaga qo'ygan edi: birinchidan, fine-tuning va RAG yondashuvlari DF sohasida samaradorlik jihatidan qanday taqqoslanadi; ikkinchidan, sozlangan LLM DF tergovlariga yordam berish uchun manbalarni qanchalik ishonchli tarzda tiklay oladi; uchinchidan, lokal LLM raqamli artefaktlar, kriminalistik vositalar va jarayonlarni tushunishni qanday yaxshilashi mumkin [2]. Mazkur maqolada ushbu uch savolga berilgan javoblar ForensicLLM'ning arxitekturaviy tuzilishi, ma'lumotlar to'plami, RAFT metodologiyasi va baholash natijalari nuqtai nazaridan batafsil yoritiladi, shuningdek uning xalqaro (Yevropa, Rossiya, Osiyo) tajriba va O'zbekiston milliy konteksti bilan bog'liqligi ko'rib chiqiladi. Tadqiqot metodologiyasi manba maqolani chuqur kontent-tahlil qilish, uni yondosh xalqaro va milliy ilmiy adabiyot bilan taqqoslash usullariga asoslanadi.

1. ForensicLLM tizimining nazariy-texnik asoslari: tokenizatsiya, joylashtirish, diqqat (attention) mexanizmi va parametrlarni samarali moslashtirish usullari

ForensicLLM'ning texnik arxitekturasini tushunish uchun avvalo katta til modellarining bir qator asosiy tushunchalarini ko'rib chiqish zarur, chunki aynan shu tushunchalar modelning DF sohasiga moslashtirilishi mumkinligini belgilaydi. Tokenizatsiya jarayoni kirish matnini kichikroq birliklarga — so'z, bo'lak-so'z yoki belgilarga — bo'lishdan iborat bo'lib, LLaMA-3 modeli 128 mingdan ortiq noyob tokenni o'z ichiga olgan tiktoken Byte Pair Encoding (BPE) tokenizatoridan foydalanadi, bu yondashuv kamdan-kam uchraydigan so'zlarni ma'lum bo'lgan bo'lak-so'z birliklariga bo'lish orqali modelning matnni qayta ishlash qobiliyatini oshiradi [2]. Token joylashtirishlari (token embeddings) — tokenning semantik ma'nosini va tokenlar o'rtasidagi bog'liqlikni o'zida mujassam etuvchi sonli vektorlar bo'lib, model ishlash jarayonida token joylashtirish qatlami oddiy qidiruv jadvali vazifasini bajarib, tokenning noyob identifikatori asosida uning vektor ko'rinishini qaytaradi; ko'pchilik zamonaviy modellarda ushbu qatlam dastlabki o'qitish jarayonida shakllansa, ba'zi hollarda word2vec yoki GLOVE kabi oldindan o'qitilgan token joylashtirish modellaridan foydalaniladi [2].

Diqqat mexanizmi (attention) LLaMA va GPT kabi transformer-asosidagi modellarning asosiy elementi bo'lib, u modelga natija generatsiya qilishda kirish ketma-ketligining tegishli qismlariga

© 2026 International Conference on Multidisciplinary Sciences and Innovations . All Rights Reserved.

