



**МЕМЛЕКЕТТІК ТІЛДЕГІ ЗАҢШЫҒАРМАШЫЛЫҚ ПРАКТИКАСЫНАН/
ИЗ ПРАКТИКИ ЗАКОНОТВОРЧЕСТВА НА ГОСУДАРСТВЕННОМ ЯЗЫКЕ/
FROM THE LAWMAKING PRACTICE IN THE OFFICIAL LANGUAGE**

ҚАЗАҚСТАН РЕСПУБЛИКАСЫНЫҢ
ЗАҢНАМА ЖӘНЕ ҚҰҚЫҚТЫҚ АҚПАРАТ
ИНСТИТУТЫНЫҢ ЖАРШЫСЫ
ҒЫЛЫМИ-ҚҰҚЫҚТЫҚ ЖУРНАЛЫ
ISSN 2788-5283
eISSN 2788-5291
ТОМ 81, НӨМІРІ 1(2026), 313-321

ӘӨЖ 34.096
FTAMP 10.01.33
DOI 10.52026/2788-5291_2026_81_1_313
Ғылыми мақала

© А.Т. Ирубаева¹, Т.Е. Нургазиев^{2*}, 2026

^{1,2} Қазақстан Республикасының Заңнама және құқықтық ақпарат институты, Астана, Қазақстан
(e-mail: ¹iruaitan@mail.ru; ²tanat.nurgaziev@mail.ru)

**ҚҰҚЫҚТЫҚ МӘТІНДЕРДІ САРАПТАУДА ЖӘНЕ АУДАРУДА
ЖАСАНДЫ ИНТЕЛЛЕКТ ПЕН САНДЫҚ КОРПУСТАРДЫ
ПАЙДАЛАНУ МҮМКІНДІГІ МЕН ТӘУЕКЕЛДЕРІ**

Аннотация. Нормативтік-құқықтық актілер көлемінің ұлғайуы, мемлекеттік және орыс тілдерінде қатарласып заң шығару үдерісінің пайда болуы және мемлекеттік құрылымдардың жасанды интеллектті пайдалана отырып автоматтандырылған жүйеге көшуі құжаттарға лингвистикалық сараптама жүргізудің жаңашыл ұстанымдары мен техникасының өзгеруін талап етіп отыр. Бұл ретте жасанды интеллект пен қарқынды дамып келе жатқан дерекқорлар (database) өзара кіріктірілген (интеграцияланған) сипат алып, мәтіндерді аудару және салыстырып талдау құралына айналуға.

Құқықтық маңызы бар заңнама мен заңға қатысы бар өзге құжаттар құқықтық дискурстың негізін құрап, өздігінен нақтылық пен бірегейлік тұтастығының көрінісі болып тұруы тиіс. Алайда жасанды интеллект пен дерекқорларда сақталған кей ұғымдардың аудармасы сөйлем мағынасына үйлеспей, кейде қазақ тіліне жат құрылым сипатын алып, дәл сол қалпында заңнамалық күші бар құжаттардың айналымға түсіп, қолданысқа еніп жатқанын аңғаруға болады.

Зерттеу мақаласында жасанды интеллект негізінде аударылған мәтіндердің лексикалық, стилистикалық, грамматикалық, орфографиялық, пунктуациялық және заң техникасына қатысты қателері бес негізгі топқа жіктеліп қарастырылған. Мақаладағы лексикалық қателердің қатарында автоматты аударма жасау барысында түсіп қалған сөздер, қате аударылған терминдер мен сөз тіркестері; стилистикалық қателер ретінде орын тәртібі ауысып кеткен сөздер мен тіркестер; грамматикалық қателер аясында орынсыз (үндестік заңы сақталмай) жалғанған жұрнақтар мен жалғаулар; орфографиялық және пунктуациялық сәйкессіздіктер; заң техникасына қатысты белгіленген стандарттардың сақталмауы сөз болады.

