ҚАЗАҚСТАН РЕСПУБЛИКАСЫ БАС ПРОКУРАТУРАСЫНЫҢ ЖАНЫНДАҒЫ
ҚҰҚЫҚ ҚОРҒАУ ОРГАНДАРЫ АКАДЕМИЯСЫ

ЖОҒАРЫ ОҚУ ОРНЫНАН КЕЙІНГІ БІЛІМ БЕРУ ИНСТИТУТЫ

АМИРАЛИЕВ НҰРТАС ҚАСЕНҰЛЫ

# Ішкі істер органдары тергеушісінің қызметінде қолжазба мәтіндерін тану технологияларын қолданудың перспективалары мен іс жүргізу шектеулері

7М12301 - «Құқық қорғау қызметі» (бейінді бағыт) білім беру бағдарламасы бойынша ұлттық қауіпсіздік және әскери іс магистрі дәрежесін алуға жоба

Ғылыми жетекші:
арнайы заң пәндері
кафедрасының
аға оқытушысы
М.Б. Садыков,
философия докторы (PhD),
әділет кеңесшісі
_____

Қосшы қ., 2026 ж.

## ТҮЙІНДЕМЕ

Жоба ішкі істер органдары тергеушісінің қызметінде қолжазба мәтінін тану технологияларын қолданудың бағыттарын және іс жүргізушілік шектеулерін зерттейді. Нәтижесінде бұл технологиялардың құжаттарды цифрландыру, іздеу, жүйелеу үшін жарамды екені, бірақ олардың нәтижелері тексерілген көмекші ақпарат ретінде ғана пайдаланылатыны негізделді.

## РЕЗЮМЕ

Проект исследует направления применения технологий распознавания рукописного текста в деятельности следователя органов внутренних дел и процессуальные ограничения их использования. Обосновано, что эти технологии пригодны для цифровизации, поиска и систематизации документов, однако их результаты могут использоваться только как проверенная вспомогательная информация.

## SUMMARY

The project examines the use of handwritten text recognition technologies in the work of investigators of internal affairs bodies and the procedural limits of such use. It substantiates that these technologies are suitable for digitizing, searching, and systematizing documents, but their results may be used only as verified auxiliary information.

МАЗМҰНЫ

АНЫҚТАМАЛАР
БЕЛГІЛЕР МЕН ҚЫСҚАРТУЛАР
КІРІСПЕ

1 ҚОЛЖАЗБА МӘТІНДЕРІН ТАНУ ТЕХНОЛОГИЯЛАРЫН ІШКІ ІСТЕР ОРГАНДАРЫ ТЕРГЕУШІСІНІҢ ҚЫЗМЕТІНДЕ ҚОЛДАНУДЫҢ ПЕРСПЕКТИВАЛАРЫ

1.1 Тергеуші қызметіндегі қолжазба құжаттардың маңызы және оларды өңдеудің практикалық мәселелері ..... 11
1.2 Қолжазба мәтіндерін тану технологияларының ұғымы, ерекшеліктері және қолданбалы мүмкіндіктері ..... 15
1.3 Ішкі істер органдары тергеушісінің ұйымдастырушылық-техникалық қызметінде қолжазба мәтіндерін тану технологияларын қолданудың негізгі бағыттары және олардың тергеу қызметін ұйымдастыруға ықпалы ..... 18

2 ҚОЛЖАЗБА МӘТІНДЕРІН ТАНУ НӘТИЖЕЛЕРІН ПАЙДАЛАНУДЫҢ ІС ЖҮРГІЗУ ШЕКТЕУЛЕРІ ЖӘНЕ ОЛАРДЫ ТЕКСЕРУ МЕН РАСТАУ ТӘРТІБІ

2.1 Қолжазба мәтіндерін тану нәтижелерінің қылмыстық процестегі құқықтық табиғаты ..... 23
2.2 Қолжазба мәтіндерін тану нәтижелерін пайдаланудың іс жүргізушілік талаптары мен шектеулері ..... 25
2.3 Қолжазба мәтіндерін тану нәтижелерін тексеру мен пайдаланудың практикалық тәртібі ..... 28

ҚОРЫТЫНДЫ ..... 32

ПАЙДАЛАНЫЛҒАН ДЕРЕККӨЗДЕРДІҢ ТІЗІМІ ..... 34

ҚОСЫМША 1 ..... 37
ҚОСЫМША 2 ..... 38
ҚОСЫМША 3 ..... 39
ҚОСЫМША 4 ..... 47

# АНЫҚТАМАЛАР

Осы жобада сәйкесінше анықтамалары бар келесі терминдер қолданылған:

Қолжазба құжат – қолмен жазылған мәтіні бар құжат; қылмыстық іс материалдарында ақпараттық немесе дәлелдемелік мәнге ие болуы мүмкін.

Қолжазба мәтінін тану технологиясы – қолмен жазылған мәтінді цифрлық мәтінге айналдыруға арналған бағдарламалық тәсілдер мен техникалық құралдардың жиынтығы.

Түпнұсқа құжат – салыстыру, тексеру және мазмұнын нақтылау үшін негізгі бастапқы материал болатын қолжазба құжат.

Электрондық қылмыстық іс – қылмыстық сот ісін электрондық форматта жүргізу кезінде қалыптастырылатын іс жүргізу материалдарының электрондық жиынтығы.

Көмекші ақпараттық ресурс – дербес дәлелдеме көзі емес, бірақ құжатты іздеуге, жүйелеуге және тексеруге жәрдемдесетін ақпарат.

Растау – танылған мәтінді түпнұсқамен салыстыру, күмәнді бөліктерді қайта тексеру және қажет болғанда арнайы білім арқылы нақтылау.

Іс жүргізушілік шектеулер – қылмыстық процесте қолжазба мәтінін тану нәтижелерін пайдалануға қойылатын құқықтық талаптар мен шектер.

# БЕЛГІЛЕР МЕН ҚЫСҚАРТУЛАР

ҚР – Қазақстан Республикасы
ҚПК – Қазақстан Республикасының Қылмыстық-процестік кодексі
ҚР ІІМ – Қазақстан Республикасының Ішкі істер министрлігі
HTR – қолжазба мәтінін тану технологиясы (Handwritten Text Recognition)
OCR – таңбаларды оптикалық тану (Optical Character Recognition)
KOHTD – қазақ тіліндегі желіден тыс қолжазба мәтіндерінің деректер жиыны (Kazakh Offline Handwritten Text Dataset)
HKR – қазақ және орыс тілдеріндегі қолжазба мәтіндерінің деректер қоры (Handwritten Kazakh & Russian Database)
ENFSI – Сот сараптамасы институттарының еуропалық желісі (European Network of Forensic Science Institutes)
NIST – АҚШ Сауда департаментінің Ұлттық стандарттар және технологиялар институты (National Institute of Standards and Technology)
READ – мұрағаттық құжаттарды тану және байыту жобасы (Recognition and Enrichment of Archival Documents)
READ-COOP – мұрағаттық құжаттарды тану және байыту жөніндегі кооператив (READ Cooperative)

# КІРІСПЕ

Жүргізілетін зерттеудің өзектілігі. Цифрлық трансформация Қазақстан үшін мемлекеттік саясаттың негізгі бағыттарының біріне айналды. Президент Қ.-Ж. Тоқаевтың 2025 жылғы Жолдауында елді үш жылдың ішінде жаппай цифрлық мемлекетке көшіру міндеті қойылды [1]. Бұл бағыт 2026 жылғы 6 қаңтардағы Президент Жарлығымен де нақтыланып, 2026 жыл Цифрландыру және жасанды интеллект жылы деп жарияланды [2]. Ресми деректер де осы үрдістің ауқымын көрсетеді: мемлекеттік қызметтердің шамамен 90%-ы онлайн форматта көрсетіледі, eGov.kz порталында 15,1 миллионнан астам тіркелген пайдаланушы бар, ал eGov Mobile қосымшасының күнделікті аудиториясы 671 мыңнан асады [3]. 2026 жылғы 15 наурыздағы республикалық референдумда қабылданған Қазақстан Республикасының Конституциясы бұл өзгерістерге жаңартылған құқықтық негіз қалады [4]. Құжат айналымының да, мемлекеттік рәсімдердің де көп бөлігі бүгінде цифрлық кеңістікке көшті.

Бірақ цифрландыру қағаз құжаттарды мүлде ығыстырып тастаған жоқ. Мұны ішкі істер органдары тергеушісінің күнделікті жұмысынан анық көруге болады. Бас прокуратураның Құқықтық статистика және арнайы есепке алу жөніндегі комитетінің деректері бойынша, 2025 жылы электрондық форматта 130 594 қылмыстық іс тергелген, бұл сотқа дейінгі тергеудегі істердің шамамен 90 пайызын құрайды [5]. Сонымен бірге қылмыстық процеске жыл сайын түсетін жазбаша материал көлемі азаймай, қайта өсіп келеді: 2023 жылғы 2,45 миллионнан 2025 жылы 2,71 миллионға жетті [6–8]. Жалпы ағын осындай. Ал жеке және заңды тұлғалардың жазбаша өтініштері бөлек алғанда біршама кеміген: 2023 жылы 271 мыңнан 2025 жылы 226 мыңға дейін [6–8]. Пайыздық үлесі азайғанымен, жүз мыңдаған өтініш әлі де қағаз түрінде келіп жатыр.

Мұнда бір нәрсені ескерген жөн. Есептерде материалдардың нақты қаншасы қолжазба екені бөлек көрсетілмеген, сондықтан жоғарыдағы сандар тергеушінің қолжазбамен жұмысының нақты көлемін емес, ең төменгі шамасын ғана көрсетеді.

Қолжазба мәтіндерін тану технологиялары шетелдік ғылыми әдебиетте кеңінен зерттелген. Ruiz-Parrado және әріптестері 1990–2020 жылдар аралығындағы желіден тыс қолжазба құжаттарды талдауға арналған 5389 жарияланымды саралап, бұл саладағы негізгі міндеттер ретінде қолжазба мәтінін тануды, қолтаңбаны тексеруді, авторды айқындауды, мәтін ішінен сөз іздеуді және ақпараттық іздеуді көрсетеді [9]. Nockels және әріптестері Transkribus платформасына қатысты 2015–2020 жылдардағы 381 жарияланымды талдап, олардың 67%-ы гуманитарлық қолданбаларға, 25%-ы технологиялық зерттеулерге, 5%-ы пайдаланушы тәжірибесіне, 3%-ы әдістемелік мәселелерге арналғанын анықтаған [10]. Бұл деректер бір жағынан қолжазба құжаттарды цифрлық талдау саласының кеңдігін, екінші жағынан нақты құралдардың ғылыми және қолданбалы айналымда орныға бастағанын көрсетеді.

Қазақстандық зерттеулерде де белгілі бір эмпирикалық негіз қалыптасып келеді. Тойғанбаева және әріптестері қазақ тіліндегі желіден тыс қолжазба мәтіндерінің KOHTD деректер жиынын 3000 студенттік емтихан парағы, 140 335 сегменттелген кескін және шамамен 922 010 таңба негізінде құрастырған; жауаптардың 99%-ы қазақ тілінде болған [11]. Нурсеитов және әріптестерінің HKR базасында 1400-ден астам толтырылған нысан, шамамен 63 000 сөйлем, 715 699 таңба және 200-ге жуық жазушы қамтылған; мәтіндердің 95%-ы орыс тілінде, 5%-ы қазақ тілінде [12]. Екі база да құнды әрі бір-бірін толықтырады: KOHTD еркін емтихан жауаптарын жинаса, HKR құрылымы шектеулі формалық мәтіндерден тұрады. Алайда тергеушінің нақты қызметіне қатысты дереккөз екеуінде де жоқ. Іс жүргізу құжаттары, арыздар, түсініктемелер мен қылмыстық процестегі өзге қолжазба материалдар бойынша ашық эмпирикалық тексеру ғылыми әдебиетте де, ресми дереккөздерде де жүргізілмеген. Зерттеудің өзектілігі осы олқылықтан туындайды.

Шешілетін практикалық міндеттің қазіргі уақыттағы жағдайын бағалау. Қолжазба мәтіндерін тану тергеуші қызметінде әлі толық шешілмеген. Техникалық база жеткілікті дамыған. Мәселе оны қылмыстық іс жүргізу тәжірибесіне бейімдеуде: бұл жұмыс әзірге аяқталмай тұр.

АҚШ Ұлттық стандарттар және технологиялар институтының 2020 жылғы есебінде қолжазбаны сот-салыстырмалы зерттеу күмәнді құжаттар саласына жатқызылады [13]. Еуропалық сот сараптамасы желісінің 2022 жылғы нұсқаулығы қолжазба құжаттарды сараптамада жиі кездесетін объект ретінде атайды [14]. Сала ретінде қолжазбамен жұмыс баяғыда қалыптасқан. Ал оны қалай ұйымдастыру, нәтижесін қалай бағалау керектігі нақты рәсімдік бақылауды талап етеді.

Техникалық тосқауылдар да белгілі. 5389 жарияланымды талдау екі негізгі түйінді көрсетеді: біріншісі – жазу мәнерлері мен тілдердің алуан түрлілігі, екіншісі – статикалық кескінде жазу жылдамдығы, штрихтардың орындалу реті сияқты динамикалық белгілердің жоғалуы [9]. Еуропалық нұсқаулық жұмыс сапасын жарық, кеңістік және хаттамалау арқылы бағалайды [14], ал американдық есеп адами факторлар мен стандартталмаған жағдайларды негізгі тәуекелдердің қатарына қосады [13]. Демек, рәсімдік бақылау болмаса, технология нәтижесіне толық сенім артуға болмайды.

Технологияның таралу деңгейі сандар арқылы да байқалады. 381 жарияланымның 66,67%-ы гуманитарлық қолданбаларға, 25,46%-ы технологиялық зерттеулерге жатады [10]. Transkribus платформасының тұрақты айлық аудиториясы 1700 пайдаланушы шамасында [10]. Бұл қолжазба мәтіндерін тану технологияларының зертханалық деңгейден шыққанын көрсетеді. Алайда олар қылмыстық іс жүргізу тәжірибесіне толық орныгып үлгерген жоқ.

Тәжірибелік мысалдардың сипаты да әртүрлі. Bias in History жобасы бір айда 100 000 бет тарихи полиция есебін өңдеген [15]. Қазақстанда Тойганбаева және әріптестері әзірлеген KOHTD деректер жиыны 3000 емтихан парағын, 140 335 сегменттелген кескінді және шамамен 922 010 таңбаны қамтиды [11]. Бірінші

мысал архивтік тарихи-құқықтық материалдарға бағытталса, екіншісі оқу ортасындағы еркін қолжазба жауаптарына негізделген. Екеуі де өз контекстінде маңызды, бірақ тергеушінің күнделікті іс жүргізу құжаттарын тікелей қамтымайды.

Қазақ-орыс аралас материалдарға арналған база да бар. Нурсеитов және әріптестерінің HKR қорында мәтіндердің 95%-ы орыс тілінде, 5%-ы қазақ тілінде берілген; база 1400-ден астам толтырылған нысанды, 715 699-дан астам таңбаны қамтиды [12]. Бұл қор пошта, мекенжай, кілт сөздер сияқты құрылымы шектеулі нысандарды тануға бейімделген. Ал тергеушінің еркін жазылған арыздарына, түсініктемелеріне немесе өзге де іс жүргізу материалдарына арналмаған.