e'tibor qaratish imkonini beradi; xususan, transformerlardagi o'z-o'ziga diqqat (self-attention) qatlami kirish ketma-ketligidagi har bir tokenning boshqa barcha tokenlarga nisbatan ahamiyatini hisoblab chiqadi, bu esa modelga uzoq masofali bog'liqliklarni aniqlash va kontekstni chuqurroq tushunish imkonini beradi [2]. Kvantizatsiya — sonli ma'lumotlarning aniqligini pasaytirish, odatda 32-bitli suzuvchi nuqtali og'irliklarni pastroq aniqlikdagi formatlarga (masalan, 16-bitli yoki 8-bitli) aylantirish orqali LLM'larni optimallashtirish usuli bo'lib, bu xotira sarfi va hisoblash xarajatlarini kamaytiradi hamda modelning aniqligini sezilarli darajada pasaytirmagan holda cheklangan resursli qurilmalarga joylashtirish imkonini beradi [2]. Jumla joylashtirishlari (sentence embeddings) va vektor bazalari token joylashtirishlariga o'xshab, uzunroq matnning semantik ma'nosini ifodalaydigan, ko'pincha yuzlab tokenlardan iborat katta kontekstlarni qayta ishlashga mo'ljallangan bo'lib, ularning asosiy maqsadi semantik axborotni qidirib topishni kuchaytirishdir; vektor joylashtirish modellari semantik jihatdan o'xshash ma'lumotlar bir-biriga yaqinroq vektor ko'rinishiga ega bo'lishi, o'xshash bo'lmagan ma'lumotlar esa bir-biridan uzoqroq joylashishi uchun o'qitiladi, bu masofa nuqta ko'paytmasi yoki kosinus masofasi kabi o'xshashlik metrikalari orqali o'lchanadi, FAISS (Facebook AI Similarity Search) va ChromaDB kabi vektor bazalari esa aynan yuqori o'lchamli vektorlarni samarali qidirib topish uchun ixtisoslashtirilgan saqlash tizimlaridir [2].

Moslashtirish (fine-tuning) — LLM'ni domenga xos ma'lumotlar to'plamida qo'shimcha o'qitish jarayoni bo'lib, bu model parametrlarini maqsadli domenning o'ziga xos talablariga moslashtirish, shu orqali tegishli vazifalardagi samaradorlikni oshirish imkonini beradi [2]. LoRA (Low-Rank Adaptation) va uning kvantizatsiyalangan varianti QLoRA usullari moslashtirish jarayonining hisoblash samaradorligini sezilarli darajada oshirish uchun ishlab chiqilgan bo'lib, iste'molchi darajasidagi GPU'larda ham kuchli LLM'larni moslashtirish imkonini beradi; LoRA modeldagi har bir og'irlik matritsasi (W) uchun ikkita past darajali (low-rank) modifikatsiya matritsasi (A va B) joriy etadi — agar W matritsa m×n o'lchamli bo'lsa, A matritsa m×r, B matritsa esa r×n o'lchamli bo'ladi, bu yerda r — daraja (rank) qiymati bo'lib, moslashtirish jarayonida butun W matritsasini yangilash o'rniga, LoRA faqat kichik A va B matritsalarini sozlaydi, natijada yangilangan og'irlik matritsasi W' = W + AB formulasi orqali hisoblanadi [2; 6]. QLoRA esa LoRA'ning kengaytmasi bo'lib, moslashtirish jarayonida past darajali modifikatsiya matritsalarini (A, B) standart 32-bitli suzuvchi nuqtali sonlar o'rniga pastroq aniqlikdagi ma'lumot turlari (masalan, 4-bitli butun sonlar) yordamida ifodalash orqali kvantizatsiya qiladi, bu jarayon Aq = quantize(A), Bq = quantize(B) va W' = W + AqBq formulalari orqali ifodalanadi, bunda quantize(·) — A va B dagi suzuvchi nuqtali qiymatlarni pastroq aniqlikdagi butun sonlarga aylantiruvchi funksiya hisoblanadi [2; 7].

ForensicLLM tizimi uchun asosiy model sifatida Meta AI tomonidan yaqinda chiqarilgan, sakkiz milliard parametrli ochiq manbali LLaMA-3.1-8B modeli tanlangan, chunki u Gemma-7B va Mistral-7B kabi shu o'lchamdagi o'xshash modellarga nisbatan yuqoriroq samaradorlik ko'rsatgan, shuningdek kvantizatsiya bilan birga ushbu model Nvidia RTX 4090 kabi iste'molchi darajasidagi GPU'da sekundiga 100 dan ortiq token tezligida ishlashi mumkin [2]. shuningdek Retrieval Augmented Generation (RAG) yondashuvi modellarga inferens vaqtida tashqi bilim manbalaridan qo'shimcha kontekst ma'lumotini qidirib topish imkonini beradi, bu esa modelning bilimini shunchaki o'quv ma'lumotlaridan tashqariga samarali kengaytiradi [2; 8]. Retrieval Augmented Fine-tuning (RAFT) esa ushbu ikki yondashuvni — moslashtirish va RAG'ni — birlashtirgan usul bo'lib, moslashtirish ma'lumotlar to'plamiga qidiruv bazasidan olingan tegishli kontekst ma'lumotini kiritish orqali, model ushbu qo'shimcha ma'lumotni o'zlashtirish va undan mulohaza yuritishni o'rganadi, bu esa aniq va yaxshi asoslangan javoblar berish qobiliyatini kuchaytiradi [2; 8]. Ushbu texnik asoslarning yig'indisi — ya'ni resurs tejamkor moslashtirish orqali yengil vaznli mahalliy modellarni amaliy muqobil yechimga aylantirish tendensiyasi — nafaqat ForensicLLM tadqiqotida, balki Yevropa (Irlandiya) tadqiqotchilari tomonidan o'tkazilgan kengroq sharh ishida [4] va Singapurda