Қазақстан Республикасында құқық саласына қатысты құжаттарды аудару және аудармаға сараптама жүргізуде жасанды интеллект құралдарын қолданудың мүмкіндіктері мен тәуекелдері сараланады. Заң құжаттарын орыс тілінен мемлекеттік тілге автоматты түрде аудару кезінде жиі кездесетін қателіктер талданып, сөз тіркестері мен жекелеген сөздердің аударма сипатындағы баламалы нұсқалары ұсынылады.

Түйінді сөздер: құқық; сараптама; аударма; жасанды интеллект; сандық корпус; тәуекелдер.

Кіріспе

Ғылым мен технологияның даму кезеңінде жасанды интеллект (ЖИ) ғылым, білім беру, зерттеу әдістерін түрлендіру, үлкен көлемдегі деректерді өңдеу барысында кеңінен қолданылып, мемлекеттік деңгейдегі кейбір үрдістерді автоматтандыру мақсатында қызмет етіп жатыр. ЖИ ұсынып отырған есептеу дәлдігі, жылдамдық пен әмбебаптылық статистикалық мәліметтерді өңдеу, графикалық дизайн, психологиялық көмек

көрсету сияқты қызметтерді табысты атқарып, технологияның ілгерілеу сипатын айқындап жатыр. Дәл осы негізде бұл тәжірибе аударма саласында да өз орнын тауып, мемлекеттік маңызы бар құжаттарды бір тілден екінші тілге аудару сипатында көрініс табуда. Дегенмен, үлкен көлемді сандық ресурстармен (электрондық нұсқадағы сөздіктер, порталдар, тіл корпустары) интеграциялық сипатта дамып, жетіліп келе жатқан ЖИ қазақ тілінің синонимдік, лексикалық қорының

* хат-хабарларға арналған автор. E-mail: tanat.nurgaziev@mail.ru.

негізіндегі автоматты аударма жасау, сөзді таңдау және сөйлемнің морфологиялық сипатын сақтап орын тәртібін айқындау, үндестік заңы ережелерін қадағалау, стилдік (әсіресе, ғылыми, публицистикалық, ресми іс-қағаздар) ерекшеліктерді ескере отырып аудару қабілеті әлі де болса жеткіліксіз. Бұл қателіктер заңдық күші бар құжаттарда жиі орын алып, бүгінгі күні аударманы талдау және сараптама жүргізу жұмысының басты мәселесіне айналды. Нормативті құжаттарды талдауда кездесетін кейбір заң терминдері баламалы сөздермен алмастырылғандықтан құжатаралық байланыс әлсіреп, аударма құжаттың заңдық күшіне қатысты және екі тілдегі сәйкестікті қамтамасыз ете алу мүмкіндігі жайлы заңгер мамандар тарапынан жиі талқыға түсіп жатады. Мысалы, жасанды интеллект моделдері кейде «жеке кәсіпкер» сөзін «дара кәсіпкер» деп аударып жатады. Бұл қатемен берілетін аудармалар «жеке» және «дара» сөзқолданысының тіркесті формада берілу барысында синонимдік қатарды дұрыс таңдай алмау себебінен болып отырады. Бұл жағдай – бұл тіркесті сөздің сандық корпуста жоқ екенін айғақтайды. Сәйкесінше, сандық корпустардағы дерекқорларға (database) талдау жүргізу және ЖИ көмегімен аударылған мәтіндердегі қателерге сараптама жасап, себебін анықтау, зерттеу мақаласының басты мақсаты болып табылады. Дерекқорлар қатарында бүгінгі күні сандық форматқа өткен «Қазақ тілінің ұлттық корпусы (ҚТҰК)», «Синонимдер сөздігі», «Антонимдер сөздігі», «Фразеологиялық сөздік», «Этимологиялық сөздік», «Sozdik.kz», т.б. корпустар бар. Бұл қатарда заң терминдерінің корпусы сияқты аса қажет болып отырған электрондық сөздік базасының болуы – заман талабы.