Тану дәлдігі де біркелкі емес. Abdallah және әріптестерінің қазақ-орыс деректері бойынша жүргізген зерттеуінде бірінші сынақ жиынында таңба қатесі 4,13%, сөз қатесі 18,91%, сөйлем қатесі 25,72% деңгейінде көрсетілген; екінші жиында тиісінше 6,31%, 23,69%, 35,16% болған [16]. Сол зерттеуде ұсынылған үлгі Bluche үлгісінің 10,55% және Puigcerver үлгісінің 20,73% таңба қатесімен салыстырылған [16]. Бұл технологияның қолданбалы әлеуеті бар екенін көрсетеді. Бірақ оның дәлдігі әр жазу мәнері, әр тілдік орта және әр құжат түрі үшін бірдей емес.

Тақырыптың практикалық маңызын Бобовкин мен Ручкиннің 2018 жылғы зерттеуі де айқын көрсетеді. Олардың деректерінше, сот-қолжазбатану сараптамасына түсетін объектілердің үштен екісіне жуығы – қолтаңбалар, тағы үштен бірі – қысқа жазбалар. Сараптамалардың басым бөлігі, шамамен 78 пайызы, алдын ала тергеу сатысында жүргізіледі [17]. Яғни қолжазба объектісі тергеу тәжірибесінен еш жоғалған жоқ, қайта оның күнделікті бөлігі болып қалып отыр. Бұл ретте қаралған еңбектердің ешбірінде тергеушінің нақты іс жүргізу құжаттарына сүйенген ашық репрезентативті тану корпусы ұсынылмаған.

Жалпы көрініс мынадай. Технологиялық алғышарттар бар, эмпирикалық негіз ішінара жасалған. Ал шешуін күтіп тұрған нақты міндет – технологияны тергеу жағдайына бейімдеу. Бұған тану нәтижелерін тексерудің бірыңғай тәртібін орнату және оларды қылмыстық процесте қолданудың шегін айқындау мәселелері де қосылады.

Зерттеу мақсаты. Ішкі істер органдары тергеушісінің қызметінде қолжазба мәтіндерін тану технологияларын қолданудың ұйымдастырушылық-техникалық мүмкіндіктері мен қылмыстық процестегі іс жүргізушілік шектеулерін айқындау, сондай-ақ алынған нәтижелерді тексеру мен пайдаланудың практикалық тәртібін негіздеу.

Зерттеу міндеттері:

1. Ішкі істер органдары тергеушісінің қызметіндегі қолжазба құжаттардың орнын, олардың іс жүргізушілік маңызын және өңдеудің практикалық қиындықтарын анықтау.

2. Қолжазба мәтінін тану технологияларының ұғымын, ерекшеліктерін және қолданбалы мүмкіндіктерін ашу.

3. Тану технологияларын тергеушінің ұйымдастырушылық-техникалық қызметінде қолданудың негізгі бағыттарын белгілеу.

4. Тану нәтижелерінің қылмыстық процестегі құқықтық табиғатын және пайдаланудың іс жүргізушілік шектеулерін айқындау.

5. Тергеуші қызметінде тану нәтижелерін тексеру мен пайдаланудың практикалық тәртібін тұжырымдау.

Осы мақсат пен міндеттер екі негізгі практикалық ұсынымға алып келеді: біріншісі – тану технологияларын тергеушінің ұйымдастырушылық-техникалық қызметінде қолдану бағыттары, екіншісі – нәтижелерді қылмыстық процесте пайдаланудың көмекші сипаты мен міндетті тексеру тәртібі.

Зерттеу объектісі. Ішкі істер органдары тергеушісінің қолжазба құжаттармен жұмыс істеуіне байланысты қалыптасатын қоғамдық қатынастар.

Зерттеу пәні. Ішкі істер органдары тергеушісінің қызметінде қолжазба мәтіндерін тану технологияларын қолданудың құқықтық, ұйымдастырушылық-техникалық және іс жүргізушілік қырлары, сондай-ақ нәтижелерді тексеру мен пайдалану тәртібі.

Зерттеуді жүргізудің әдістері мен әдіснамалық негіздері. Жұмыста бірнеше әдіс үйлестірілді: формальды-құқықтық әдіс, жүйелік-құрылымдық және салыстырмалы талдау, ғылыми еңбектерді мазмұндық талдау, сондай-ақ құқық қолдану практикасын зерделеу.

Формальды-құқықтық әдіс Қылмыстық-процестік кодексті, қылмыстық сот ісін электрондық форматта жүргізу жөніндегі нұсқаулықты, ақпараттандыру мен дербес деректерді қорғау туралы заңнаманы талдауға негіз болды. Жүйелік-құрылымдық талдау қолжазба құжатпен жұмысты біртұтас процесс ретінде көруге мүмкіндік берді: құжат алдымен айқындалады, цифрланады, танылады, түпнұсқамен салыстырылады, тексеріледі және іс жүргізуде пайдаланылады. Шетелдік қолданбалы тәжірибені қазақстандық құқықтық реттеумен қатар қою үшін салыстырмалы әдіс пайдаланылды. Технологиялардың қолданбалы мүмкіндіктерін мазмұндық талдау арқылы, ал олардың тергеудегі нақты маңызын құқық қолдану практикасы арқылы бағаладық.

Жұмыстың әдіснамалық тірегі – құқықтық, ұйымдастырушылық-техникалық және іс жүргізушілік қырларды бір-бірінен бөлмей, өзара байланыста қарайтын кешенді тәсіл.

Ғылыми жаңалықтың негіздемесі. Жобада қолжазба мәтіндерін тану технологиясын тергеуші қызметінде қолданудың екі қыры зерттелді: оның ұйымдастырушылық-техникалық бағыттары мен қылмыстық процестегі іс жүргізушілік шектеулері.

Технология тергеуші үшін көмекші құрал ретінде көрсетілді: ол қолжазба құжаттарды цифрлауға, мәтін бойынша іздеуге, материалдарды жүйелеуге және электрондық есепке енгізуге жарайды. Ал тану нәтижесінің өзі дәлелдеме болып саналмайды. Бұл – тексеруді қажет ететін туынды ақпарат. Оның дұрыстығы түпнұсқамен салыстыру және қолмен тексеру арқылы, күрделі жағдайда маман не сарапшы қорытындысы арқылы бағаланады.

Жобаның жаңалығы дәл осы екі қырды бір тақырып аясында байланыстыра қарауда.

Практикалық ұсынымдар:

1. Қолжазба мәтіндерін тану технологиясы тергеушінің ұйымдастырушылық-техникалық жұмысын оңтайландырудың басым бағыттарының бірі. Жоба аясында оны қолдану моделі (қосымша 1) және технологияның тергеу жұмысын ұйымдастыруға ықпалын көрсететін кесте (қосымша 2) әзірленді.

2. Тану нәтижесі қылмыстық іс жүргізуде көмекші сипатқа ие және процессуалдық шешімнің дербес негізі бола алмайды; ол алдын ала тексеруден өтуі тиіс. Осы талап негізінде әзірленген алгоритм Ішкі істер министрлігі Тергеу департаментінің практикалық қызметіне енгізілді (қосымша 3, 4).

Апробация және нәтижелерді енгізу. Зерттеу нәтижелері ғылыми және практикалық деңгейде апробациядан өтті. Тақырып бойынша ғылыми мақала ішкі істер органдары тергеушісінің қызметінде жасанды интеллект негізіндегі қолжазба мәтіндерін тануды енгізудің процестік шектеулері мен дәлелдемелік мәртебесіне арналды. Әзірленген алгоритм Қазақстан Республикасы Ішкі істер министрлігінің Тергеу департаментінің практикалық қызметіне енгізілді. Бұл енгізу актісімен расталады.

# 1 Қолжазба мәтіндерін тану технологияларын ішкі істер органдары тергеушісінің қызметінде қолданудың перспективалары

## 1.1 Тергеуші қызметіндегі қолжазба құжаттардың маңызы және оларды өңдеудің практикалық мәселелері

Тергеуші үшін қолжазба құжат жай мәтін тасымалдағыш емес. Бұл сот-криминалистикалық зерттеуге жататын, іс жүргізу салмағы бар объект. АҚШ Сауда Департаментінің Ұлттық стандарттар және технологиялар институтының 2020 жылғы есебінде қолжазбаны сот-салыстырмалы зерттеу күмәнді құжаттар саласының құрамдас бөлігі ретінде сипатталған [13]. Сол есепте құжат материалдық түрде бекітілген коммуникация нысаны деп берілсе, күмәнді құжат түпнұсқалығы, шығу көзі немесе жасалу тәсілі тексерілуге жататын объект ретінде айқындалған [13]. Сот сараптамасы институттарының еуропалық желісінің 2022 жылғы нұсқаулығы әртүрлі құжаттардың сараптамалық тәжірибеде тұрақты түрде кездесетінін және оларды зерттеу қылмысты тергеуге мәнді мәліметтер беретінін атап өтеді [14].

Қазақстандық және халықаралық жарияланымдар құжаттың тергеудегі орнын әртүрлі ауқымда сипаттайды. Аратұлы және әріптестері құжаттар қылмыстық әрекеттердің барлық дерлік көрінісінде кездесетінін көрсетеді – графология, сот автортануы және техникалық-криминалистикалық зерттеу осы саладағы үш негізгі бағыт [18]. Deviterne-Lapeyre-дің 2020 жылғы Interpol шолуы күмәнді құжаттарды зерттеу саласын сегіз бағытқа топтастырады: қолжазбаларды салыстырмалы зерттеу, сия құрамын талдау, баспа мәтіндерін талдау, қиылысқан штрихтарды зерттеу, қағазды талдау, бедерленген іздерді анықтау, өзгертілген құжаттарды зерттеу және қорғаныш элементтерін зерттеу [19]. Бұл ауқыммен салыстырғанда NIST есебінің қамтуы тарлау: онда сия мен қағаз, машинкамен басылған мәтін және алдын ала өңдеу тәсілдері қарастырылмаған, негізгі назар қолжазбалық бөлікке аударылған [13]. Отандық әдебиетте құжат қылмыстық құқық бұзушылықпен байланысты әмбебап объект ретінде кеңірек ұғынылады [18].

Қолжазба құжаттарын өңдеудің практикалық қиындықтары бірнеше деңгейде көрінеді. Ruiz-Parrado және әріптестері 1990–2020 жылдар аралығындағы офлайн режимде қолжазба құжаттарды талдауға арналған 5389 жарияланымды саралап, негізгі қиындықтар ретінде жазу мәнерлерінің, әліпбилердің және тілдердің алуан түрлілігін, статикалық скан-кескінмен жұмыс кезінде жоғалатын динамикалық белгілерді көрсеткен [9]. Офлайн жүйелерде жазу жылдамдығы, қаламды қағаздан көтеру сәттері, қысым өзгерістері, штрихтардың орындалу реті тіркелмейді [9]. ENFSI нұсқаулығы зерттеу нәтижесінің құндылығын құжат қаптамасының жай-күйі, жарық сапасы, жұмыс кеңістігінің талапқа сәйкестігі, ластануды болдырмау шаралары және әрекеттерді жүйелі түрде хаттамалаумен тікелей байланыстырады [14]. NIST есебі адами факторлар қатарына жеткіліксіз дайындықты, артық мәнмәтіндік ақпаратты,

стресті, басқарудың әлсіздігін және стандартталмаған жұмыс жағдайларын кіргізген [13]. Тергеуші қызметіндегі қолжазба құжаттардың маңызы мен өңдеудің практикалық қиындықтары кесте 1-де жинақталған. Қолжазба құжатпен жұмыс процессінде оны сақтау, тіркеу, салыстыру және алынған нәтижені дұрыс түсіндіру кезеңдері бар.

Кесте 1 – Тергеуші қызметіндегі қолжазба құжаттардың маңызы және оларды өңдеудің практикалық мәселелері

| Талдау бағыты | Негізгі мазмұны | Дереккөз | Тергеуші қызметі үшін мәні |
|---|---|---|---|
| Қолжазба құжаттың процессуалдық мәні | Қолжазба құжат тек мәтін сақталатын материал емес, іс жүргізу тұрғысынан маңызы бар және сот-криминалистикалық зерттеуге жататын объект болып саналады | АҚШ Сауда Департаментінің Ұлттық стандарттар және технологиялар институты (NIST), 2020 (қолжазбаларды соттық зерттеу туралы есеп); Сот сараптамасы институттарының еуропалық желісі (ENFSI), 2022 (сараптамалық нұсқаулық) | Құжатты ақпарат көзі ретінде ғана емес, зерттеу объектісі ретінде де қарауға мүмкіндік береді |
| Құжаттарды зерттеу ауқымы | Күмән тудыратын құжаттарды зерттеу саласына қолжазбаларды салыстыру, сияны, қағазды, өзгертілген құжаттарды және қорғаныш элементтерін зерттеу кіреді | Deviterne-Lapeyre, 2020 (күмәнді құжаттар бойынша халықаралық шолу) | Тергеуші үшін құжаттың мәтінімен қатар оның материалдық және техникалық белгілері де маңызды екенін көрсетеді |

Кесте 1 жалғасы

| | | | |
|---|---|---|---|
| Қолжазба құжаттарды өңдеудің техникалық қиындықтары | Офлайн талдауда жазу жылдамдығы, қаламды көтеру сәттері, қысым өзгерісі және штрихтардың орындалу реті сияқты динамикалық белгілер жоғалады | Ruiz-Parrado және әріптестер, 2022 (5389 жарияланымға жасалған талдау) | Цифрланған көшірме түпнұсқадағы барлық белгіні толық сақтамайтынын көрсетеді |
| Ұйымдастырушылық және рәсімдік талаптар | Қаптама жай-күйі, жарық сапасы, жұмыс кеңістігі, ластануды болдырмау және әрекеттерді хаттамалау зерттеу нәтижесіне әсер етеді | Сот сараптамасы институттарының еуропалық желісі (ENFSI), 2022 (қолжазбаларды сараптамалық зерттеу тәртібі) | Құжатпен жұмыс сапасы тек технологияға емес, рәсімнің сақталуына да тәуелді екенін көрсетеді |
| Адами факторлардың әсері | Жеткіліксіз дайындық, артық мәнмәтіндік ақпарат, стресс, басқарудың әлсіздігі және стандартталмаған жағдайлар қате қаупін арттырады | АҚШ Сауда Департаментінің Ұлттық стандарттар және технологиялар институты (NIST), 2020 (адами факторлар мен қателік тәуекелі туралы есеп) | Қолжазба құжаттарды өңдеу мен түсіндіруде адами қателік ықтималдығын ескеруді талап етеді |
| HTR технологиясының халықаралық қолданылуы | Transkribus платформасына қатысты 381 жарияланымның 66,67%-ы гуманитарлық қолданбаларға, 25,46%-ы технологиялық зерттеулерге арналған | Nockels және әріптестері, 2022 (Transkribus туралы жүйелі шолу) | Қолжазба мәтінін тану технологиясының мұрағаттық және ғылыми ортада орныққанын көрсетеді |