© 2026 International Conference on Multidisciplinary Sciences and Innovations . All Rights Reserved.

nashr etilgan mintaqaviy tahliliy bobda [5] ham DF sohasida LLM integratsiyasining asosiy strategik yo'nalishi sifatida alohida ta'kidlanadi.

2. ForensicLLM'ning ma'lumotlar to'plami, Retrieval Augmented Fine-tuning (RAFT) metodologiyasi va miqdoriy baholash natijalari

ForensicLLM ikki asosiy ma'lumot manbai — DF ilmiy maqolalari va kriminalistik artefaktlar — asosida moslashtirilgan. Birinchi komponent 'Forensic Science International: Digital Investigation' jurnali va uning 'Digital Investigation' jurnalidan yuklab olingan jami 1082 ta tadqiqot maqolasidan iborat bo'lib, bu maqolalar turli DF mavzularini qamrab oladi [2]. Ikkinchi komponent sifatida Grajeda va boshqalar tomonidan ishlab chiqilgan Artifact Genome Project (AGP) — kriminalistik artefaktlarning saralangan repozitoriysi (CUFAs) — ishlatilgan bo'lib, tadqiqot vaqtida AGP jami 1390 ta DF artefaktini o'z ichiga olgan (shundan 1126 tasi fayl turidagi, 211 tasi Windows Registry, 19 tasi xotira, 19 tasi tarmoq paketi va 15 tasi boshqa turdagi artefaktlar), har bir artefakt sarlavha, tur, qurilma, joylashuv yo'li, tavsif, izohlar, qidiruv teglari va tegishli ma'lumot maydonlarini o'z ichiga oladi [2; 12]. Metodologik jarayon bir necha bosqichdan iborat: dastlab yuklab olingan PDF fayllardan annotatsiya va minnatdorchilik bo'limlari orasidagi matn mazmuni pdftotext kutubxonasi yordamida ajratib olingan, adabiyotlar ro'yxati esa vektor qidiruvida tegishsiz mosliklarni oldini olish uchun chiqarib tashlangan; so'ngra matn LangChain'ning RecursiveCharacterTextSplitter vositasi yordamida taxminan 2000 belgidan iborat kichikroq bo'laklarga bo'lingan, bunda bo'lish jarayoni ikki qatorli bo'shliqlarga, keyin nuqta-qator ketma-ketligiga, so'ngra bitta qator bo'shlig'iga va nihoyat nuqtalarga ustuvorlik bergan, 600 belgidan kam bo'lgan bo'laklar (odatda sarlavha yoki footnote kabi tegishsiz ma'lumotlar) esa filtrlab tashlangan [2]. Har bir maqolaning muallif va sarlavha ma'lumotlari bo'laklarga metama'lumot sifatida qo'shilgan, so'ngra bo'laklar Massive Text Embedding Benchmark (MTEB) natijalariga asoslanib yuqori samaradorligi va nisbatan past xotira sarfi tufayli tanlangan UAE_Large-V1 joylashtirish modeli orqali qayta ishlanib, ularning vektor ko'rinishlari va tegishli metama'lumotlar Chroma DB vektor bazasida saqlangan [2].