Құқық саласындағы сөздер мен сөз тіркестерін жүйелеу, орыс, ағылшын тілдеріндегі сәйкестігін анықтау, таңдалған лексиканың функционалдық мүмкіндіктерін саралау, жасанды интеллект арқылы автоматтандырылған аударманың тәуекелдері мен оңтайлы тұстарын анықтап алу, талдау және қажетті ұсынымдарды әзірлеу жүйелі жүргізілетін әдістердің кешенін талап етеді. Құқықтану жайлы сөз қозғалғанда ең бірінші назарда болуы тиіс

мәселе – бұл үрдістің заңдылық деңгейін анықтау және зерделеу. 2025 жылдың 17 қарашасында Қазақстан Республикасының «Жасанды интеллект туралы» Заңы қабылданды. Бұл Заң «Егемен Қазақстан» газетінің № 222 (31202) санында ресми түрде жарияланды. Заң Қазақстан Республикасындағы сандық технологияларды дамыту мен бақылаудың құқықтық негіздерін белгілей отырып, жасанды интеллект жүйелерінің деректерді жинау, талдау мәселелеріне баса назар аударды. Аталған Заңның 23-бабы жасанды интеллектінің авторлық құқық мәселесіне қатысын реттейді. Заңның 21-бабының 3-тармағында қандай да бір жасанды интеллект құрастырған синтетикалық өнім болған жағдайда пайдаланушыны ақпараттандыру жүргізу керек екені анықтайды. Алайда, Заң аясында автоматтандырылған аударма мәселесі қарастырылмайды.

Жасанды интеллектті дамытудың 2024-2029 жылдарға арналған тұжырымдамасында аударма жасау мәселесі 4-бағыт, 2-тармағында «Қазақ тілінің ауқымды тілдік моделі (KazLLM)» ретінде көрсетілген. Тұжырымдама аясында қарастырылған «Тілмаш» жүйесі әлі де болса дамуды қажет етіп отырғанын байқауға болады. Сондықтан мақаладағы зерттеу әдістерін таңдау барысында терминологиялық сөздіктердің семантикалық, құрылымдық ерекшеліктерін, сандық корпустардың базалық әлеуеті және жасанды интеллект моделдері өзара салыстырмалы түрде қарастырылып, ерекшеліктері ескеріледі.

Материалдар мен әдістер

Мемлекеттік маңызы бар құжаттарды аударудағы терминологиялық базаны саралау, құқықтық семантика мәселесіне мән бере отырып лингвистикалық нормалардың логикалық және құрылымдық бірізділігін сақтаудың жолын қарастыру, сандық сипаттағы сөздік корпустарға талдау жүргізу, автоматты аудармаларға сараптама жасау, ақпаратты жинақтау, саралау, қорыту және статистикалық есебін шығару зерттеу мақаласының негізгі әдістемелік базасын құрайды. Бұл ретте материал ретінде жасанды интеллект тарапынан (GPTChat, Grok, Deepseek)

аударылған үлгілік мәтіндердің орысша және қазақша нұсқалары лексикалық, стилистикалық, грамматикалық, орфографиялық, пунктуациялық және заң техникасына сәйкес талдау жүргізу мақсатында семантикалық тұтастық, тезаурустық салыстыру, контекстуалды, дистрибутивтік талдау, графемалық және морфемалық салыстыру, құқықтық-логикалық талдау, терминологиялық сараптама әдістері қолданылады. Қазақстан Республикасының «Жасанды интеллект туралы» Заңы, «Жасанды интеллектті дамытудың 2024-2029 жылдарға арналған тұжырымдамасы» талдамалық (аналитикалық) әдіс арқылы талданады.

Нәтижелер мен талқылау

Құқықтық мәтіндерді сараптауда және аударуда жасанды интеллект пен сандық корпустарды пайдалану мүмкіндігі орасан зор. «Қазіргі таңда әлемнің көптеген мемлекеттері сандық технологияларды құқықтық жүйеге енгізу арқылы әртүрлі деңгейдегі табыстарға қол жеткізуде» [1, 384-б.]. Жасанды интеллектінің ең бірінші артықшылығы – аудармадағы құрылымдық-логикалық байланысты жүйелі беруі. Мысалы, заңнамалық актілердің, құжаттардың жүйелі берілетін ой тәртібі мен сөйлем құрылымы бар. Дерекқорларда сақталған электронды құжаттар негізінде үйренген, жаттыққан ЖИ моделі үшін техникалық сәйкестікті сақтай отырып аудару