Кесте 1 жалғасы

| | | | |
|---|---|---|---|
| Құқық қолдану мазмұны бар мұрағаттық мысал | 100000-нан астам бет тарихи полиция есептері негізінде деректер жиыны жасалып, қолжазба мәтінін тану моделі оқытылған | Bias in History, 2022–2023 (тарихи полиция құжаттарымен жұмыс тәжірибесі) | Полициялық сипаттағы құжаттар массивіне HTR-ды бейімдеу мүмкіндігін көрсетеді |
| Қазақстандық білім беру практикасы | KOHTD деректер жиыны студенттердің жазбаша емтихан жұмыстары негізінде құрылып, 922010 таңба мен 140335 сегменттелген кескінді қамтыған | Тойганбаева және әріптестері, 2021 (қазақ тіліндегі қолжазбаларды тану бойынша зерттеу) | Қазақ тіліндегі қолжазба мәтінін тануға арналған жергілікті эмпирикалық база бар екенін дәлелдейді |
| Қолданбалы шектеу | Ішкі істер органдары тергеушілерінің нақты іс жүргізу құжаттарына қатысты ашық репрезентативті корпус табылмаған | Табылған материалдар бойынша | Тергеу құжаттарына бейімделген эмпирикалық база әзірге шектеулі екенін көрсетеді |

Transkribus платформасы – қолжазба мәтінін тану (Handwritten Text Recognition, HTR) үдерістеріне арналған қолданбалы орта. Мұрағаттық және кітапханалық цифрландыру практикасында ол берік орынға ие. Nockels және әріптестері Google Scholar, Scopus және Web of Science базаларынан 2015–2020 жылдарға жататын 381 жарияланымды талдаған [10]. Олардың 66,67%-ы гуманитарлық қолданбаларға, 25,46%-ы технологиялық зерттеулерге, 4,99%-ы пайдаланушыларға бағдарланған зерттеулерге, 2,89%-ы оқыту материалдарына жатады [10]. Сол зерттеуде қолжазба мәтінін танудан (OCR, Optical Character Recognition) бұрын қалыптасқан, баспа мәтініндегі жеке таңбаларды дәл тануға бағытталған технология ретінде сипатталады; күрделі графикада, өзгермелі жазу үлгілерінде және қолжазбада ол шектеулі нәтиже береді [10]. HTR сегменттелген мәтін жолындағы таңбалар мен сөздерді тануға бейімделген, тарихи қолжазбаларды транскрипциялау мен толықмәтінді іздеуді жеделдетеді [10]. Transkribus 2015 жылы READ (Recognition and Enrichment of Archival Documents)

жобасы аясында онлайн құрал ретінде іске қосылды; кейін READ-COOP кооперативі арқылы дамыды [10]. 2020 жылдан бері ақылы үлгіге көшті. Мәдени мұра саласындағы тұрақты айлық аудиториясы шамамен 1700 пайдаланушы [10]. Жарияланымдардың 67,98%-ы ағылшын тілінде, 10,76%-ы неміс тілінде шыққан [10]. Платформа ғылыми айналымда, ең алдымен, еуропалық және ағылшынтілді ортада орныққан.

Қолданбалы мүмкіндіктер мұрағаттық материалдармен бітпейді. Бельгиядағы Bias in History жобасы мұны айқын көрсетті. Жоба материалдарында Lith есімді зерттеуші жасанды интеллект модельдерінің көмегімен 100 000 бет тарихи дереккөзді бір ай ішінде транскрипциялай алғаны баяндалған; мұндай жұмысты қолмен орындау он екі жылға созылар еді [15]. Жобаның 2023 жылғы баяндамасында 1876–1945 жылдарға жататын 100 000-нан астам бет қолжазба полиция есептері негізінде деректер жиыны бірлесіп қалыптастырылғаны айтылса, 2022 жылғы пайдаланушылар конференциясындағы баяндамасында 100 000 қолжазба полиция жазбасына HTR моделін үйрету әрекеттері сипатталған [15]. Бұл мысал маңызды: мұрағаттық ортада дамыған технология құқық қолдану мазмұны бар құжаттар массивіне бейімделе алады. Бірақ жүктелген материалдарда мұндай тәсілдің Қазақстан ішкі істер органдары тергеушілерінің нақты іс жүргізу құжаттарына қатысты нәтижелері туралы ашық эмпирикалық дерек жоқ [15].

Қазақстандық білім беру практикасы өз бағытында дамып келеді. Тойганбаева және әріптестері 2021 жылы KOHTD (Kazakh Offline Handwritten Text Dataset) деректер жиынын студенттердің жазбаша емтихан жұмыстары негізінде құрастырған; оған шамамен 922 010 таңба мен 140 335 сегменттелген кескін енгізілген [11]. Халықаралық мұрағаттық ортада Transkribus туралы 381 жарияланым талданса [10], қазақстандық мысалда нақты оқу процесінен алынған қолжазба материалдарының сандық корпусы ұсынылған [11]. Айырмашылық тура көрінеді. Бірінде платформа сипатталады, екіншісінде нақты деректер жиыны мен оқыту базасы беріледі.

## 1.2 Қолжазба мәтіндерін тану технологияларының ұғымы, ерекшеліктері және қолданбалы мүмкіндіктері

Қолжазба мәтінін тану – Handwritten Text Recognition, HTR – қолмен жазылған мәтінді машина оқи алатын таңбалық форматқа айналдыру міндетін орындайды. Garrido-Munoz және әріптестерінің 2025 жылғы шолуында бұл міндет үлгілерді тану мен машиналық оқытудың дербес бағыты ретінде сипатталған; авторлар HTR-дің қолданылу салаларына тарихи құжаттарды сақтау, автоматтандырылған деректер енгізу, цифрлық жазбалар және көру қабілеті шектеулі адамдарға арналған қолжетімділік шешімдерін кіргізеді [20]. Сол еңбекте HTR күрделілігі қолжазбаның жоғары вариативтілігімен – жазу стилі, көлбеулігі, өлшемі, әріп безендірілуінің дара айырмашылықтарымен –

байланыстырылған [20]. Abdallah және әріптестері бұл міндетті айнымалы екіөлшемді кескіннен таңбалар тізбегін алу үдерісі деп нақтылайды [16].

OCR мен HTR бір міндетті шешпейді. Nockels және әріптестері 2022 жылы OCR-ды баспа мәтініндегі жеке таңбаларды дәл тануға бағытталған ертерек технология ретінде сипаттайды [10]. Күрделі графикада, өзгермелі жазу үлгілерінде және қолжазбада ол шектеулі нәтиже береді [10]. HTR керісінше сегменттелген мәтін жолындағы таңбалар мен сөздерді тануға бейімделген, тарихи қолжазбаларды транскрипциялау мен толықмәтінді іздеуді жеделдетеді [10]. Айырмашылық технологиялардың қолданылу шегін нақты ажыратады. Бірі баспа мәтінін цифрландыруға тиімдірек. Екіншісі қолжазба деректеріне арналған. Олардың салыстырмалы сипаттамасы кесте 2-де берілген.

Кесте 2 – Қолжазба мәтінін тану технологиясының негізгі сипаттамалары

| Өлшем | Қолжазба мәтінін тану | Таңбаларды оптикалық тану |
|---|---|---|
| Танылатын объект | Қолмен жазылған мәтін | Баспа мәтіні |
| Бастапқы дерек | Сканерленген немесе цифрланған қолжазба кескіні | Сканерленген баспа құжаты |
| Негізгі міндет | Қолжазба мәтінді машина оқи алатын форматқа айналдыру | Баспа таңбаларын автоматты тану |
| Күрделілік көзі | Жазу мәнерінің, көлбеуліктің, өлшемнің, әріп байланыстарының өзгермелілігі | Қаріп сапасы, бет сапасы, орналасу құрылымы |
| Қолданылатын тәсілдер | Ықтималдық модельдер, қайталанатын нейрондық желілер, назар механизмдері, түрлендіргіш архитектуралар | Таңбаларды үлгі бойынша тану, оптикалық өңдеу, кейінгі машиналық оқыту тәсілдері |
| Нәтиже | Қолжазба мәтіннің танылған цифрлық нұсқасы | Баспа мәтіннің танылған цифрлық нұсқасы |
| Қолданылу саласы | Қолжазба құжаттар, тарихи материалдар, оқу жұмыстары, цифрландырылған жазбалар | Баспа құжаттар, кітаптар, формалар, стандартты мәтіндік материалдар |

HTR күрделілігі тек жазуды оқуда емес, оны модельдеу деңгейінде де көрінеді. Жұмыс логикасы өзара байланысты бірнеше кезеңнен тұрады. Жалпы құрылым 1-суретте берілген.




Сурет 1 – Қолжазба мәтінін тану технологиясының жалпы жұмыс кезеңдері

Garrido-Munoz және әріптестерінің 2025 жылғы шолуында бұл бағыттың эволюциясы көрсетілген: эвристикалық және ережеге негізделген әдістерден ықтималдық модельдерге, одан қайталанатын нейрондық желілерге, одан әрі назар механизмдері мен түрлендіргіш архитектураларға өткен [20]. Сол мақалада тану деңгейлері екі топқа бөлінген: бір жол көлеміндегі тану және абзац не құжат деңгейіндегі тану [20]. Бірінші деңгейде сөз бен жол танылады. Екінші деңгейде бірнеше жолдан тұратын құрылым тұтас оқылады [20]. Бұл жіктеу қолданбалы мәнге ие – қолжазба мәтінді өңдеу таңбаны жеке танумен бітпейді, ол тұтас құрылымды қалпына келтіруді де қамтиды [20].

Қазақстандық зерттеулерде технология нақты деректер арқылы көрсетілген. Тойганбаева және әріптестері KOHTD-ны 3000 студенттік емтихан парағы негізінде құрастырған [11]. Жиында 140 335 сегменттелген кескін мен шамамен 922 010 таңба бар [11]. Авторлар бұл деректер жиынын алгоритмдердің сапасын салыстыруға және жаңа модельдерді оқытуға жарамды стандартты база ретінде сипаттайды [11]. Жауаптардың 99%-ы қазақ тілінде, 1%-ы орыс тілінде [11]. Білім беру практикасы HTR-ды теориялық емес, нақты оқу материалдары арқылы тексеру мүмкіндігін берді [11].

Қазақ және орыс тілдеріне арналған тағы бір база аралас тілдік ортадағы тану міндетін сипаттайды. Нурсеитов және әріптестерінің HKR базасында мәтіндердің шамамен 95%-ы орыс тілінде, 5%-ы қазақ тілінде; құрамында 1400-ден астам толтырылған форма, шамамен 63 000 сөйлем және 715 699-дан астам таңба [12]. KOHTD еркін емтихан жауаптарына сүйенсе, HKR пошта, мекенжай мен кілт сөздерге жақын формалық мәтіндерге сүйенеді [11], [12]. Айырмашылық технологияның қолданбалы икемділігін көрсетеді. Бір жағдайда

еркін оқу жауабы өңделеді. Екінші жағдайда құрылымы алдын ала шектелген мәтіндер танылады [11], [12].

Модельдік нәтижелер технологияның практикалық әлеуетін сандармен көрсетеді. Abdallah және әріптестері 2020 жылы қазақ және орыс тілдеріндегі деректерде назар механизміне негізделген модельді сынаған [16]. Бірінші тест жиынында таңба қателігінің деңгейі 4,13%, сөз қателігінің деңгейі 18,91%, сөйлем қателігінің деңгейі 25,72% болған [16]. Екінші тест жиынында бұл көрсеткіштер тиісінше 6,31%, 23,69% және 35,16% [16]. Сол мақалада ұсынылған модель Bluche пен Puigcerver модельдерімен салыстырылған [16]. Bluche моделінің таңба қателігі бірінші тест жиынында 10,55%, Puigcerver моделінікі 20,73% [16]. Архитектура нәтижеге тікелей әсер етеді.

Орыс тіліндегі зерттеулер дерек көлемінің жеткіліксіздігі мен почерк алуан түрлілігі тану сапасына тікелей әсер ететінін көрсетеді. Богатенкова, Беляева және Перминов 2023 жылы синтетикалық деректерді қосу екі ашық кириллдік жиында және екі түрлі модельде орта есеппен тану дәлдігін 6%-ға арттырғанын көрсеткен [21]. Скрыпников және әріптестері 2021 жылы HTR күрделілігін бір адамның бір таңбаны әртүрлі жазуымен, әріптердің қосылуымен және сканерлеу ақауларымен байланыстырады [22]. Бір зерттеу шешімнің тиімділігін көрсетеді. Екіншісі бастапқы қиындықтың табиғатын ашады. Қайшылық емес – екі қыр.

Құқық қорғау қызметімен байланысты қолданбалы мүмкіндік HTR-ды дербес дәлелдеу құралы ретінде емес, құжат массивін өңдеу құралы ретінде қарастыруға негіз береді. Бобовкин мен Ручкиннің 2018 жылғы статистикасы қолжазба объектілерінің құқық қорғау тәжірибесінде тұрақты қолданылатынын дәлелдейді [17]. Соған қарамастан, табылған деректерде ішкі істер органдары тергеушілерінің нақты іс жүргізу құжаттарына арналған ашық репрезентативті тану корпусы көрсетілмеген [11], [12], [15], [17]. Сондықтан HTR-ды қазіргі кезеңде тергеуші үшін, ең алдымен, қолжазба құжаттарды бастапқы цифрландыруға, мәтін бойынша іздеуге бейімдеуге, жүйелеуге және электрондық өңдеуге арналған көмекші цифрлық құрал деп сипаттаған орынды.

## 1.3 Ішкі істер органдары тергеушісінің ұйымдастырушылық-техникалық қызметінде қолжазба мәтіндерін тану технологияларын қолданудың негізгі бағыттары және олардың тергеу қызметін ұйымдастыруға ықпалы

HTR-ды тергеушінің ұйымдастырушылық-техникалық қызметінде қолдану бастапқы цифрландырудан басталады. Garrido-Munoz және әріптестері HTR жүйелері сөз деңгейіндегі модельдерден жол деңгейіндегі және құжат деңгейіндегі шешімдерге өткенін көрсетеді [20]. Nockels, Gooding, Ames және Terras бұл технология кітапханалар мен архивтерде бастапқы дереккөздерді транскрипциялауды жеделдетіп, толық мәтін бойынша іздеу мен ірі көлемдегі деректерді талдауға мүмкіндік беретінін жазады [10]. Қолжазба құжатты сканерленген бейне күйінен машина өңдей алатын мәтіндік форматқа көшіру – дербес қолданбалы бағыт [10], [20].

Шетелдік мұрағаттық орта мен қазақстандық білім беру практикасы бұл бағыттың әртүрлі тәжірибелік негізде жұмыс істейтінін көрсетеді. Nockels және әріптестері Transkribus туралы 381 жарияланымды талдап, олардың 67%-ын гуманитарлық қолдану саласына, 25%-ын технологиялық зерттеулерге, 5%-ын пайдаланушылар тәжірибесіне, 3%-ын әдістемелік нұсқауларға жатқызған [10]. Тойганбаева және әріптестері KOHTD-ны 3000 студенттік емтихан парағы, 140 335 бөлшектелген кескін және шамамен 922 010 таңба негізінде сипаттайды [11]. Бір дерек цифрландырудың платформалық таралуын көрсетсе, екіншісі нақты оқу материалынан алынған қазақтілді қолжазба массивінің өңделу мүмкіндігін көрсетеді [10], [11].

Келесі бағыт – мәтіндік іздеу мен жүйелеу. Nockels және әріптестері HTR-ды толық мәтін бойынша іздеу және ірі көлемдегі деректерді талдау үшін пайдаланатын құрал деп сипаттайды [10]. Ruiz-Parrado, Heradio, Aranda-Escolastico және әріптестері желіден тыс қолжазба құжаттарды талдау саласында үш міндетті ажыратады: қолжазба мәтінді тану, мәтін ішінен нақты сөздерді табу және қажетті ақпаратты іздеу [9]. Тергеуші үшін мұның мәні бар. Қолжазба жазбалар тек оқылып қана қоймайды. Олар мәтін бойынша іздеуге бейімделеді, ұқсас фрагменттер табылады, көп материал тақырыптық не мазмұндық белгі бойынша бөлінеді [9], [10].