RAFT (Retrieval Augmented Fine-tuning) uchun zarur bo'lgan belgilangan ma'lumotlar to'plamini yaratish maqsadida, DF sohasida bunday tayyor to'plam mavjud emasligi va uni qo'lda yaratish amalda imkonsiz bo'lgani sababli, GPT-4 Turbo modelidan foydalanilgan holda har bir tadqiqot maqolasining to'liq matni va maxsus ishlab chiqilgan ko'rsatma (dolzarblik, matn tilining asl maqoladagidek saqlanishi, JSON formatidagi chiqish tuzilmasi, savollarda maqola sarlavhasi yoki mualliflariga bevosita ishora qilmaslik, javoblarda APA formatidagi iqtibos keltirish majburiyati va javoblarning batafsil bo'lishi kabi talablarni o'z ichiga olgan) asosida savol-javob (Q&A) juftliklari generatsiya qilingan, bu jarayon 1082 ta maqoladan taxminan 10 000 ta Q&A juftligini yaratgan, so'ngra qo'lda tekshiruv orqali taxminiy 10 foizi 'yuqoridagi tadqiqotda' kabi tashqi hujjatga ishora qiluvchi nomaqbul iboralarga ega ekanligi aniqlanib, filtrlab tashlangan [2]. Ushbu Q&A juftliklarini generatsiya qilish uchun ishlatilgan ko'rsatma DF sohasida 19 yildan ortiq tajribaga ega mualliflardan birining ekspert bilimi yordamida ishlab chiqilgan bo'lib, savollar raqamli artefaktlar va ularning joylashuvi (tarmoq, disk, xotira), metama'lumot elementlari (vaqt tamg'alari va identifikatorlar), kriminalistik jarayonlar (olish, autentifikatsiya qilish), vositalar, topilmalar va kriminalistik puxtalik kabi mavzularni qamrab olishi, savollarning o'zi esa hujjatning sarlavhasi yoki mualliflariga bevosita ishora qilmasligi talab qilingan [2].

Har bir savol uchun Chroma DB'dan eng o'xshash o'nta bo'lak qidirib topilgan va yakuniy ma'lumotlar to'plami Alpaca formatida (ko'rsatma, kirish — qidirib topilgan bo'laklar va savolning o'zi, chiqish — javob va manba ma'lumoti) tuzilgan, so'ngra 75/25 nisbatida trenirovka (6739 namuna) va test (2244 namuna) to'plamlariga bo'lingan [2]. Moslashtirish jarayonining o'zi Axolotl vositasi yordamida amalga oshirilgan, LLaMA-3.1-8B modeli bitsandbytes kutubxonasi orqali 4-bitli butun sonlarga kvantizatsiya qilingan va QLoRA usuli (daraja — 32, alfa — 32, dropout — 0,05, Paged

© 2026 International Conference on Multidisciplinary Sciences and Innovations . All Rights Reserved.