промтын (тапсырма беру) дұрыс беру арқылы жүзеге асырылады. Мысалы, жүйелі аударма жасау үшін GPTChat жасанды интеллектісіне келесі промтты жазу ұсынылады: «Осы мәтінді орыс тілінен қазақ тіліне аудар. Нормативтік-құқықтық актілерге тән құрылымдық логиканы, терминдердің бірізділігін және құқықтық стильдің дәлдігін сақта. Баптар, тармақтар, тармақшалар арасындағы мағыналық байланыстарды бұзбай, ресми-іскери стильде аудар. Терминдерге балама берме, тек құқықтық қолданыста қалыптасқан нұсқасын пайдалан. Егер орыс тіліндегі сөйлемде артық баяндау немесе екіұштылық болса, оны қазақ тіліндегі құқықтық тіл нормаларына сай нақты әрі түсінікті етіп бер». Осы промттан кейін GPTChat мәтінді максималды түрде жүйемен аударуды көздеп, біршама дұрыс аударманы ұсынады.

Алайда кейбір сөздерді (сөйлемдерді) аудару барысында жасанды интеллектінің жиі қате таңдау жасайтынын да байқауға болады. Бұл қазақ тілінің тілдік құрылымында синонимдік қатардың көптігімен түсіндіріліп, мағыналық тұрғыда ЖИ терминология саласындағы сөздерді аударуда әлсіздік танытатынын көрсетеді.

Мысалы, төменде салыстырмалы түрде GPTChat көмегімен қазақ тілінен орыс тіліне аударылған мәтіннің суреттері берілген.

ИИ обеспечивает:

- автоматическую терминологическую сверку;
- согласование синтаксических структур;
- учёт контекста в несколько абзацев (что недоступно статистическому переводу).

ИИ қамтамасыз етеді:

- автоматты терминологиялық тексеріс
- синтаксистік құрылымдардың сәйкестендірілуі
- бірнеше абзацтық контексті ескеру (бұл статистикалық аудармада жоқ)

Сурет 1. GPTChat көмегімен аударылған мәтін

Аударма мәтінде төмендегі қателіктер орын алған:

Орыс тілінен қазақ тіліне жасалған аударма	Қазақ тіліндегі дұрыс аударма	Қатенің түрі
ИИ	ЖИ (жасанды интеллект)	Қысқарған сөзді аударудағы қателік.
терминологиялық тексеріс	терминологиялық сәйкестендіру	Етістікті аударудағы қателік.
бұл	статистикалық аудармада жоқ	Артық сілтеу есімдігінің қолданылуы.

Кесте 1. GPTChat көмегімен аударылған мәтіндегі қателіктер

Талдау кестесінен байқалғандай, аудармаға берілген 16 сөздің 3 сөзі (сөз тіркесі) қате аударылып тұр. Есептік көрсеткіште бұл аударманың 18,7% құрайды. Кейде, жасанды интеллект көмегімен аударылған мәтіндерде бағалау сипатындағы сөздер, әрекет

субъектісін белгілеп тұрған сөз орамдары жиі түсіп қалып отырады. Нәтижесінде аударма мәтінінде жасырын, кейде екіұшты мағына пайда болады.

Дәл осы мәтінді GrokAI жасанды интеллектісі келесі суретте көрсетілген нұсқа бойынша аударды:

Жасанды интеллект қамтамасыз етеді:

- автоматты терминологиялық сәйкестендіру;
- синтаксистік құрылымдардың үйлесімділігі;
- бірнеше абзацқа созылатын контексті ескеру (бұл статистикалық аудармаға қолжетімсіз).

Балама, сәл басқа стильде:

ИИ мыналарды қамтамасыз етеді:

- терминдерді автоматты түрде сәйкестендіру;
- сөйлем құрылымдарының үйлесімді болуын;
- бірнеше абзац көлеміндегі контексті ескеру (статистикалық аудармада мұндай мүмкіндік жоқ).