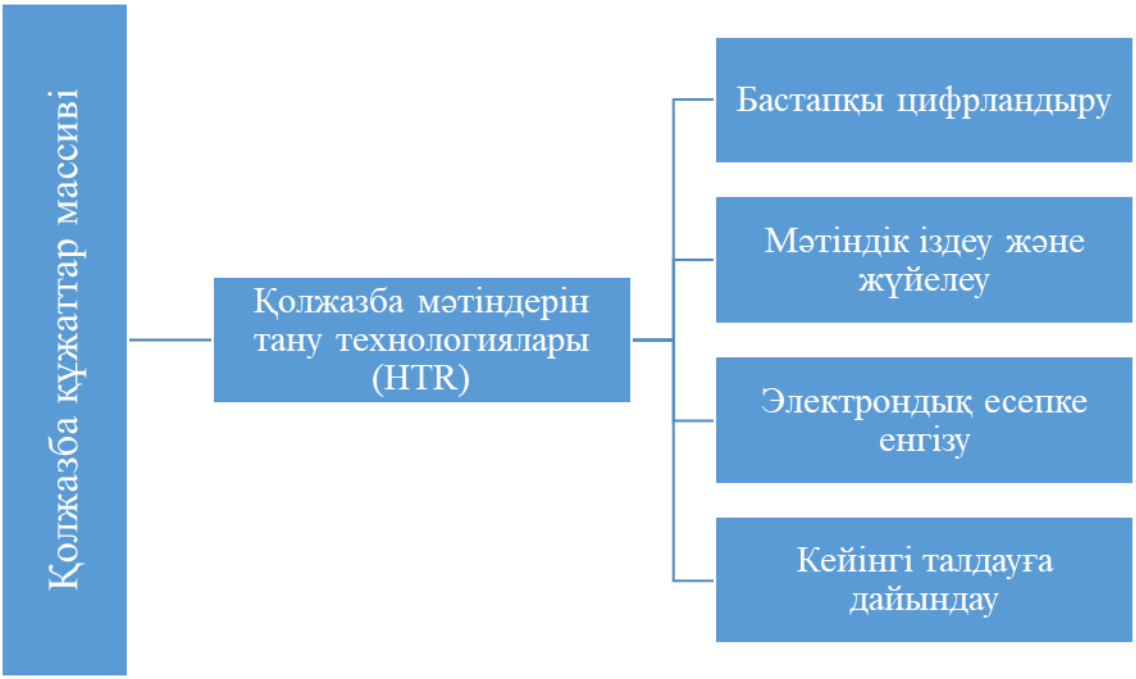
Қазақ тіліндегі және қазақ-орыс аралас деректер жиындары екі ұйымдастырушылық моделді көрсетеді. KOHTD Сәтбаев университеті мен әл-Фараби атындағы Қазақ ұлттық университеті студенттерінің емтихан жұмыстары негізінде жасалған [11]. Жауаптардың 99%-ы қазақ тілінде, 1%-ы орыс тілінде; материал жол және сөз деңгейінде бөлінген [11]. HKR дерекқоры 95% орысша және 5% қазақша сөздер мен сөйлемдерден тұрады; құрамында 1400-ден астам толтырылған нысан, шамамен 63 000 сөйлем, 715 699-дан астам таңба, шамамен 200 жазушының қолжазбасы [12]. Біріншісі еркін мәтінді жүйелеуге бейімделген. Екіншісі құрылымы алдын ала шектелген үлгілік нысандарды электрондық өңдеуге жақын [11], [12].

Электрондық есепке енгізу мен көп көлемді материалдарды іріктеуге дайындау жеке бағыт. KOHTD сипаттамасында емтихан парақтарын автоматты тану үшін контурларды бағалау, кескіннің бұрылуын түзету, жолдар мен сөздерге бөлу кезеңдері берілген [11]. HKR сипаттамасында деректерді жинау, дюймге 300 нүкте сапасында сканерлеу, нысандарды алдын ала өңдеу, шекаралық артефактілерді алып тастау және сөздерді бір өлшемге келтіру рәсімдері көрсетілген [12]. Операциялардың бағыты бір: тану нәтижесін бірден дәлелдеме ретінде пайдалану емес, құжаттарды электрондық қорда сақтау, қайта табу, сүзу және кейінгі сараптамалық не талдамалық жұмысқа дайындау [11], [12].

Кейінгі талдауға жарамды цифрлық мәтіндік қор қалыптастыру тағы бір бағыт. Тойганбаева және әріптестері KOHTD-ны алгоритмдердің өнімділігін салыстыруға арналған стандартталған және тазартылған деректер жиыны деп сипаттайды [11]. Нурсеитов және әріптестері HKR-ды қазақ және орыс тілдеріндегі сөздерді тану зерттеулеріне арналған оқыту мен сынақ жиыны

ретінде ұсынады [12]. Қазақстандық оқу ортасындағы еркін жауаптар жиыны мен пошта-адрестік нысандарға ұқсас құрылымдалған деректер жиыны арасындағы айырма тергеуші жұмысы үшін екі қолданбалы салдар береді: еркін мазмұнды қолжазбаларды іздеу мен топтауға бейімдеу және үлгілік нысандағы жазбаларды электрондық есепке енгізу [11], [12].

Сандық нәтижелер бұл бағыттардың техникалық орындалу мүмкіндігін көрсетеді. Бірақ абсолютті дәлдікті растамайды. Abdallah, Hamada және Нурсеитов қазақ және орыс тілдеріндегі деректерде назар тетігіне негізделген толық қақпаланған конволюциялық нейрондық желі мен екі бағытты қақпаланған рекуррентік блокқа құрылған модельді сынаған [16]. Бірінші сынақ жиынында таңба қатесінің деңгейі 0,045, сөз қатесінің деңгейі 0,192, тізбек қатесінің деңгейі 0,253 болған [16]. Екінші сынақ жиынында көрсеткіштер тиісінше 0,064, 0,24 және 0,361 [16]. KOHTD материалында бір ерікті пайдаланушы бойынша жеке сөздердің ұқсастық көрсеткіштері 60%-дан 100%-ға дейін құбылған; жазылуы ұқсас «ң» мен «қ», «м» мен «ш», «н» мен «и», «л» мен «е» сияқты әріптер тануды күрделендірген [11]. Бір дерек модельдің жұмысқа қабілетті екенін көрсетеді. Екіншісі қолжазбаның дара мәнері тану нәтижесіне тікелей әсер ететінін дәлелдейді [11], [16].




Сурет 2 – Қолжазба мәтіндерін тану технологияларының тергеушінің ұйымдастырушылық-техникалық қызметіндегі қолдану моделі

Құқық қорғау қызметіне қатысты ұйымдастырушылық салдар қолжазба объектілерінің құқық қолдану практикасында тұрақты кездесетінінен туындайды (сурет 2). Бобовкин мен Ручкиннің 2018 жылғы зерттеуі сот қолжазбатану сараптамасында жиі зерттелетін объектілердің 65%-ын қолтаңбалар, 30%-ын

қысқа жазбалар құрайтынын, мұндай сараптамалардың 78%-ы алдын ала тергеу сатысында жүргізілетінін көрсетеді [17]. Қолжазба объектілері бар құжаттар массиві құқық қорғау қызметінің нақты айналымында жұмыс істейді. Алайда Бобовкин мен Ручкин 2018 тергеу сатысындағы сараптамалық айналымды сипаттайды, ал шетелдік HTR жөніндегі еңбектер көбіне мұрағаттық және білім беру деректеріне сүйенеді. Ішкі істер органдары тергеушілерінің нақты іс жүргізу құжаттары негізінде жасалған ашық әрі өкілді деректер жиыны қарастырылған материалдар ішінде жоқ. Бұл мәселе бойынша тәжірибелік база жеткіліксіз.

Осы деректер HTR-дың тергеуші қызметіндегі қолданбалы мүмкіндіктері мен шектеулерін нақтылайды. Қолжазба мәтіндерін тану технологиялары тергеушінің ұйымдастырушылық-техникалық қызметінде бастапқы цифрландыруға, мәтіндік іздеуге, жүйелеуге, электрондық есепке енгізуге, көп көлемді материалдарды іріктеуге және ақпаратты кейінгі талдауға дайындауға пайдаланылады (кесте 3).

Кесте 3 – Қолжазба мәтіндерін тану технологияларын қолдану бағыттары және олардың тергеу қызметін ұйымдастыруға ықпалы

| Қолдану бағыты | Мазмұны | Тергеуші қызметіне ықпалы | Шектеуі |
|---|---|---|---|
| Бастапқы цифрландыру | Қолжазба құжатты сканерленген бейнеден мәтіндік форматқа көшіру | Құжатпен электрондық жұмыс істеуге мүмкіндік береді | Сканерлеу сапасы мен қолжазба анықтығына тәуелді |
| Мәтіндік іздеу | Танылған мәтін ішінен сөздер мен фрагменттерді табу | Көп материал ішінен қажетті ақпаратты жылдам табуға көмектеседі | Қате танылған сөздер іздеу нәтижесін бұрмалайды |
| Жүйелеу | Құжаттарды мазмұндық немесе тақырыптық белгі бойынша бөлу | Материалдар массивін реттеуге жағдай жасайды | Автоматты бөлу тергеушінің тексеруін қажет етеді |
| Электрондық есепке енгізу | Қолжазба деректерді электрондық қорда сақтауға жарамды форматқа келтіру | Құжаттарды қайта табу, сүзу және салыстыру жеңілдейді | Нәтиже бірден дәлелдеме ретінде қолданылмайды |
| Кейінгі талдауға дайындау | Цифрлық мәтіндік қор қалыптастыру | Ірі көлемдегі қолжазба материалдарды талдауға негіз болады | Нақты тергеу құжаттары бойынша ашық деректер жеткіліксіз |

Бұл технологиялар құқықтық бағалау мен дәлелдемелік шешімді алмастырмайды. Олардың рөлі – қолжазба құжаттар массивімен жұмысты реттеу, құжаттық ақпаратты іздеуге жарамды форматқа көшіру, тергеушінің құжат айналымын оңтайландыру. Осы шектегі қызмет – HTR-ды тергеушінің ұйымдастырушылық-техникалық жұмысын жетілдіру құралы деп қарастыруға негіз.

Қолжазба құжаттар тергеуші қызметінде тұрақты кездеседі. Олармен жұмыс көпсатылы техникалық және ұйымдастырушылық әрекеттерді талап етеді. HTR осы үдерістердің бір бөлігін цифрлық форматқа көшіруге мүмкіндік береді. Қарастырылған еңбектер мен деректер жиындары бұл технологиялардың қолжазба құжаттарды бастапқы цифрландыруға, мәтін бойынша іздеуге бейімдеуге, жүйелеуге, электрондық есепке енгізуге, көп көлемді материалдарды іріктеуге және кейінгі талдауға дайындауға қолданылатынын дәлелдейді. Сонымен бірге олар құқықтық бағалау мен дәлелдемелік шешімді алмастырмайды. Негізгі қолданылу шегі тергеушінің ұйымдастырушылық-техникалық қызметімен байланысты. Бұл шектеу қолжазба мәтінін тану технологияларын тергеушінің құжат айналымын оңтайландыру, ақпаратты өңдеу рәсімдерін жүйелеу және қолжазба материалдар массивін электрондық форматта пайдалануға бейімдеу үшін перспективалы құрал ретінде көрсетеді.

## 2 Қолжазба мәтіндерін тану нәтижелерін пайдаланудың іс жүргізу шектеулері және оларды тексеру мен растау тәртібі

### 2.1 Қолжазба мәтіндерін тану нәтижелерінің қылмыстық процестегі құқықтық табиғаты

HTR нәтижелерінің қылмыстық процестегі құқықтық табиғатын ҚПК дәлелдемелер туралы нормалары айқындайды [23]. ҚПК-нің 111-бабында дәлелдемелер қылмыстық заңда көзделген әрекеттің болғанын немесе болмағанын, оны нақты адамның жасағанын не жасамағанын, кінәлілік пен өзге де мән-жайларды анықтайтын заңды түрде алынған нақты деректер ретінде берілген [23]. 125-бапқа сәйкес әрбір дәлелдеме қатыстылығы, жол берілетіндігі және анықтығы тұрғысынан, ал барлық дәлелдемелер жиынтығы істі шешу үшін жеткіліктілігі тұрғысынан бағаланады [23]. 126-бап қылмыстық процесті жүргізетін органға, қорғаушыға, адвокатқа, маманға және сарапшыға дәлелдеу процесінде ғылыми-техникалық құралдарды пайдалануға жол береді [23]. Осы нормалар жиынтығы HTR нәтижесін дербес дәлелдеме көзі емес, дәлелдеу процесінде қолданылуы мүмкін ғылыми-техникалық өңдеу өнімі ретінде сипаттауға негіз береді [23].

Электрондық қылмыстық іс тәртібі бұл тұжырымды нақтылай түседі. Қылмыстық сот ісін электрондық форматта жүргізу жөніндегі нұсқаулықтың ұғымдық бөлігінде «Электрондық қылмыстық іс» модулі электрондық қылмыстық істі дайындауды, жүргізуді, жөнелтуді және сақтауды ұйымдастыруға арналғаны, ал «электрондық қылмыстық іс» - қылмыстық қудалау органы е-ҚІ модулінің функционалы арқылы қалыптастыратын іс жүргізу нысаны екені бекітілген [24]. Қазақстандық нұсқаулық электрондық айналымды ұйымдастыруға басымдық береді. ENFSI-дің 2022 жылғы нұсқаулығы сараптамалық қорытындыны талдау, салыстыру, бағалау және шолу кезеңдері арқылы қалыптастыруды көздейді [14]. Бірінші модель құжаттық ағынды ұйымдастырады. Екіншісі қолжазба белгісін сараптамалық бағалауға бағытталған. HTR нәтижесі электрондық іске енгізілуі мүмкін, бірақ оның процессуалдық салмағы автоматты тану фактісінен емес, кейінгі тексеру мен бағалаудан туындайды [14], [24].

Ақпарат саласындағы құқықтық реттеу нәтижелердің екінші қырын ашады. «Ақпараттандыру туралы» заңның 1-бабында автоматтандыру ақпаратты жасауды, іздестіруді, жинауды, жинақтауды, сақтауды, өңдеуді, алуды және пайдалануды қамтитын процесс ретінде айқындалған [25]. HTR құқықтық мағынада мазмұнды бағалау емес – ақпаратты автоматтандырылған өңдеу нысандарының бірі [25]. Мұндай нәтижелер қылмыстық іс үшін маңызға ие болуы мүмкін. Бірақ олардың құқықтық статусы бастапқы қолжазба құжаттың орнын алмастырмайды [23], [25].