Adam 32-bit optimallashtiruvchisi, 0,0002 dan boshlanuvchi kosinus o'quv tezligi jadvali, 100 ta isinish qadami, jami 4 epoxa, 2-epoxada erta to'xtatish) qo'llanilgan holda o'tkazilgan [2]. Baholash bosqichida uchta model — sof LLaMA-3.1-8B, RAG bilan mustahkamlangan LLaMA-3.1-8B va ForensicLLM — uchta metrika (BERTScore, BAAI'ning BGE-M3 va GPT-4o asosidagi G-Eval) bo'yicha solishtirilgan, natijalar shuni ko'rsatdi: ForensicLLM BERTScore F1 bo'yicha 0,9232 ballga (bazaviy modelning 0,8872 va RAG modelining 0,8923 balliga nisbatan yuqori), BGE-M3 kosinus o'xshashligi bo'yicha 0,9091 ballga (0,8623 va 0,8805 ga nisbatan) hamda G-Eval umumiy ko'rsatkichi bo'yicha 2,7544 ballga (2,3787 va 2,6329 ga nisbatan) erishgan, bunda ForensicLLM'ning javoblari o'rtacha 156,34 token uzunlikda bo'lib, bu bazaviy modelning 128,94 va RAG modelining 226,38 token uzunligi o'rtasida joylashgan, ya'ni model bir vaqtning o'zida ixcham va semantik jihatdan aniq javoblar generatsiya qila olgan [2]. Manba ko'rsatish (source attribution) bo'yicha o'tkazilgan tahlilda 2244 ta javobning 2243 tasi iqtibos o'z ichiga olgan, shulardan 1823 tasida (81,2 foiz) iqtibos ham sarlavha, ham mualliflarga to'g'ri kelgan, 119 tasida faqat mualliflarga to'g'ri kelgan, natijada model 86,6 foiz hollarda manbani (sarlavha yoki muallif orqali) to'g'ri ko'rsatgan, biroq 300 ta javobda (13,4 foiz) iqtibos berilgan kontekstdagi na sarlavha, na mualliflarga mos kelmagan, bu esa hatto RAG bilan mustahkamlangan moslashtirilgan modelda ham gallyutsinatsiya xavfi saqlanib qolishini ko'rsatadi [2]. Ushbu ko'p metrikali, miqdoriy jihatdan qat'iy baholash yondashuvi Rossiya ilmiy adabiyotida chuqur o'qitishga asoslangan kiberxavfsizlik yechimlarini baholashda ham xuddi shunday ko'p mezonli empirik tekshiruv zarurligi alohida ta'kidlanganligi bilan hamohang keladi, bu esa AI-asosidagi xavfsizlik vositalariga ishonch faqat qat'iy va ko'p qirrali sinovdan keyin shakllanishi lozimligini yana bir bor tasdiqlaydi [9]. Shuni alohida ta'kidlash lozimki, ForensicLLM'ning o'rtacha ixcham javob uzunligi (156,34 token) uni faqat yuqori aniqlik nuqtai nazaridangina emas, balki amaliy foydalanish nuqtai nazaridan ham RAG modelidan farqlantiradi, chunki tergov sharoitida ixcham va aniq javob ko'pincha keraksiz batafsillikdan ko'ra afzalroq hisoblanadi, ayniqsa vaqt tanqisligi yuqori bo'lgan dastlabki tergov bosqichlarida.

3. Foydalanuvchi tadqiqoti natijalari, model cheklovlari va ForensicLLM'ning xalqaro-huquqiy kontekstdagi ahamiyati

ForensicLLM'ning amaliy foydaliligini haqiqiy DF mutaxassislari nuqtai nazaridan baholash maqsadida, mualliflar Flask asosida qurilgan va Microsoft Azure'da joylashtirilgan veb-ilova orqali foydalanuvchi tadqiqotini o'tkazgan, bunda ishtirokchilar dastlab demografik va kasbiy ma'lumot beruvchi so'rovnomani to'ldirgan, so'ngra ikkita farazий ssenariy — bank talon-tarojligi va yo'qolgan shaxsni qidirish ishi — bo'yicha savollar berib, uchta model konfiguratsiyasi (LLaMA-3.1-8B, LLaMA-3.1-8B+RAG va ForensicLLM) javoblarini foydalilik, to'g'rilik, dolzarblik, iqtibos aniqligi va tushunishni yaxshilash mezonlari bo'yicha baholagan [2]. Jami 32 ishtirokchi kamida bitta so'rov taqdim etgan, 23 tasi tegishsiz (masalan, tasodifiy yoki modelni 'buzishga' urinuvchi) so'rov sifatida filtrlab tashlangandan so'ng 157 ta yaroqli so'rov qolgan, ishtirokchilar orasida 11 tadqiqotchi, 10 DF eksperti va boshqa turli kasb egalari bo'lgan [2]. Yig'ilgan 157 ta savolning mavzu taqsimoti shuni ko'rsatdiki, eng ko'p so'ralgan mavzular artefaktlar va dalillar (80 ta savol), keyin vositalar (38 ta), olish jarayoni (33 ta) va protsedura (33 ta) bo'lgan, bu esa amaliyotchilarning kundalik DF ishida aynan shu yo'nalishlarga eng ko'p ehtiyoj sezishini ko'rsatadi [2]. Og'irlangan ballar tahlili shuni ko'rsatdiki, ham RAG, ham ForensicLLM modellari bazaviy LLaMA modelidan barcha mezonlar bo'yicha sezilarli darajada ustun bo'lgan, bunda ForensicLLM ayniqsa 'dolzarblik' (108 ball) va 'to'g'rilik' (83 ball) mezonlari bo'yicha yuqori natija ko'rsatgan, RAG modeli esa ko'proq batafsil javoblari tufayli 'tushunishni yaxshilash' mezonida biroz ustunlikka ega bo'lgan, bazaviy LLaMA modeli esa 'foydalilik' va 'iqtibos' mezonlarida manfiy ball olgan [2].