Сурет 2. GrokAI көмегімен аударылған мәтін

GrokAI жасанды интеллектісі аударған мәтін (Сурет 2), GPTChat ұсынған аудармаға қарағанда (Сурет 1) әлдеқайда дәл әрі нақты берген. Алайда баламалы нұсқа ретінде берген екінші мәтінінің бірінші тармағында стилистикалық сәйкессіздіктер орын алған. Сонымен бірге лексикалық қателердің қатарында «Бұл статистикалық аудармаға

қолжетімсіз» сөйлемінде «мүмкіндік» сөзі түсіп қалғанын аңғаруға болады.

Deepseek жасанды интеллектісінің V3 моделі бұл мәтінді GPTChat 5.2 немесе GrokAI 3 моделдеріне қарағанда дұрыс форматта аударып, формалық ауытқудың автоматтандырылған аудармаған тән сипатын байқатады.

Сәлеметсіз бе! Мына мәтінді сіз үшін аударамын:

Жасанды интеллект келесі мүмкіндіктерді ұсынады:

- терминологияны автоматты түрде сәйкестендіру;
- синтаксистік құрылымдарды үйлестіру;
- бірнеше абзац көлеміндегі контексті есепке алу (бұл мүмкіндік статистикалық аудармаға қол жетімсіз).

Сурет 3. Deepseek көмегімен аударылған мәтін

Жасанды интеллект түрлерінің (GPTChat 5.2, GrokAI 3, Deepseek V3) аудармаларын салыстыру барысында

№ р/с	Жіберген қатенің түрі	GPTChat 5.2	GrokAI 3	Deepseek V3
1.	Лексикалық	3	1	0
2.	Стилистикалық	2	1	1
3.	Грамматикалық	0	0	0
4.	Орфография және пунктуация	0	0	0
Барлық қате саны:		5	3	1

Кесте 2. Аудармалардағы қателерге салыстырмалы талдау



Сурет 4. Салыстырмалы қателер кестесі

Салыстырмалы кестеден ЖИ көмегімен аударылған мәтіндерде орфографиялық, пунктуациялық, грамматикалық қателер өте сирек кездесетіні анықталды. Бұл жасанды интеллектінің сөздерді дайын ресурстардан сөздікке сай алатынын және сингармонизм заңдылықтарын (үндестік заңын) математикалық дәлдікпен қолданатынын айғақтайды. GPTChat 5.2 және GrokAI 3 ЖИ моделдері лексикалық және стилистикалық қателерді көп жіберетіні анықталды. Лексикалық және стилистикалық қателер орынсыз сөз тастап кету немесе тіркес түрінде қолданылмайтын сөздерді тіркестіру сипатында көрініс табады.

«Тұжырымдамалық деңгейде заң шығару процесіне жасанды интеллектті енгізу келесі элементтерді қарастыруы керек: тиісті Стратегияны әзірлеу; басымдықтарды белгілеу; оқыту; енгізу процесінің өзі; басқару. Заң шығару процесіне жасанды интеллектті толыққанды енгізу үшін бірыңғай оператор құру қажеттілігі бар. Бұл оператор жасанды интеллект құралдарын әзірлейді, тиімді енгізеді және құқықтың әрбір саласы бойынша оқытуды жүзеге асырады» [2]. Алайда заңнамалық актілер мен өзге де мемлекеттік маңызы бар құжаттарды ЖИ көмегімен аудару барысында міндетті түрде аударма сапасын тексеріп, мәтінге күрделі сараптама жасап отыру қажет. Себебі, «Адамзат мүддесі үшін жасанды интеллект әлеуетін іске асыру технологиялық және этикалық аспектілерді қамтитын кешенді тәсілді қажет етеді. Тиімді ғана емес, сонымен