Дербес деректерді қорғаудың құқықтық тәртібі мәселені одан әрі шектейді. «Дербес деректер және оларды қорғау туралы» заңның 5-бабында дербес

деректерді жинау, өңдеу және қорғау қағидаттары бекітілген; 8-бап субъектінің дербес деректерді жинауға және өңдеуге келісім беру тәртібін белгілейді [26]. Қолжазба арыздар, түсініктемелер, қолхаттар мен жазбаларда жәбірленуші, куә, күдікті және өзге тұлғалар туралы сәйкестендіруші мәліметтер жиі болады [26]. Қазақстандық құқықтық режим мұндай ақпаратты жинау мен өңдеуді арнайы заң арқылы шектейді. NIST 2020 мен ENFSI 2022 контекстік ақпараттың сараптамалық шешімге әсерін басқару қажеттігін көрсетеді [13], [14], [26]. Бірінші бағыт деректер субъектісінің құқықтарын қорғайды. Екіншісі сарапшының бейтараптығын. Екі бағыт та HTR нәтижесін бақылаусыз және түпкі процессуалдық өнім ретінде пайдалануға жол бермейді [13], [14], [26].

Ғылыми-қолданбалы материалдар құқықтық шектеуді сандық деректермен растайды. NIST-тің 2020 жылғы есебінде автоматтандырылған жүйелер қолжазбаны сараптамалық зерттеу үдерісін қолдау үшін пайдаланылатыны көрсетілген [13]. Сол есепте Sargur Srihari тобы 1500 адамның қолжазба үлгілері негізінде жүргізген зерттеу нәтижелері берілген [13]. Кеңейтілген қолжазба үлгілері бойынша жазуды орындаушыны айқындаудағы қателік деңгейі егіздер үлгілерінде 13%, ал егіз емес үлгілерде 4% [13]. Бұл сандар автоматтандырылған тану немесе салыстыру жүйесінің нәтижесін өздігінен процессуалдық ақиқат ретінде қабылдауға болмайтынын көрсетеді [13]. Мұндай нәтиже тек аралық талдамалық дерек ретінде бағаланады [13], [23].

Еуропалық сараптамалық нұсқаулық пайымды өзге ұғымдық аппарат арқылы растайды. ENFSI-дің 2022 жылғы нұсқаулығында қорытынды дәлелдемелік мәнді бағалау негізінде қалыптасатыны, ал бағалау кезеңі өзара бәсекелес жорамалдарды салыстырумен және кейінгі тексеру рәсімдерімен байланыстырылатыны көрсетілген [14]. ҚПК дәлелдемелерді бағалауды ішкі сенім және дәлелдемелер жиынтығының жеткіліктілігі қағидаттарымен байланыстырады [23]. ENFSI сараптамалық қорытындының сенімділігін қосымша тексеру рәсімі арқылы күшейтеді. Бобовкин мен Ручкин 2018 құқық қорғау практикасында қолжазба объектілерінің нақты үлес салмағын сандармен сипаттайды [17]. Қорытынды бір нүктеге жинақталады: қолжазба материалымен жұмыс қылмыстық процесте елеулі орын алады, бірақ HTR нәтижесі сот сараптамасының қорытындысымен теңестірілмейді [14], [17], [23].

ҚПК нормаларын электрондық қылмыстық іс жүргізу нұсқаулығымен қатар қарастырғанда HTR нәтижесінің қылмыстық процестегі орны айқындала түседі [23], [24]. Ол бастапқы құжаттың өзіне теңестірілмейтін, одан туындайтын цифрлық қабат ретінде көрінеді. Цифрлық қабат құжатты электрондық қылмыстық іске енгізу, мәтіндік іздеуге бейімдеу, материалдарды жүйелеу және кейінгі талдауға даярлау мақсатында пайдаланылады [24], [25]. Дәлелдемелік мәні дербес түрде емес – түпнұсқа қолжазба құжаттың өзімен, оны алу мен тіркеудің заңдылығымен, салыстырып тексеру мүмкіндігімен және процестік бағалау нәтижесімен байланыста айқындалады [23], [24]. ҚПК-нің 111, 125 және 126-баптары ғылыми-техникалық құралдарды қолдануға жол береді. Бірақ

олардың көмегімен алынған нәтижені автоматты түрде дербес дәлелдеме көзі ретінде танымайды [23].

Талданған нормативтік және ғылыми-қолданбалы материалдарда HTR нәтижелерінің қылмыстық процестегі құқықтық табиғатын нақтылайтын мысалдар берілмеген. Бұл мәселе бойынша дерек жеткіліксіз.

HTR нәтижелері Қазақстан қылмыстық процесінде дербес дәлелдеме көзіне жатпайды [23]-[26]. Олар ғылыми-техникалық құралдарды қолдану арқылы алынатын, электрондық қылмыстық іс шеңберінде айналымға енуі мүмкін туынды ақпараттық өнім ретінде бағаланады [23]–[25]. Құқықтық мәні түпнұсқа құжатпен салыстыру, өзге дәлелдемелермен тексеру, қажет болғанда маман немесе сарапшы қатысуымен растау арқылы қалыптасады [14], [17], [23], [24].

## 2.2 Қолжазба мәтіндерін тану нәтижелерін пайдаланудың іс жүргізушілік талаптары мен шектеулері

HTR нәтижесін қылмыстық процесте пайдалану оның техникалық түрде алынуымен емес, кейінгі тексерілуімен айқындалады. ҚПК-нің 111-бабы дәлелдемелерді заңды түрде алынған нақты деректер ретінде бекітеді. 125-бап оларды қатыстылығы, жол берілетіндігі және анықтығы тұрғысынан бағалауды талап етеді. 126-бап ғылыми-техникалық құралдарды қолдануға жол береді [23]. Үш норма бірге HTR нәтижесін процестік айналымнан тыс қалдырмайды. Бірақ оны тексерілмеген күйінде дербес негіз ретінде қабылдауға да жол бермейді.

Құжаттың бастапқы нұсқасы шешуші мәнін жоғалтпайды. Электрондық қылмыстық іс жүргізу нұсқаулығы электрондық істі дайындау, жүргізу, жөнелту және сақтау тәртібін белгілейді [24]. Бұл реттеу құжат қозғалысын цифрлық ортада ұйымдастырады. Еуропалық сараптамалық тәжірибеде қолжазбамен жұмыс одан тереңірек тексеру логикасына құрылады. ENFSI-дің 2022 жылғы нұсқаулығында қорытынды талдау, салыстыру, бағалау және кейін қайта тексеру арқылы қалыптасатыны көрсетілген. Қазақстандық тәртіп электрондық істі жүргізуге басымдық берсе, еуропалық нұсқаулық қолжазба белгісін тексерудің өзіне назар аударады. Тоғысатын жері біреу: автоматты түрде танылған мәтін түпнұсқамен салыстырылмайынша, оның мазмұны процестік тұрғыдан жеткілікті деп қабылданбайды.

Автоматты тану мәтінді оқуды жеңілдетеді. Бірақ мағынасын түпкілікті бекітпейді. NIST-тің 2020 жылғы есебінде мұндай жүйелер қолжазбаны зерттеу үдерісін қолдауға арналғаны, нәтижеге адами факторлар, артық мәнмәтіндік ақпарат және тексеру сапасы әсер ететіні талданған. Күмән туғызатын сөздер, бұрмаланған тіркестер немесе екіұшты оқылатын бөліктер тергеуші тарапынан қайта қаралуға тиіс. Қайта қарау техниканы жоққа шығару емес. Ол қылмыстық іс үшін елеулі мазмұнды тексерілмеген күйінде бекітпеу талабынан туындайды.

HTR нәтижесі электрондық іске енгізілген кезде оның қайдан алынғаны мен қалай өңделгені де ашық болуы керек. «Ақпараттандыру туралы» заңның 1-

бабында автоматтандыру ақпаратты жасау, іздестіру, жинау, жинақтау, сақтау, өңдеу, алу және пайдалану үдерісі ретінде анықталған [25]. Анықтама HTR нәтижесінің құқықтық шегін көрсетеді: мұнда сөз құқықтық қорытынды туралы емес, ақпаратты өңдеу туралы. Танылған мәтіннің іс материалында қандай құжаттан алынғаны, қандай бейнеге сүйенгені және кейін кімнің тексергені көрініп тұруға тиіс. Әйтпесе ол тек оқуға ыңғайланған мәтін күйінде қалады.

Мәселе дербес деректерді қорғау талабымен күрделене түседі. «Дербес деректер және оларды қорғау туралы» заңның 5-бабы жинау мен өңдеудің қағидаттарын бекітеді, 8-бабы келісім беру тәртібін айқындайды [26]. Қолжазба арыздарда, түсініктемелерде және қолхаттарда адамдарды сәйкестендіретін мәліметтер жиі болады. Мұндай мәтінді машина оқи алатын нысанға көшіру тек техникалық қадам емес. Ол дербес деректерді өңдеуге де айналады. Қазақстандық заңнама бұл мәліметтерді құқық иесінің мүддесі тұрғысынан қорғайды. NIST 2020 міндетке қатысы жоқ мәнмәтіндік ақпараттың талдауға артық ықпал ету қаупін көрсетеді. Біріншісі деректер иесінің құқығын шектен тыс өңдеуден сақтайды. Екіншісі талдаушының бейтараптығын қорғайды. Тексерілмеген және бақылаусыз өңделген мәтін қылмыстық процесте сенімді негіз ретінде қаралмайды.

Даулы жағдайларда арнайы білімнің маңызы күрт артады. Танылған мәтін түпнұсқамен сәйкес келмесе, жазудың мазмұны бірнеше түрде оқылса немесе құжаттағы белгі авторлығы мен түпнұсқалығы тұрғысынан мәнге ие болса, тергеушінің техникалық пайымы жеткіліксіз. Сандар қолжазба объектілерінің құқық қорғау қызметінде шеткері емес екенін дәлелдейді.

Қолданылу шегі осыдан көрінеді. HTR нәтижесі қылмыстық іс бойынша шешім қабылдаудың жалғыз тірегі бола алмайды [23]. HTR нәтижелерін қылмыстық процесте пайдалануға қойылатын негізгі талаптар мен шектеулер 4-кестеде жинақталған.

Кесте 4 – Қолжазба мәтіндерін тану нәтижелерін қылмыстық процесте пайдаланудың іс жүргізушілік шарттары мен шектеулері

| Іс жүргізушілік шарт | Мазмұны | Құқықтық негізі | Шектеуі |
|---|---|---|---|
| Түпнұсқа құжатпен міндетті салыстыру | Танылған мәтін бастапқы қолжазба құжатпен салыстырылғаннан кейін ғана пайдаланылуға тиіс | ҚР ҚПК 111, 125-баптары; электрондық қылмыстық іс жүргізу жөніндегі нұсқаулық; ENFSI, 2022 | Түпнұсқамен салыстырылмаған мәтін процестік тұрғыдан жеткілікті деп қабылданбайды |

Кесте 4 жалғасы

| | | | |
|---|---|---|---|
| Қате немесе күмәнді бөліктерді қолмен қайта тексеру | Автоматты түрде танылған мәтіндегі бұрмаланған, толық танылмаған немесе екіұшты фрагменттер адам тарапынан қайта қаралуға тиіс | ҚР ҚПК 125, 126-баптары; NIST, 2020; ENFSI, 2022 | Қолмен нақтыланбаған фрагменттер мазмұнды бұрмалауы мүмкін |
| Электрондық іске енгізудің рәсімдік дұрыстығы | Танылған мәтіннің қай құжаттан алынғаны, қандай бейне негізінде қалыптасқаны және кейін кімнің тексергені айқын болуы керек | Электрондық қылмыстық іс жүргізу жөніндегі нұсқаулық; «Ақпараттандыру туралы» заң | Шығу тегі мен өңделу жолы көрсетілмеген мәтін тек техникалық өнім күйінде қалады |
| Дербес деректерді қорғау талаптарын сақтау | Танылған мәтінде тұлғаларды сәйкестендіретін мәліметтер болса, оларды өңдеу заң талаптарына сәйкес жүзеге асырылуға тиіс | «Дербес деректер және оларды қорғау туралы» заңның 5, 8-баптары | Бақылаусыз өңдеу мен тарату заңдылықты бұзады және мұндай нәтижеге деген сенімді әлсіретеді |
| Маман немесе сарапшыны тарту қажеттігі | Даулы, толық оқылмайтын немесе түпнұсқалығы мәнге ие қолжазба объектілері бойынша арнайы білім иесін тарту қажет болуы мүмкін | ҚР ҚПК 126-бабы; Bobovkin, Ruchkin, 2018; ENFSI, 2022 | Тергеушінің техникалық пайымы барлық жағдайда жеткілікті бола бермейді |
| Процестік шешімнің жалғыз негізі ретінде пайдаланбау | Танылған мәтін өзге дәлелдемелермен бірге ғана бағалануға тиіс | ҚР ҚПК 111, 125-баптары; NIST, 2020 | HTR нәтижесі дербес дәлелдеме көзіне жатпайды және өздігінен шешім қабылдауға жеткіліксіз |

Нәтиже арызды, түсініктемені немесе өзге де жазбаны жылдам оқуға көмектесуі мүмкін. Материалдарды іздеуге, жүйелеуге немесе электрондық іске енгізуге қолайлы жағдай жасай алады. Бірақ түпнұсқа құжаттың орнын

алмастырмайды және өзге дәлелдемелерден бөлек бағаланбайды. ENFSI-дің 2022 жылғы нұсқаулығында қайта тексеру жеке кезең ретінде қарастырылуы да осыған сәйкес. Автоматты түрде танылған мәтін процестік шешімге бастайтын көмекші құрал. Шешімнің дербес негізі емес.

Қолданыстағы нормативтік материалдарда автоматты түрде танылған мәтінді іс жүргізу құжатында белгілеудің, оның күмәнді бөліктерін жеке көрсетудің және тексерілген нұсқасын ресімдеудің бірыңғай егжей-тегжейлі тәртібі толық ашылмаған. Электрондық қылмыстық іс жүргізу нұсқаулығы электрондық ортаны ұйымдастыруға бағытталғанымен, танылған мәтіннің қай жері дәл, қай жері күмәнді екенін көрсетудің жеке үлгісін арнайы бекітпейді [24]. Бұл мәселе бойынша дерек жеткіліксіз.

2026 жылы Республикалық референдумда қабылданған Қазақстан Республикасының Конституциясы [4] мен қабылданған Цифрлық кодекс [27] осы тармақта қолданыстағы іс жүргізушілік реттеудің тікелей қайнар көзі ретінде емес, қылмыстық іс жүргізуді одан әрі цифрландыруға байланысты құқықтық дамудың алдағы бағытын сипаттайтын құжаттар ретінде ғана ескеріледі. Олар әлі күшіне енген жоқ. Сондықтан HTR нәтижелерін пайдаланудың қазіргі іс жүргізушілік шарттары мен шектеулері қолданыстағы ҚПК, электрондық қылмыстық іс жүргізу нұсқаулығы және ақпараттандыру мен дербес деректерді қорғау туралы заңнама негізінде айқындалады [23]-[26].

Талдау HTR нәтижесін қылмыстық процесте пайдалану тек техникалық алынуымен бітпейтінін көрсетеді. Мәтін түпнұсқамен салыстырылған кезде ғана мәнге ие болады. Күмән туғызатын бөліктері қайта тексерілмейінше, мазмұны процестік тұрғыдан жеткілікті деп қабылданбайды. Электрондық іске енгізілген жағдайда ол ақпараттық заңнамаға және дербес деректерді қорғау талаптарына бағынады. Даулы жағдайда арнайы білімді тарту қажеттігі туындайды. Шарттар сақталмаған жерде HTR нәтижесі қылмыстық процесте сенімді негіз ретінде пайдаланылмайды.