© 2026 International Conference on Multidisciplinary Sciences and Innovations . All Rights Reserved.

Statistik tahlil natijalari DF tajribasi ko'proq bo'lgan ishtirokchilarning javoblarni ancha qattiqroq baholaganini ko'rsatdi — xi-kvadrat testi dolzarblik ($\chi^2 = 11,02$; $p = 0,026$), iqtibos ($\chi^2 = 10,15$; $p = 0,038$) va tushunishni yaxshilash ($\chi^2 = 11,52$; $p = 0,021$) mezonlari bo'yicha statistik ahamiyatli farqlarni aniqladi, bu esa tajribali mutaxassislarning texnik tafsilot va aniqlikka nisbatan yuqoriroq talab qo'yishi bilan izohlanadi [2]. Sifat jihatidan yig'ilgan fikr-mulohazalarda o'n uch ishtirokchi javoblarni haddan tashqari umumiy va tafsilotsiz deb topgan bo'lsa, uch nafari aksincha, javoblarni haddan tashqari texnik deb baholagan, bu esa turli foydalanuvchi talablariga mos ravishda javob darajasini moslashtirish zarurligini ko'rsatuvchi qiziqarli muammoni ochib beradi; shu bilan birga, iqtibos xususiyati beshta ishtirokchi tomonidan ijobiy baholangan bo'lsa-da, o'n holatda ishtirokchilar keltirilgan maqolalarni internetda topa olmagan, bu esa keltirilgan manbalarga to'g'ridan-to'g'ri havolalar qo'shish zarurligini ko'rsatadi [2]. Mualliflar tomonidan qayd etilgan asosiy cheklov shundan iboratki, hatto to'g'ri iqtibos keltirilgan taqdirda ham model manba maqoladagi xatolarni o'zlashtirib olishi mumkin — buni bir ishtirokchining 'taqdim etilgan ma'lumot texnik jihatdan noto'g'ri, ammo bu LLM'ning aybi emas, chunki u iqtibos keltirgan tadqiqot maqolasining o'zida vositalar haqida noto'g'ri tavsif berilgan edi' degan izohi yaqqol tasdiqlaydi [2]; bundan tashqari, GPT-4'ning chiqish tokenlari 4096 bilan cheklanganligi sababli har bir maqoladan atigi 10 ta Q&A juftligi generatsiya qilingan, shuningdek AGP ma'lumotlar to'plamidan faqat metama'lumotlar ishlatilgan, log yoki konfiguratsiya fayllari kabi chuqurroq artefakt ma'lumotlari esa kelgusi tadqiqotlar uchun qoldirilgan [2]. Ushbu natijalar ForensicLLM'ning Daubert mezoniga muvofiqligini alohida yorqin ko'rsatadi: model faqat tengdoshlar tomonidan taqrizdan o'tgan DF adabiyoti asosida o'qitilgani va uning iqtiboslari tekshiriluvchi bo'lgani sababli, u Daubert mezonining 'tengdoshlar tomonidan taqrizdan o'tganlik' talabini bevosita, 'ma'lum xatolik darajasi' talabini esa 86,6 foizlik aniqlik va 13,4 foizlik gallyutsinatsiya ko'rsatkichi orqali qisman qondiradi — bu shaffoflik darajasi manbalari tekshirib bo'lmaydigan yopiq bulutli umumiy modellarda mavjud emas [2].