қатар қауіпсіз және бейтарап жасанды интеллект алгоритмдерін құру және дамыту қажет» [3,66-б.]. Бұл ретте, әуелде «жасанды интеллект» ұғымының заңдық күші мен авторлыққа ие болу құқығын айқындау қажеттілігі туындап отыр. «Қазақстан Республикасындағы жасанды интеллектті құқықтық реттеу мәселесін қарастырғанда, «жасанды интеллект» термині заң актілерінде қолданылғанымен, оған жеке анықтама берілмегені байқалады» [4,42-б.]. Бүгінгі күні Қазақстан Республикасының «Жасанды интеллект туралы» Заңы, «Жасанды интеллектті дамытудың 2024-2029 жылдарға арналған тұжырымдамасы» аясында айналымға енгізілген ЖИ синтетикалық өнімі заңи тұрғыда жалпылама беріліп, әлі де болса әр сала бойынша нақтылануы тиіс. Бұл ретте мәтіндердің автоматты түрде аудару мәселесі де тысқары қалмай, ғылыми жұмыстарда, мақала жазуда, кәсіби аударма жасауда тәуекелдер мен нақты шектеулер заң тұрғысынан реттелуі тиіс шара. Жасанды интеллектті дамытудың 2024-2029 жылдарға арналған тұжырымдамасында аударма жасау мәселесі 4-бағыт, 2-тармағында «Қазақ тілінің ауқымды тілдік моделі (KazLLM)» жасауды көздеп отыр. Ауқымды модель тілдік корпусстардан жасалатынын ескерсек, заң терминдерінің тілдік корпусын жасау, басқа тілдердегі аудармалары мен баламаларын беру заң тілінің сапасын жоғарылатып, автоматтандырылған сапалы аударма жүйесіне әкелетін бірден-бір құрал бола алады.

Жасанды интеллект моделін дамыту үшін әуелде бастапқы қор (немесе сандық корпус) мәселесін айқындап алған дұрыс. Сандық корпус¹ «Қазақ тілінің ұлттық корпусы» аясында қалыптасып, бүгінгі күні сандық жүйеге түскен 90 млн сөз қолданысын қамтитын қорға айналды. Корпус өз алдына 20 негізгі ішкорпуста құрастырылған. Бұл корпустардың қатарында заң шеңберінде қолданылатын құқықтық терминологияны әлі де дамыту қажеттілігін айғақтайды. Бұл ретте синонимдік қатар мен кей заңнамалық сөз қолданыстарын сандық форматқа көшіріп, сандық корпустарды (database) сөз ғана емес, жекелеген ғылым салаларындағы сөз тіркестері формасында жетілдіру жұмысының жүргізілуін қажет етеді.

Қорытынды

«Тәуелсіздік алғаннан кейін қазақ тілінің әлеуметтік-экономикалық, ғылыми-техникалық және жалпы мәдениет саласындағы қызметі өсіп, қолдану ауқымы мейілінше кеңейе түсті» [5, 53-б.]. Құқықтық мәтіндерді аудару және аударма мәтінді сараптау жұмысы бүгінгі күні ерекше маңызды сипат алып жатыр. Бұл қажеттілік мемлекеттік құрылымдардағы электронды үкімет

платформалары және жеке тұлғаларға көрсетілетін қызметтердің онлайн форматқа көшуімен түсіндіріледі.

Алайда құжат дайындау, аудару және тиісті адресатқа жіберудің нақты шекті уақыты болғандықтан, мемлекеттік құрылымдар жиі жасанды интеллект құралдарын қолданатынын байқауға болады.

Жасанды интеллект өз алдына кейбір термин сөздерді, морфологиялық тұтастықты сақтамай аударарды. Бұл заңнамалық құжаттарды аудару барысында үлкен мәселені туындатады.

Жасанды интеллект электронды ресурстардан, сандық корпус атауына ие дерекқорлардан (database) алған дайын сөздік (тіркестік) бірліктердің материалдарын қолдана отырып аударма жасайды. Сондықтан кез келген ЖИ көмегімен жасалған аударма міндетті түрде сараптамадан өткізіліп, түзету жасауды қажет етеді. Сонымен қатар күн санап дамып келе жатқан тілдің сандық корпусы жүйеленіп, әсіресе терминология саласында қазақ тілінің құқық саласындағы заңға қатысты сөздіктер электронды форматқа көшіріліп, жасанды интеллект тарапынан меңгеріле алатындай жалпы қолданысқа енгізілуі тиіс.