## 2.3 Қолжазба мәтіндерін тану нәтижелерін тексеру мен пайдаланудың практикалық тәртібі

Алдыңғы кіші бөлімдерде HTR нәтижесінің қылмыстық процестегі құқықтық табиғаты мен оны пайдаланудың іс жүргізушілік шектері айқындалды. Бұдан кейін мәселе теориялық емес, қолданбалы сипат алады. Тергеуші үшін шешуші сұрақ нәтиженің құқықтық мәртебесінде ғана емес. Онымен күнделікті жұмыстың қандай тәртіппен жүргізілетінінде. Сондықтан бұл кіші бөлімде HTR нәтижесін тексеру мен пайдаланудың практикалық тәртібі түпнұсқа құжатты сақтау, танылған мәтінді онымен салыстыру, күмәнді фрагменттерді қайта қарау және содан кейін ғана нәтижені көмекші ақпарат ретінде қолдану логикасы арқылы ашылады. ҚПК-нің 111, 125 және 126-баптары дәлелдемелердің заңды түрде алынуын, жиынтық тұрғыдан бағалануын және ғылыми-техникалық құралдарды қолдану мүмкіндігін бекітеді [23]. Нормалар HTR нәтижесін

процестік айналымнан шығармайды. Бірақ оны тексерусіз қабылдауға да жол бермейді [23].

Практикалық тәртіп түпнұсқа қолжазба құжатты айқындаудан басталады. Тергеуші алдымен құжаттың іс үшін маңызын белгілейді. Сосын бастапқы нысанын сақтайды. Электрондық қылмыстық іс жүргізу нұсқаулығы электрондық істі дайындауды, жүргізуді, жөнелтуді және сақтауды ұйымдастыруға арналғанын көрсетеді [24]. Бірақ автоматты түрде танылған мәтінді түпнұсқаның орнына жүретін дерек ретінде бекітпейді [24]. ENFSI-дің 2022 жылғы нұсқаулығында құжатпен жұмыс талдау, салыстыру, бағалау және қайта тексеру кезеңдері арқылы жүргізіледі [14]. Қазақстандық ведомстволық реттеу электрондық айналымды ұйымдастыруға басымдық береді, ал еуропалық сараптамалық үлгі тексерудің сатылылығын күшейтеді [14], [24]. Айырма практикалық тәртіптің өзегін айқындайды: цифрлық өңдеу түпнұсқаны ығыстырмайды – оны қайта қарауға қолайлы қосымша қабат жасайды [14], [24].

Содан кейін құжат цифрлық нысанға көшіріліп, машина оқи алатын мәтіндік нұсқа алынады. «Ақпараттандыру туралы» заңның 1-бабында автоматтандыру ақпаратты жасау, іздестіру, жинау, жинақтау, сақтау, өңдеу, алу және пайдалануды қамтитын үдеріс ретінде айқындалады [25]. HTR осы мағынада құқықтық қорытынды емес. Ол ақпаратты өңдеу тәсілі [25]. Практикалық тұрғыдан бұл кезең тергеушіге мәтінді оқуға, іздеуге және жүйелеуге ыңғайлы нұсқа береді. Бірақ алынған мәтін әлі де аралық өнім. Процестік маңызы келесі сатыда – түпнұсқамен салыстыру және мазмұнды тексеру арқылы ашылады.

Танылған мәтінді тексерудің негізгі бөлігі түпнұсқамен тікелей салыстыруға және күмәнді фрагменттерді қолмен қайта оқуға тіреледі. ENFSI-дің 2022 жылғы нұсқаулығында талдау мен салыстыру аяқталғаннан кейін ғана бағалау жүргізіледі; бағалау жеке белгіге емес, барлық анықталған ұқсастықтар мен айырмашылықтардың жиынтығына сүйенеді [14]. Сол құжатта қайта тексеру мен жазбаларды тіркеу жеке рәсім ретінде берілген [14]. NIST-тің 2020 жылғы есебі қолжазбаны зерттеу үдерісінде адами факторлардың, жұмыс жағдайының және міндетке қатысы жоқ мәнмәтіндік ақпараттың нәтижеге ықпал етуі мүмкін екенін талдайды [13]. Еуропалық нұсқаулық тексерудің құрылымын қатаңдайды. Американдық есеп сол тексеруді не үшін қажет екенін түсіндіреді [13], [14]. Практикалық қорытынды бір: танылған мәтіндегі күндер, сандар, тегі, мекенжайы, қысқартулар немесе істің мәніне әсер ететін сөздер түпнұсқаға қайта қарап нақтыланады [13], [14].

Салыстыру барысында мәтіннің оқылуы бірнеше нұсқаға бөлінсе, түпнұсқа мен танылған мәтін арасында елеулі алшақтық анықталса немесе қолжазба белгісінің авторлығы мен түпнұсқалығы процестік мәнге ие болса, арнайы білім иесі тартылады. Ресейлік материал қолжазба объектілерінің құқық қорғау айналымындағы нақты салмағын көрсетеді. ҚПК ғылыми-техникалық құралдармен қатар маман мен сарапшының қатысуына құқықтық негіз береді

[17], [23]. HTR нәтижесі даулы учаскеде тергеушінің техникалық пайымымен ғана шектелмейді [17], [23].

Кесте 5 – Тергеуші қызметінде қолжазба мәтіндерін тану нәтижелерін тексеру және пайдаланудың практикалық тәртібі

| Кезең | Тергеушінің әрекеті | Практикалық мақсаты | Іс жүргізушілік мәні |
|---|---|---|---|
| 1 | Қолжазба құжатты айқындау және түпнұсқасын сақтау | Құжаттың іс үшін маңызын белгілеу және бастапқы материалды жоғалтпау | Түпнұсқа кейінгі салыстыру мен тексерудің негізгі объектісі болып қалады |
| 2 | Құжатты цифрландыру және тануға дайындау | Құжатты машина оқи алатын өңдеуге жарамды күйге келтіру | Электрондық ортада жұмыс істеуге мүмкіндік береді, бірақ түпнұсқаны алмастырмайды |
| 3 | Қолжазба мәтінін тану нәтижесін алу | Мәтіннің цифрлық нұсқасын қалыптастыру | Нәтиже тек аралық және көмекші ақпараттық өнім ретінде қаралады |
| 4 | Танылған мәтінді түпнұсқамен салыстыру | Мәтіннің дұрыстығын тексеру | Тексерілмеген мәтінге іс жүргізушілік сенім берілмейді |
| 5 | Күмәнді фрагменттерді қолмен қайта қарау | Қате танылған не екіұшты бөліктерді нақтылау | Күні, тегі, мекенжайы, сома, қысқартулар сияқты маңызды бөліктер нақтыланады |
| 6 | Қажет болған жағдайда маман немесе сарапшыны тарту | Даулы учаскелерді арнайы білім арқылы түсіндіру | Танылған мәтін мен түпнұсқа арасында алшақтық болса, арнайы білім қолданылады |
| 7 | Танылған мәтінді көмекші ақпарат ретінде пайдалану | Мәтін бойынша іздеу, жүйелеу, іріктеу және кейінгі талдауға дайындау | Нәтиже дербес дәлелдеме емес, процестік шешімді қолдайтын қосымша құрал ретінде пайдаланылады |

Тексеріліп, қажет болғанда нақтыланғаннан кейін ғана танылған мәтін электрондық қылмыстық іс құрамында көмекші ақпарат ретінде пайдаланылады. Мұндай пайдалану мәтін бойынша іздеуге бейімдеуді, материалдарды жүйелеуді, ұқсас құжаттарды іріктеуді және кейінгі талдауға дайындауды қамтиды. Бірақ тағы бір шек сақталады. «Дербес деректер және оларды қорғау туралы» заңның 5-бабы жинау, өңдеу және қорғау қағидаттарын, 8-бабы келісім беру тәртібін

бекітеді [26]. Қолжазба арыздарда, түсініктемелерде және өзге де құжаттарда жәбірленуші, куә, күдікті және басқа тұлғалар туралы сәйкестендіретін мәліметтер жиі кездеседі [26]. Танылған мәтінді электрондық ортаға енгізу тек техникалық рәсім емес. Ол дербес деректерді өңдеу әрекеті де [25], [26]. Қазақстандық заңнама деректерді қорғау талабын қояды. NIST қолжазбаны зерттеуде артық мәнмәтіндік ақпаратты шектеу қажеттігін көрсетеді [13]. Біріншісі құқық иесін қорғайды. Екіншісі талдаудың бейтараптығын сақтайды [13], [26].

Тергеуші қызметінде HTR нәтижесін тексеру мен пайдаланудың практикалық тәртібі бірізді төрт өзекке жинақталады (кесте 5): түпнұсқа құжатты сақтау, танылған мәтінді онымен салыстыру, күмәнді бөліктерді қолмен қайта қарау және содан кейін ғана нәтижені көмекші ақпарат ретінде қолдану. Даулы жағдайда тәртіп арнайы білімді тартумен толықтырылады. HTR нәтижесі қылмыстық процесте дербес дәлелдеме ретінде емес – түпнұсқаға тәуелді, тексерілетін және верификациядан кейін ғана пайдаланылатын қосымша ақпарат ретінде қаралады.

# ҚОРЫТЫНДЫ

Жүргізілген зерттеу қолжазба құжаттардың ішкі істер органдары тергеушісінің қызметінде тұрақты кездесетінін және олардың тек мәтіндік ақпарат көзі емес, іс жүргізу тұрғысынан маңызы бар объект екенін көрсетті. Зерттелген материалдарда қолжазба құжаттарды өңдеу мәтінді оқумен бітпейтіні, оған сақтау, тіркеу, салыстыру, цифрландыру және түсіндіру әрекеттері кіретіні анықталды. Ruiz-Parrado және әріптестері 1990–2020 жылдар аралығындағы 5389 жарияланымды талдап, офлайн қолжазба құжаттарды өңдеудегі негізгі қиындықтарды жазу мәнерлерінің әркелкілігімен, тілдік айырмашылықтармен және динамикалық белгілердің жоғалуымен байланыстырған. Тергеуші үшін қолжазба материалмен жұмыс әрдайым көпсатылы ұйымдастырушылық және техникалық әрекеттерді талап етеді.

Қолжазба мәтінін тану технологияларының қолданбалы әлеуеті халықаралық және қазақстандық деректермен расталды. Nockels және әріптестері 2015–2020 жылдар аралығындағы 381 жарияланымды саралап, олардың 66,67%-ын гуманитарлық қолданбаларға, 25,46%-ын технологиялық зерттеулерге жатқызған. Тойганбаева және әріптестері сипаттаған KOHTD 3000 емтихан парағын, 140 335 сегменттелген кескінді және шамамен 922 010 таңбаны қамтыған. Халықаралық ортада технологияның платформалық орнығуы байқалады. Қазақстандық тәжірибеде нақты тілдік және графикалық база қалыптасқан. Бұл деректер HTR-ды абстрактілі жаңалық емес, нақты құжаттық материалмен жұмыс істеуге бейімделетін құрал ретінде бағалауға мүмкіндік береді.

Тергеуші қызметіндегі қолдану шегі осы технологиялардың ұйымдастырушылық-техникалық мәнімен айқындалды. Зерттеу нәтижелері оларды бастапқы цифрландыруға, мәтін бойынша іздеуге бейімдеуге, жүйелеуге, электрондық ортаға енгізуге және кейінгі талдауға дайындауға пайдалануға болатынын көрсетті. HKR дерекқорында 1400-ден астам толтырылған нысан, шамамен 63 000 сөйлем және 715 699-дан астам таңба қамтылған. KOHTD базасында жауаптардың 99%-ы қазақ тілінде және 1%-ы орыс тілінде. Бірінші мысал құрылымы шектеулі жазбалармен жұмыс істеуге бейімділікті көрсетеді. Екінші мысал еркін оқу мәтінін цифрландыру мүмкіндігін көрсетеді. Қорғауға ұсынылатын бірінші тұжырымның мазмұны осыдан туындайды: қолжазба мәтінін тану технологиялары ішкі істер органдары тергеушісінің қызметінде, ең алдымен, арыздарды, түсініктемелерді, қолхаттарды, жазбаларды және өзге де қолжазба құжаттарды бастапқы цифрландыру, мәтіндік іздеу, жүйелеу және электрондық есепке енгізу кезінде қолданылатын ұйымдастырушылық-техникалық құрал ретінде қарастырылады.

Сонымен бірге зерттеу бұл технологиялардың нәтижесі құқықтық бағалау мен дәлелдемелік қорытындыны алмастыра алмайтынын көрсетті. Abdallah және әріптестері 2020 қазақ және орыс тілдеріндегі деректерде бірінші сынақ жиынында таңба қатесінің деңгейін 4,13%, сөз қатесінің деңгейін 18,91%, сөйлем

қатесінің деңгейін 25,72% деп көрсеткен; екінші жиында бұл көрсеткіштер 6,31%, 23,69% және 35,16%. NIST-тің 2020 жылғы есебінде Srihari тобы 1500 адамның үлгілері бойынша жазуды орындаушыны айқындаудағы қателік деңгейі егіздер үлгілерінде 13%, егіз емес үлгілерде 4% болғанын келтіреді. Қазақстандық-ресейлік тілдік материалдағы модельдік нәтиже мен американдық салыстырмалы зерттеудегі қателік көрсеткіштері бір-біріне ұқсамайды. Бірақ екеуі де бір ортақ шекке әкеледі: автоматты нәтиже барлық жағдайда бірдей сенімді емес. Оны тексерусіз процессуалдық ақиқат ретінде қабылдауға болмайды.

Қылмыстық-процестік талдау шекті нақтылады. ҚПК-нің 111-бабы дәлелдемелерді заңды түрде алынған нақты деректер ретінде бекітеді, 125-бап оларды жиынтық тұрғыдан бағалауды талап етеді, 126-бап ғылыми-техникалық құралдарды қолдануға жол береді. Электрондық қылмыстық іс жүргізу нұсқаулығы танылған мәтінді электрондық ортаға енгізуге мүмкіндік береді. Бірақ оны түпнұсқа құжаттың орнын басатын дерек ретінде бекітпейді. «Ақпараттандыру туралы» заң автоматтандыруды ақпаратты өңдеу үдерісі ретінде сипаттайды. Дербес деректерді қорғау туралы заң қолжазба мәтінінде болатын сәйкестендіретін мәліметтерді өңдеудің шегін айқындайды. Құқықтық негіздер HTR нәтижесінің мәртебесін анық көрсетеді: ол дербес дәлелдеме көзі емес – түпнұсқа құжатқа тәуелді, тексерілетін және бағаланатын туынды ақпараттық өнім.

Практикалық тексеру тәртібі қорытындыны бекітті. ENFSI-дің 2022 жылғы нұсқаулығында талдау, салыстыру, бағалау және қайта тексеру кезеңдері бөлінеді. Еуропалық сараптамалық тәсіл тексерудің құрылымын айқындайды. Ресейлік дерек қолжазба объектілерінің құқық қорғау практикасындағы нақты салмағын көрсетеді. Қорғауға ұсынылатын екінші тұжырымның мазмұны осыдан қалыптасады: қолжазба мәтінін тану нәтижелері тергеуші үшін тек көмекші ақпараттық ресурс ретінде қаралады; олар қылмыстық іс бойынша процестік шешім қабылдаудың дербес негізі бола алмайды және түпнұсқа құжатпен салыстыруды, қате танылған бөліктерді қолмен тексеруді, ал даулы жағдайда маман не сарапшы қатысуымен растауды талап етеді.