Xalqaro miqyosda parallel sa'y-harakatlarni ko'rib chiqsak, Irlandiyalik tadqiqotchilar Wickramasekara, Breitinger va Scanlon tomonidan tayyorlangan kengroq sharh ham xuddi shunday domenga bog'lanish va shaffoflik zarurligini ta'kidlaydi [4], Singapurda nashr etilgan Chaudhary, Iyengar va Saidakhrarovich tomonidan yozilgan bobda esa LLM'larni DF jarayoniga integratsiyalashuvi mintaqalararo tadqiqot yo'nalishi sifatida, ehtiyotkor baholash freymvorklarini talab qiluvchi masala sifatida ko'rsatiladi [5].

O'zbekiston kontekstida esa, 'Kiberxavfsizlik to'g'risida'gi O'RQ-764-son Qonunga muvofiq Davlat xavfsizlik xizmati kiberxavfsizlik sohasidagi vakolatli davlat organi sifatida belgilangan bo'lib [1], ForensicLLM'ning tengdoshlar tomonidan taqrizdan o'tgan manbalarga asoslangan, tekshiriluvchi arxitekturasi kelgusida o'zbek tilidagi domenga xos DF modeli yaratilishi uchun uzatiladigan (transferable) loyihalash namunasi bo'lib xizmat qilishi mumkin, bu fikr O'zbekistonda kompyuter va raqamli kriminalistika sohasida qonunchilikni modernizatsiya qilish hamda ilmiy-texnik salohiyatni oshirish zarurligini ta'kidlovchi milliy tadqiqotlar bilan ham uyg'un keladi [10]. Yakuniy umumlashtirish sifatida shuni ta'kidlash mumkinki, ForensicLLM tajribasi shuni ko'rsatadiki, DF sohasida sun'iy intellektni joriy etishning eng istiqbolli yo'nalishi katta, umumiy maqsadli bulutli modellarni ko'r-ko'rona qo'llash emas, balki kichikroq, ma'lumotlar suvereniteti va manba shaffofligini ta'minlovchi, tor domenga chuqur moslashtirilgan modellarni ishlab chiqishdir, bu yondashuv nafaqat texnik samaradorlikni, balki huquqiy-institutsional ishonchni ham bir vaqtning o'zida ta'minlash imkonini beradi.

Xulosa

Mazkur maqolada ForensicLLM — raqamli kriminalistika uchun ishlab chiqilgan birinchi ixtisoslashtirilgan mahalliy katta til modelining nazariy-texnik asoslari, ma'lumotlar to'plami, RAFT

© 2026 International Conference on Multidisciplinary Sciences and Innovations . All Rights Reserved.

metodologiyasi hamda miqdoriy va sifat jihatidan baholash natijalari chuqur tahlil qilindi. O'tkazilgan tahlil shuni ko'rsatdiki, LoRA/QLoRA kabi resurstejamkor moslashtirish usullari va RAFT yondashuvi orqali kichik, ochiq manbali modellarni domenga xos ma'lumotlar asosida sozlash orqali bulutli umumiy maqsadli modellardan yuqori samaradorlikka erishish mumkin, bunda ForensicLLM barcha uchta baholash metrikasi (BERTScore, BGE-M3, G-Eval) bo'yicha ustunlik ko'rsatgan va 86,6 foizlik manba ko'rsatish aniqligiga erishgan. Shu bilan birga, tahlil shuni ham ko'rsatdiki, hatto ushbu ilg'or yondashuv ham gallyutsinatsiya xavfidan (13,4 foiz noto'g'ri iqtibos) va manba maqoladagi xatolarni o'zlashtirib olish muammosidan butunlay xoli emas, bu esa har qanday AI-yordamchi kriminalistik vositaning natijalari inson tomonidan tekshirilishi zarurligini yana bir bor tasdiqlaydi. ForensicLLM'ning taqrizdan o'tgan, tekshiriluvchi manbalarga asoslangan arxitekturasi uni AQSh Daubert mezoniga eng yaqin keluvchi AI-kriminalistika yechimlaridan biriga aylantiradi. Xalqaro (Yevropa, Rossiya, Singapur) tajribani O'zbekiston milliy konteksti bilan solishtirish shuni ko'rsatdiki, ushbu loyihalash tamoyillari — ya'ni faqat tekshirilgan ilmiy manbalar asosida o'qitish va iqtiboslarni tekshiriluvchi qilish — kelgusida o'zbek tilidagi domenga xos DF modellarini yaratishda ham asosiy yo'l-yo'riq bo'lib xizmat qilishi mumkin.