Әдебиеттің тізімі:

1. Журунова М.Қ., Исабай Н., Ибраев А.С. Жер қатынастары саласындағы цифрлық реттеу мен заңнаманы бағалаудың құқықтық аспектілері // Қазақстан Республикасының Заңнама және құқықтық ақпарат институтының Жаршысы. Ғылыми-құқықтық журнал. 2025. Т.80. №4. С.381–397. https://doi.org/10.52026/2788-5291_2025_80_4_381.
2. Қазақстанда заңдарды жасанды интеллект әзірлейтін болады: миф немесе шындық // Парламентаризм институтының вебсайты. Сілтеме арқылы қол жетімді: <https://parlaminst-mtqb.qr-pib.kz/p/32375> (қаралған күні: 15.05.2024).
3. Жалтырбаева Р.С., Джангабулова А.Қ., Тілембаева Ж.О. Жасанды интеллект ұғымының құқықтық дефинициясы: теориялық негіздер мен заңнамалық тұжырымдар // Қазақстан Республикасының Заңнама және құқықтық ақпарат институтының Жаршысы. Ғылыми-құқықтық журнал. 2025. Т.80. №4. https://doi.org/10.52026/2788-5291_2025_80_4_35.
4. Копежанов И.Ж. Жасанды интеллект: қиындықтар мен мүмкіндіктер // Arts Academy. 2024. Т.10. №2. С.65–73. <https://doi.org/10.56032/2523-4684.2024.2.10.64>.
5. Абдуова Б.С., Төлешева Г.Қ. Қазақ тілінің мәртебесі – ұлт болашағына бағдар. М.Қозыбаев атындағы Солтүстік Қазақстан университетінің Хабаршысы. 2022. №1 (53):53-58. <https://doi.org/10.54596/2309-6977-2022-1-53-58>.

¹ «Қазақ тілінің ұлттық корпусы (ҚТҰК) – қазақ тілінің лексика-грамматикалық жүйесін толық қамтыған (терең аннотацияланған) миллиондаған сөзқолданыстан тұратын электронды пішіндегі көлемді мәтіндер жинағы, қазақ тілі жайлы барлық мәліметті жинақтаушы «ақылды», мамандандырылған білім қоры. Жалпы сөзқолданыс саны – 90 000 000». URL: <https://qazcorpus.kz/> (қаралған күні: 14.01.2026 ж.).

© А.Т. Ирубаева¹, Т.Е. Нургазиев², 2026

^{1,2} Институт законодательства и правовой информации Республики Казахстан, Астана, Казахстан
(e-mail: ¹iruaitan@mail.ru; ²tanat.nurgaziev@mail.ru)

ВОЗМОЖНОСТИ И РИСКИ ИСПОЛЬЗОВАНИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА, ЦИФРОВЫХ КОРПУСОВ ПРИ ЭКСПЕРТИЗЕ И ПЕРЕВОДЕ ПРАВОВЫХ ТЕКСТОВ

Аннотация. Рост объёма нормативных правовых актов, появление параллельного законотворчества на государственном и русском языках, а также переход госструктур к автоматизированным системам с использованием искусственного интеллекта требуют пересмотра подходов и техники лингвистической экспертизы документов. В такой ситуации искусственный интеллект и стремительно развивающиеся базы данных постепенно превращаются в единый инструмент для перевода и сравнительного анализа текстов.

Законодательные и иные правовые документы, обладающие юридической значимостью, формируют основу правового дискурса и должны демонстрировать точность и смысловую цельность. Но нередко можно увидеть, что переводы некоторых терминов и выражений, извлечённых из ИИ моделей или корпусов, не вписываются в контекст, создают для казахского языка нетипичные конструкции и в таком виде попадают в оборот как документы имеющие юридическую силу.