Зерттеу барысында тағы бір қолданбалы шектеу анықталды. Қарастырылған материалдарда ішкі істер органдары тергеушілерінің нақты іс жүргізу құжаттарына арналған ашық репрезентативті корпус табылмады. Bias in History жобасында 100 000-нан астам бет тарихи полиция есептері негізінде деректер жиыны қалыптастырылғаны көрсетілген. Бірақ бұл тәжірибе тарихи мұрағаттық материалға жатады. Қазақстандық база білім беру материалдары негізінде жасалған. Құқық қолдану мазмұны бар құжаттар массивіне технологияны бейімдеу мүмкіндігі көрсетілгенімен, ішкі істер органдарының нақты іс жүргізу құжаттарына бейімделген ашық эмпирикалық база жөніндегі дерек шектеулі. Бұл мәселе бойынша дерек жеткіліксіз.

# ПАЙДАЛАНЫЛҒАН ДЕРЕККӨЗДЕРДІҢ ТІЗІМІ

1. Мемлекет басшысы Қасым-Жомарт Тоқаевтың 2025 жылғы 8 қыркүйектегі Қазақстан халқына Жолдауы «Жасанды интеллект дәуіріндегі Қазақстан: өзекті мәселелер және оны түбегейлі цифрлық өзгерістер арқылы шешу». [Электрондық ресурс] – Қолжетімділік режимі: https://adilet.zan.kz/kaz/docs/K25002025_1 (жүгінген күні: 25.05.2026).
2. Қазақстан Республикасы Президентінің 2026 жылғы 6 қаңтардағы № 1151 Жарлығы «Цифрландыру және жасанды интеллект жылын жариялау туралы». [Электрондық ресурс] – Қолжетімділік режимі: https://adilet.zan.kz/kaz/docs/U2600001151 (жүгінген күні: 20.05.2026).
3. 2025 жылдың қорытындылары: қазақстандықтар мемлекеттік қызметтерді смартфон арқылы жиі ала бастады. [Электрондық ресурс] – Қолжетімділік режимі: https://www.gov.kz/memleket/entities/aktobe-irgiz/press/news/details/1134935 (жүгінген күні: 25.04.2026).
4. Қазақстан Республикасының Конституциясы. 2026 жылғы 15 наурызда республикалық референдумда қабылданды. [Электрондық ресурс] – Қолжетімділік режимі: https://adilet.zan.kz/kaz/docs/K2600000000 (жүгінген күні: 25.04.2026).
5. Қазақстан Республикасы Бас прокуратурасының Құқықтық статистика және арнайы есепке алу жөніндегі комитеті. Құқық қорғау органдары қызметінің негізгі көрсеткіштері туралы мәліметтер. 2025 жыл. [Электрондық ресурс] – Қолжетімділік режимі: https://qamqor.gov.kz/crimestat/files/2025/12/19/202512_osnpok_ru.xlsx (жүгінген күні: 25.04.2026).
6. Қазақстан Республикасы Бас прокуратурасының Құқықтық статистика және арнайы есепке алу жөніндегі комитеті. Қылмыстық құқық бұзушылықтар туралы арыздарды, хабарламаларды және өзге де ақпаратты қарау туралы есеп. 2023 жыл. [Электрондық ресурс] – Қолжетімділік режимі: https://qamqor.gov.kz/crimestat/files/2023/12/19/202312_kui_00300___ru.xlsx (жүгінген күні: 25.04.2026).
7. Қазақстан Республикасы Бас прокуратурасының Құқықтық статистика және арнайы есепке алу жөніндегі комитеті. Қылмыстық құқық бұзушылықтар туралы арыздарды, хабарламаларды және өзге де ақпаратты қарау туралы есеп. 2024 жыл. [Электрондық ресурс] – Қолжетімділік режимі: https://qamqor.gov.kz/crimestat/files/2024/12/19/202412_kui_00300___ru.xlsx (жүгінген күні: 25.04.2026).
8. Қазақстан Республикасы Бас прокуратурасының Құқықтық статистика және арнайы есепке алу жөніндегі комитеті. Қылмыстық құқық бұзушылықтар туралы арыздарды, хабарламаларды және өзге де ақпаратты қарау туралы есеп. 2025 жыл. [Электрондық ресурс] – Қолжетімділік режимі: https://qamqor.gov.kz/crimestat/files/2025/12/19/202512_kui_00300___ru.xlsx (жүгінген күні: 25.04.2026).

9. Ruiz-Parrado V., Heradio R., Aranda-Escolastico E. et al. A bibliometric analysis of off-line handwritten document analysis literature (1990–2020) // Pattern Recognition. – 2022. – Vol. 125. – Article 108513. DOI: 10.1016/j.patcog.2021.108513.
10. Nockels J., Gooding P., Ames S., Terras M. Understanding the application of handwritten text recognition technology in heritage contexts: a systematic review of Transkribus in published research // Archival Science. – 2022. – Vol. 22. – № 3. – P. 367–392. DOI: 10.1007/s10502-022-09397-0.
11. Toiganbayeva N., Kasem M., Abdimanap G. et al. KOHTD: Kazakh Offline Handwritten Text Dataset [Электрондық ресурс] // arXiv. – 2021. – Қолжетімділік режимі: https://arxiv.org/abs/2110.04075 (қаралған күні: 25.04.2026).
12. Nurseitov D., Bostanbekov K., Kurmankhojayev D. et al. HKR For Handwritten Kazakh & Russian Database [Электрондық ресурс] // arXiv. – 2020. – Қолжетімділік режимі: https://arxiv.org/abs/2007.03579 (қаралған күні: 25.04.2026).
13. Expert Working Group for Human Factors in Handwriting Examination. Forensic Handwriting Examination and Human Factors: Improving the Practice Through a Systems Approach. [Электрондық ресурс] – Gaithersburg: National Institute of Standards and Technology, 2020. – NISTIR 8282. – Қолжетімділік режимі: https://nvlpubs.nist.gov/nistpubs/ir/2020/NIST.IR.8282.pdf (жүгінген күні: 28.04.2026).
14. ENFSI. Best Practice Manual for the Forensic Handwriting Examination. [Электрондық ресурс] – 2022. – Edition 04. – Қолжетімділік режимі: https://enfsi.eu/wp-content/uploads/2023/02/BPM-Handwriting-Ed.-4.pdf (жүгінген күні: 03.05.2026).
15. Bias in History. Past activities. [Электрондық ресурс] – Қолжетімділік режимі: https://www.bias-in-history.eu/past-activities/ (жүгінген күні: 07.05.2026).
16. Abdallah A., Hamada M., Nurseitov D. Attention-based fully gated CNN-BGRU for Russian handwritten text // Journal of Imaging. – 2020. – Vol. 6. – P. 141.
17. Бобовкин М.В., Ручкин В.А. Судебно-почерковедческая экспертиза в правоохранительной деятельности Российской Федерации // Правовая парадигма. – 2018. – Т. 17. – № 2. – С. 63–70.
18. Аратұлы Қ., Тапалова Р.Б., Даубасова С.Ш., Төлегенова А.А. Құжаттарды криминалистикалық зерттеудің құрылымы туралы // ҚазҰУ Хабаршысы. Заң сериясы. – 2020. – № 2 (94). – Б. 113–122. DOI: 10.26577/JAPJ.2020.v94.i2.13.
19. Deviterne-Lapeyre M. Interpol review of questioned documents 2016–2019 // Forensic Science International: Synergy. – 2020. – Vol. 2. – P. 429–441. DOI: 10.1016/j.fsisyn.2020.01.012.
20. Garrido-Munoz C., Rios-Vila A., Calvo-Zaragoza J. Handwritten Text Recognition: A Survey [Электрондық ресурс] // arXiv. – 2025. – Қолжетімділік режимі: https://arxiv.org/abs/2502.08417 (қаралған күні: 25.04.2026).
21. Богатенкова А.О., Беляева О.В., Перминов А.И. Генерация изображений рукописного текста на русском языке // Труды Института системного программирования РАН. – 2023. – Т. 35. – Вып. 2. – С. 19–34. DOI: 10.15514/ISPRAS-2023-35(2)-2.

22. Скрыпников А.В., Денисенко В.В., Хитров Е.Г., Евтеева К.С., Савченко И.И. Распознавание рукописного текста с использованием нейронных сетей // Современные наукоемкие технологии. – 2021. – № 6-1. – С. 91–95. DOI: 10.17513/snt.38703.
23. Қазақстан Республикасының Қылмыстық-процестік кодексі. Қазақстан Республикасының Кодексі 2014 жылғы 4 шілдедегі № 231-V ҚРЗ. [Электрондық ресурс] – Қолжетімділік режимі: https://adilet.zan.kz/kaz/docs/K1400000231 (жүгінген күні: 25.04.2026).
24. «Қылмыстық сот ісін электрондық форматта жүргізу жөніндегі нұсқаулықты бекіту туралы» Қазақстан Республикасы Бас Прокурорының 2018 жылғы 3 қаңтардағы № 2 бұйрығы. Қазақстан Республикасының Әділет министрлігінде 2018 жылғы 23 қаңтарда № 16268 болып тіркелді. [Электрондық ресурс] – Қолжетімділік режимі: https://adilet.zan.kz/kaz/docs/V1800016268 (жүгінген күні: 30.04.2026).
25. «Ақпараттандыру туралы» Қазақстан Республикасының Заңы. Қазақстан Республикасының Заңы 2015 жылғы 24 қарашадағы № 418-V ҚРЗ. [Электрондық ресурс] – Қолжетімділік режимі: https://adilet.zan.kz/kaz/docs/Z1500000418 (жүгінген күні: 27.04.2026).
26. «Дербес деректер және оларды қорғау туралы» Қазақстан Республикасының Заңы. Қазақстан Республикасының 2013 жылғы 21 мамырдағы № 94-V Заңы. [Электрондық ресурс] – Қолжетімділік режимі: https://adilet.zan.kz/kaz/docs/Z1300000094 (жүгінген күні: 26.04.2026).
27. Қазақстан Республикасының Цифрлық кодексі. Қазақстан Республикасының Кодексі 2026 жылғы 9 қаңтардағы № 255-VIII ҚРЗ. [Электрондық ресурс] – Қолжетімділік режимі: https://adilet.zan.kz/kaz/docs/K2600000255 (жүгінген күні: 28.04.2026).

# ҚОСЫМША 1

Сурет. Қолжазба мәтіндерін тану технологияларының тергеушінің ұйымдастырушылық-техникалық қызметіндегі қолдану моделі



# ҚОСЫМША 2

Кесте. Қолжазба мәтіндерін тану технологияларын қолдану бағыттары және олардың тергеу қызметін ұйымдастыруға ықпалы

| Қолдану бағыты | Мазмұны | Тергеуші қызметіне ықпалы | Шектеуі |
|---|---|---|---|
| Бастапқы цифрландыру | Қолжазба құжатты сканерленген бейнеден мәтіндік форматқа көшіру | Құжатпен электрондық жұмыс істеуге мүмкіндік береді | Сканерлеу сапасы мен қолжазба анықтығына тәуелді |
| Мәтіндік іздеу | Танылған мәтін ішінен сөздер мен фрагменттерді табу | Көп материал ішінен қажетті ақпаратты жылдам табуға көмектеседі | Қате танылған сөздер іздеу нәтижесін бұрмалайды |
| Жүйелеу | Құжаттарды мазмұндық немесе тақырыптық белгі бойынша бөлу | Материалдар массивін реттеуге жағдай жасайды | Автоматты бөлу тергеушінің тексеруін қажет етеді |
| Электрондық есепке енгізу | Қолжазба деректерді электрондық қорда сақтауға жарамды форматқа келтіру | Құжаттарды қайта табу, сүзу және салыстыру жеңілдейді | Нәтиже бірден дәлелдеме ретінде қолданылмайды |
| Кейінгі талдауға дайындау | Цифрлық мәтіндік қор қалыптастыру | Ірі көлемдегі қолжазба материалдарды талдауға негіз болады | Нақты тергеу құжаттары бойынша ашық деректер жеткіліксіз |

ҚОСЫМША 3

**ҚАЗАҚСТАН РЕСПУБЛИКАСЫ БАС ПРОКУРАТУРАСЫНЫҢ ЖАНЫНДАҒЫ ҚҰҚЫҚ ҚОРҒАУ АКАДЕМИЯСЫ**

ІШКІ ІСТЕР ОРГАНДАРЫ ТЕРГЕУШІСІНІҢ ҚЫЗМЕТІНДЕ ҚОЛЖАЗБА МӘТІНДЕРІН ТАНУ ТЕХНОЛОГИЯЛАРЫН ҚОЛДАНУДЫҢ АЛГОРИТМІ

**Қосшы**
**2026 жыл**

Кіріспе

Алгоритм қолжазба мәтіндерін тану технологияларын ішкі істер органдары тергеушісінің қызметінде қолданудың практикалық тәртібін және құқықтық шектерін айқындау үшін әзірленді. Негізі – ҚР Қылмыстық-процестік кодексі [1], электрондық қылмыстық сот ісін жүргізу нұсқаулығы [2], «Ақпараттандыру туралы» [3] және «Дербес деректер және оларды қорғау туралы» заңдар [4].

HTR жүйелері сөз және жол деңгейінен абзац пен құжат деңгейіне дейін дамыған. Garrido-Munoz, Rios-Vila және Calvo-Zaragoza 2025 жылы мұндай жүйелердің тұтас құжат құрылымын өңдеуге бағытталғанын көрсетті [5]. Қазақстандық тәжірибеде Тойганбаева және әріптестері сипаттаған KOHTD деректер жиыны 3000 емтихан парағын, 140 335 бөлшектелген кескінді және шамамен 922 010 таңбаны қамтыған [6]. Бір дерек технологияның жалпы даму деңгейін көрсетеді. Екіншісі оның нақты оқу материалында жұмыс істегенін дәлелдейді.

Алгоритм екі бөліктен тұрады. Бірінші бөлім – тергеушінің HTR нәтижесімен тікелей жұмыс істеу реті. Екінші бөлім осы тәртіптің құқықтық және әдістемелік негіздерін жинақтайды.

Пайдаланылған қысқартулар

ҚР ҚПК – Қазақстан Республикасының Қылмыстық-процестік кодексі
ҚР ҚК – Қазақстан Республикасының Қылмыстық кодексі
HTR – қолжазба мәтінін тану технологиясы
Электрондық қылмыстық іс – электрондық форматта жүргізілетін қылмыстық іс

# 1-БӨЛІМ. ТЕРГЕУШІНІҢ ҚОЛЖАЗБА МӘТІНДЕРІН ТАНУ НӘТИЖЕЛЕРІМЕН ЖҰМЫС ІСТЕУІНІҢ ПРАКТИКАЛЫҚ ТӘРТІБІ

## 1.1. Қолжазба құжатты айқындау және бастапқы материалды сақтау

Тергеуші ең алдымен қолжазба мәтіні бар құжаттың іс үшін маңызын белгілейді. Мұндай құжатқа арыз, түсініктеме, қолхат, хат, жеке жазба немесе өзге де қолмен толтырылған мәтін жатады.

Бастапқы материалдың сақталуы шешуші мәнге ие. Қолжазбаның мазмұны ғана емес, графикалық нысаны да маңызды. Қолтаңба, түзету іздері, қысқартулар, жазу бағыты, әріптердің өзара байланысы – бұлар кейінгі салыстырып тексеруге қажет болуы мүмкін. Сондықтан тану технологиясын қолданғанға дейін түпнұсқа құжат бөлек материалдық объект ретінде сақталады.