Foydalanilgan adabiyotlar

I. Qonunchilik hujjatlari

[1] O'zbekiston Respublikasining "Kiberxavfsizlik to'g'risida"gi Qonuni, O'RQ-764-son, 2022-yil 15-aprel (2022-yil 17-iyuldan kuchga kirgan). Manba: O'zbekiston Respublikasi qonun hujjatlari milliy ma'lumotlar bazasi — lex.uz. Havola: https://lex.uz/uz/docs/-5960604

II. Ilmiy maqolalar

[2] Sharma B., Ghawaly J., McCleary K., Webb A.M., Baggili I. ForensicLLM: A local large language model for digital forensics // Forensic Science International: Digital Investigation. — 2025. — Vol. 52, Art. 301872. DOI: 10.1016/j.fsidi.2025.301872. URL: https://www.sciencedirect.com/science/article/pii/S2666281725000113

[3] Scanlon M., Breitinger F., Hargreaves C., Hilgert J.-N., Sheppard J. ChatGPT for digital forensic investigation: the good, the bad, and the unknown // Forensic Science International: Digital Investigation. — 2023. — Vol. 46, Art. 301609.

[4] Wickramasekara A., Breitinger F., Scanlon M. Exploring the Potential of Large Language Models for Improving Digital Forensic Investigation Efficiency // arXiv preprint. — 2024. arXiv:2402.19366. URL: https://arxiv.org/pdf/2402.19366

[5] Chaudhary N.K., Iyengar S.S., Saidakhrarovich G.S. Large Language Models for Digital Forensic Investigations // Information Security, Privacy and Digital Forensics. ICISPD 2024. Lecture Notes in Electrical Engineering, vol. 1456. — Singapore: Springer, 2026. DOI: 10.1007/978-981-95-2196-8_18

[6] Hu E.J., Shen Y., Wallis P., Allen-Zhu Z., Li Y., Wang S., Wang L., Chen W. LoRA: Low-Rank Adaptation of Large Language Models // arXiv preprint. — 2021. arXiv:2106.09685.

[7] Dettmers T., Pagnoni A., Holtzman A., Zettlemoyer L. QLoRA: Efficient Finetuning of Quantized LLMs // Advances in Neural Information Processing Systems. — 2024. — Vol. 36.

[8] Zhang T., Patil S.G., Jain N., Shen S., Zaharia M., Stoica I., Gonzalez J.E. RAFT: Adapting Language Model to Domain Specific RAG // arXiv preprint. — 2024. arXiv:2403.10131.

[9] Использование больших языковых моделей для поиска угроз кибербезопасности на основе методов глубокого обучения: анализ современных исследований // КиберЛенинка (электронная научная библиотека). URL: https://cyberleninka.ru/article/n/ispolzovanie-bolshih-yazykovyh-modeley-dlya-poiska-ugroz-kiberbezopasnosti-na-osnove-metodov-glubokogo-obucheniya-analiz

© 2026 International Conference on Multidisciplinary Sciences and Innovations . All Rights Reserved.

[10] Kompyuter va raqamli kriminalistika sohasida xalqaro hamkorlik va qonunchilikni modernizatsiya qilish masalalari // International Journal of Science and Technology (O'zbekiston). — 2026. — B. 3–8. Havola: https://science-technology.uz/index.php/journal/article/download/695/651/1256

III. Ilmiy tezislar va dissertatsiyalar

[11] Sharma B. ForensicLLM: A Local Large Language Model for Digital Forensics // LSU Master's Theses, No. 6059. — Louisiana State University, 2024. URL: https://repository.lsu.edu/gradschool_theses/6059/

[12] Grajeda C., Sanchez L., Baggili I., Clark D., Breitinger F. Experience constructing the Artifact Genome Project (AGP): managing the domain's knowledge one artifact at a time // Digital Investigation. — 2018. — Vol. 26, S47–S58.