## 1.2. Құжатты цифрландыру және тануға дайындау

Келесі кезеңде құжат цифрлық нысанға көшіріледі. Бейненің сапасы, анықтылығы және тануға жарамдылығы тексеріледі.

KOHTD деректер жиынында кескінді алдын ала өңдеу, бұрылуын түзету, жолдарға және сөздерге бөлу әрекеттері қолданылған [6]. Испаниялық шолу құжат деңгейіндегі тану күрделілігі артқан сайын бастапқы өңдеудің маңызы да өсетінін көрсетеді [5]. Қазақстандық оқу тәжірибесі бұл кезеңді нақты техникалық қадам ретінде көрсетеді. Шетелдік шолу оның жалпы әдістемелік маңызын ашады.

## 1.3. Қолжазба мәтінін тану нәтижесін алу

Цифрландырудан кейін HTR жүргізіледі. Нәтиже – түпнұсқа құжат емес, машина оқи алатын мәтіндік нұсқа.

Құжат көлемі ұлғайған сайын мұндай нәтиженің ұйымдастырушылық пайдасы артады. Ол бірнеше бетті материалды жедел оқуға, мәтін бойынша іздеуге және кейінгі талдауға дайындауға мүмкіндік береді. Бірақ алынған мәтін бастапқы құжаттың орнын алмастырмайды. Ол тек көмекші цифрлық қабат [5].

## 1.4. Танылған мәтінді түпнұсқамен салыстыру

Техникалық танудан кейін бірден түпнұсқамен салыстыру жүргізіледі. Бұл – практикалық тәртіптің өзегі.

ENFSI-дің 2022 жылғы нұсқаулығында қолжазбамен жұмыс талдау, салыстыру, бағалау және қайта тексеру арқылы жүргізіледі [7]. NIST-тің 2020 жылғы есебі автоматты жүйелердің қолжазбаны зерттеу үдерісін қолдауға арналғанын талдайды [8]. Еуропалық тәсіл салыстырудың сатылылығын күшейтеді. Американдық тәсіл техникалық көмектің шегін көрсетеді. Танылған мәтін түпнұсқамен тексерілмейінше, оны әрі қарай пайдалануға жол берілмейді.

## 1.5. Күмәнді фрагменттерді қолмен тексеру

Мәтіннің жекелеген бөліктері бұрмаланып танылса, оларды қолмен қайта қарау қажет. Бүкіл мәтінді қайта көшіру емес – күмән туғызатын бөліктерді нақтылау.

KOHTD тәжірибесі тану сапасы жазу мәнеріне, әріп ұқсастығына және материал сапасына тәуелді екенін көрсетеді [6]. Даталар, тегі, мекенжайы, қысқартулар және мағынаны өзгертетін сөздер тергеуші тарапынан міндетті түрде қайта оқылады.

1.6. Қажет болған жағдайда маман немесе сарапшыны тарту

Мәтіннің оқылуы даулы болса, түпнұсқа мен танылған мәтін арасында алшақтық байқалса немесе қолжазба белгісінің авторлығы мен түпнұсқалығы мәселе туғызса, арнайы білім иесі тартылады.

Бобовкин мен Ручкин 2018 жылы қолжазбатану сараптамасында зерттелетін объектілердің 65%-ы қолтаңбалар, 30%-ы қысқа жазбалар, ал мұндай сараптамалардың 78%-ы алдын ала тергеу сатысында жүргізілетінін көрсеткен [9]. Сандар қолжазба объектілерінің құқық қорғау практикасында тұрақты кездесетінін айтады. Мұндай жағдайда маман не сарапшыны тарту процестік жағынан негізді.

1.7. Нәтижені көмекші ақпарат ретінде пайдалану

Тексерілгеннен кейін танылған мәтін көмекші ақпарат ретінде пайдаланылады. Ол құжатты электрондық ортаға енгізуге, мәтін бойынша іздеуге бейімдеуге, материалдарды жүйелеуге және кейінгі талдауға дайындауға жарайды.

Nockels және әріптестері 2022 жылы мұрағаттық және кітапханалық ортада HTR бастапқы дереккөздерді көшіруді жеделдетіп, толық мәтін бойынша іздеуге және ірі көлемдегі талдауға жағдай жасайтынын көрсетеді [10]. Қазақстандық оқу тәжірибесінде KOHTD деректер жиыны алгоритмдерді оқыту мен салыстыруға арналған стандартталған база ретінде беріледі [6]. Шетелдік мұрағаттық тәжірибе мәтіндік іздеудің пайдасын көрсетеді. Қазақстандық оқу материалы тану нәтижесін құрылымдауға болатынын дәлелдейді. Тергеуші қызметінде де осы шек сақталады: танылған мәтін құжатпен жұмысты жеңілдетеді, бірақ дербес дәлелдеме ретінде пайдаланылмайды.

Практикалық тәртіп HTR нәтижесімен жұмыс істеудің негізгі логикасын айқындайды. Тергеуші бастапқы құжатты сақтайды. Кейін оны цифрландырады. Техникалық тану нәтижесін алады. Оны түпнұсқамен салыстырады. Күмәнді бөліктерді қолмен қайта тексереді. Қажет болғанда маман немесе сарапшыны тартады. Содан кейін ғана танылған мәтін құжат айналымын жеңілдететін көмекші ақпарат ретінде пайдаланылады.

# 2-БӨЛІМ. ҚОЛЖАЗБА МӘТІНДЕРІН ТАНУ ТЕХНОЛОГИЯЛАРЫН ҚОЛДАНУДЫҢ ҚҰҚЫҚТЫҚ ЖӘНЕ ӘДІСТЕМЕЛІК НЕГІЗДЕРІ

## 2.1. Нормативтік-құқықтық негіздер

HTR нәтижесін пайдалану қолданыстағы қылмыстық-процестік және ақпараттық заңнама шегінде жүргізіледі. ҚР ҚПК 111-бабы дәлелдемелерді заңды түрде алынған нақты деректер ретінде айқындайды. 125-бап оларды қатыстылығы, жол берілетіндігі және анықтығы тұрғысынан бағалауды талап етеді. 126-бап ғылыми-техникалық құралдарды қолдануға жол береді [1]. Нормалар тану нәтижесін іс материалымен жұмыс істеуді жеңілдететін техникалық өңдеу өнімі ретінде пайдалануға мүмкіндік береді. Бірақ оны дербес дәлелдеме көзіне теңестірмейді.

Электрондық қылмыстық сот ісін жүргізу нұсқаулығы электрондық істі дайындау, жүргізу, жөнелту және сақтау тәртібін белгілейді [2]. «Ақпараттандыру туралы» заңда автоматтандыру ақпаратты жасау, жинау, сақтау, өңдеу және пайдалану үдерісі ретінде айқындалады [3]. Бірінші құжат электрондық айналымды ұйымдастырады. Екіншісі ақпаратты өңдеудің жалпы құқықтық шегін белгілейді. HTR нәтижесі құқықтық қорытынды емес — өңделген ақпарат.

Дербес деректерді қорғау туралы заңнама нәтижені қолданудың қосымша шегін көрсетеді. Қолжазба арыздарда, түсініктемелерде және өзге де құжаттарда адамдарды сәйкестендіретін мәліметтер жиі болады. Мұндай мәтінді машина оқи алатын нысанға көшіру тек техникалық қадам емес. Ол дербес деректерді өңдеу әрекеті де [4].

## 2.2. Қолданудың іс жүргізушілік шектері мен тәуекелдері

HTR нәтижесінің практикалық маңызы тексеру сапасына тәуелді. ENFSI-дің 2022 жылғы нұсқаулығында қолжазбамен жұмыс талдау, салыстыру, бағалау және қайта тексеру арқылы жүргізіледі [7]. Танылған мәтін түпнұсқамен салыстырылмайынша, оның мазмұнына іс жүргізушілік сенім берілмейді.

NIST-тің 2020 жылғы есебі автоматты жүйелердің шегін сандармен көрсетеді. Srihari тобы 1500 адамнан жиналған үлгілер бойынша жүргізген салыстыруда жазуды орындаушыны айқындау қателігінің деңгейі егіздер үлгілерінде 13%, ал егіз емес үлгілерде 4% болған [8]. Тану нәтижесі барлық жағдайда бірдей сенімді емес. Күмән туғызатын сөздер мен екіұшты оқылатын бөліктер қайта тексеріледі.

Даулы жағдайларда арнайы білімнің маңызы артады. Сандар қолжазба объектілерінің шеткері емес екенін айтады. Танылған мәтін мен түпнұсқа арасында алшақтық байқалса немесе мәтін бірнеше түрде оқылса, маман не сарапшы тартылады.

Тағы бір шек – танылған мәтінді жалғыз негізге айналдырмау. ҚПК дәлелдемелерді жиынтық тұрғыдан бағалауды талап етеді [1]. Танылған мәтін шешім қабылдауды жеңілдетуі мүмкін. Бірақ жалғыз тірегі бола алмайды.

Қолданыстағы нормативтік материалдарда автоматты түрде танылған мәтінді іс материалында белгілеудің, оның күмәнді бөліктерін жеке көрсетудің және тексерілген нұсқасын ресімдеудің бірыңғай егжей-тегжейлі тәртібі толық ашылмаған. Бұл мәселе бойынша дерек жеткіліксіз.

2.3. Қорытындылар мен ұсынымдар

Жоғарыда қарастырылған деректер HTR нәтижесін тергеуші қызметінде пайдаланудың нақты шегін көрсетеді. Мұндай нәтиже түпнұсқа құжатпен салыстырылған кезде ғана мәнге ие. Күмән туғызатын бөліктері қолмен қайта тексеріледі. Даулы жағдайда арнайы білім иесі тартылады. Танылған мәтін электрондық ортада айналымға ене алады. Бірақ дербес дәлелдеме емес және процестік шешімнің жалғыз негізі ретінде пайдаланылмайды.

Практикалық ұсынымдар бірнеше бағытта шоғырланады. Түпнұсқа құжат міндетті түрде сақталады. Танылған мәтін онымен салыстырылып тексеріледі. Күмәнді бөліктер қолмен нақтыланады, қажет болғанда маман немесе сарапшы тартылады. Электрондық іске енгізілген мәтінмен жұмыс істеу кезінде ақпараттық заңнама мен дербес деректерді қорғау талаптары қатаң сақталады. Тәсіл HTR-ды тергеуші қызметінде көмекші, реттелген құрал ретінде пайдалануға мүмкіндік береді.

Пайдаланылған әдебиеттер тізімі

1. Қазақстан Республикасының Қылмыстық-процестік кодексі. Қазақстан Республикасының Кодексі 2014 жылғы 4 шілдедегі № 231-V ҚРЗ. [Электрондық ресурс] – Қолжетімділік режимі: https://adilet.zan.kz/kaz/docs/K1400000231 (жүгінген күні: 20.02.2026).
2. «Қылмыстық сот ісін электрондық форматта жүргізу жөніндегі нұсқаулықты бекіту туралы» Қазақстан Республикасы Бас прокурорының 2018 жылғы 3 қаңтардағы № 2 бұйрығы. Қазақстан Республикасының Әділет министрлігінде 2018 жылғы 23 қаңтарда № 16268 болып тіркелді. [Электрондық ресурс] – Қолжетімділік режимі: https://adilet.zan.kz/kaz/docs/V1800016268 (жүгінген күні: 25.02.2026).
3. «Ақпараттандыру туралы» Қазақстан Республикасының Заңы. Қазақстан Республикасының Заңы 2015 жылғы 24 қарашадағы № 418-V ҚРЗ. [Электрондық ресурс] – Қолжетімділік режимі: https://adilet.zan.kz/kaz/docs/Z1500000418 (жүгінген күні: 10.03.2026).
4. «Дербес деректер және оларды қорғау туралы» Қазақстан Республикасының Заңы. Қазақстан Республикасының 2013 жылғы 21 мамырдағы № 94-V Заңы. [Электрондық ресурс] – Қолжетімділік режимі: https://adilet.zan.kz/kaz/docs/Z1300000094 (жүгінген күні: 05.03.2026).
5. Garrido-Munoz C., Rios-Vila A., Calvo-Zaragoza J. Handwritten Text Recognition: A Survey [Электрондық ресурс]. 2025. arXiv:2502.08417.
6. Тойганбаева N., Kasem M., Abdimanap G., Bostanbekov K., Abdallah A., Alimova A., Нурсеитов D. KOHTD: Kazakh Offline Handwritten Text Dataset [Электрондық ресурс]. 2021. arXiv:2110.04075.
7. ENFSI. Best Practice Manual for the Forensic Handwriting Examination. Edition 04. ENFSI-BPM-FHX-01. 2022.
8. Expert Working Group for Human Factors in Handwriting Examination. Forensic Handwriting Examination and Human Factors: Improving the Practice Through a Systems Approach. NISTIR 8282. U.S. Department of Commerce, National Institute of Standards and Technology, 2020.
9. Бобовкин М.В., Ручкин В.А. Судебно-почерковедческая экспертиза в правоохранительной деятельности Российской Федерации // Правовая парадигма. 2018. Т. 17. № 2. С. 63-70. DOI: 10.15688/lc.jvolsu.2018.2.10.
10. Nockels J., Gooding P., Ames S., Terras M. Understanding the application of handwritten text recognition technology in heritage contexts: a systematic review of Transkribus in published research // Archival Science. 2022. Vol. 22. No. 3. P. 367-392. DOI: 10.1007/s10502-022-09397-0.

# ҚОСЫМША 4

«БЕКІТЕМІН»
Қазақстан Республикасы ІІМ
Тергеу департаменті
Аймақтық бақылау
басқармасының орынбасары
_____С.Ж.Хамитбеков
2026 жылғы «21» наурыз

Қазақстан Республикасы Бас прокуратурасының жанындағы
Құқық қорғау органдары академиясының магистранты
Амиралиев Нұртас Қасенұлының
«Ішкі істер органдары тергеушісінің қызметінде қолжазба мәтіндерін тану технологияларын қолданудың перспективалары мен іс жүргізу шектеулері» тақырыбы бойынша зерттеу нәтижелерін енгізу актісі

Комиссия құрамы:
төраға — 1) ҚР ІІМ Тергеу департаментінің Аймақтық бақылау басқармасының орынбасары С.Ж.Хамитбеков

және комиссия мүшелері — 2) Аймақтық бақылау басқармасының аса маңызды істер жөніндегі аға тергеушісі Данияр Калдыбекович Нысанбаев

3) Аймақтық бақылау басқармасының аса маңызды істер жөніндегі аға тергеушісі Қайрат Серікбайұлы Касенов

«Ішкі істер органдары тергеушісінің қызметінде қолжазба мәтіндерін тану технологияларын қолданудың перспективалары мен іс жүргізу шектеулері» магистрлік жобаның нәтижелері Ішкі істер органдары тергеушісінің қызметінде қолжазба мәтіндерін тану технологияларын қолданудың алгоритмі түрінде Қазақстан Республикасы Ішкі істер министрлігі Тергеу департаментінің практикалық қызметіне енгізілгені туралы осы актіні жасады.

Енгізу күні: 2026 жылғы «21» наурыз

Төраға: 1) _____ (Қолы) С.Ж.Хамитбеков

Комиссия мүшелері: 2) _____ (Қолы) Д.К.Нысанбаев

3) _____ (Қолы) Қ.С.Касенов