В статье рассматриваются пять основных групп ошибок встречающихся в текстах переведённых с помощью искусственного интеллекта: лексические, стилистические, грамматические, орфографические, пунктуационные и связанные с нарушением требований юридической техники. К лексическим ошибкам относятся пропуски слов, неверно переведённые термины и словосочетания. К стилистическим – переставленные местами слова и выражения. В грамматическом плане встречаются неправильно присоединённые суффиксы и окончания, не соответствующие законам сингармонизма. Отдельно отмечаются орфографические и пунктуационные несоответствия, а также нарушения установленных стандартов технического оформления.

Также анализируются возможности и риски применения инструментов искусственного интеллекта при переводе и экспертизе правовых документов в Казахстане. Отдельное внимание уделено типичным ошибкам, возникающим при автоматическом переводе документов с русского на государственный язык, и приводятся более корректные варианты передачи отдельных слов и словосочетаний.

Ключевые слова: право; экспертиза; перевод; искусственный интеллект; цифровой корпус; риски.

© А.Т. Irubayeva¹, T.E. Nurgaziyev², 2026

^{1,2} Institute of Legislation and Legal Information of the Republic of Kazakhstan, Astana, Kazakhstan
(e-mail: ¹iruaitan@mail.ru; ²tanat.nurgaziev@mail.ru)

POTENTIALS AND RISKS OF USING ARTIFICIAL INTELLIGENCE AND DIGITAL CORPORA IN THE EXPERT ANALYSIS AND TRANSLATION OF LEGAL TEXTS

Abstract. The growing volume of normative legal acts, the emergence of parallel lawmaking in the Kazakh and Russian languages and the transition of government institutions to automated systems powered by artificial intelligence necessitate a revision of the approaches and methods used in linguistic expertise of legal documents. Artificial intelligence and rapidly expanding databases are gradually turning into an integrated tool for translating and comparatively analyzing texts.

However, it is often observed that translations of certain terms and expressions generated by AI models or extracted from corpora do not fit the contextual structure of the target language. Such translations produce atypical constructions for Kazakh and nevertheless enter official circulation as documents with legal force.

The article identifies five major groups of errors found in AI translated texts: lexical, stylistic, grammatical, orthographic, punctuation-related, and those associated with violations of legal drafting requirements. Lexical errors include omitted words, inaccurately rendered terms and incorrectly translated collocations. Stylistic errors involve misplaced words and expressions. In grammatical terms, typical issues include incorrectly attached suffixes and endings that contradict the principles of vowel harmony. Orthographic and punctuation inconsistencies, as well as deviations from established technical formatting standards are examined separately.

The study also analyzes the opportunities and risks of applying artificial intelligence tools in the translation and linguistic expertise of legal documents in Kazakhstan. Special attention is given to common errors that occur in automatic translations from Russian into the state language, with more accurate variants of selected words and expressions proposed where appropriate.

Keywords: law; expertise; translation; artificial intelligence; digital corpus; risks.

References:

1. Zhurunova M.K., Isabai N., Ibraev A.S. Zher katynastary salasyn-dagy cifrlyk retteu men zaңnamany bagalau-dyn kuckytyk aspektileri // Kazakstan Respublikasynyn Zannama zhane kuckytyk aqparat institutynyn Zharsysy. Gylymi-kuckytyk zhurnal. 2025. T.80. №4. S.381–397. https://doi.org/10.52026/2788-5291_2025_80_4_381.

2. Kazakstanda zaңdardy zhasandy intellekt azirleitin bolady: mif nemese shyndyk // Parlamentarizm institutynyn veb-saity. Rezhim dostupa: <https://parlaminst-mtqb.qr-pib.kz/p/32375>.

3. Zhaltyrbaeva R.S., Zhangabulova A.K., Tilembaeva Zh.O. Zhasandy intellekt ugymynyn kukyqtyk definiciyasi: teoriyalyk negizder men zaңnamalyk tuzhyrymdar // Kazakstan Respublikasynyn Zannama zhane kukyqtyk aqparat institutynyn Zharsysy. Gylymi-kukyqtyk zhurnal. 2025. T.80. №4. https://doi.org/10.52026/2788-5291_2025_80_4_35.

4. Kopezhanov I.Zh. Zhasandy intellekt: kiyindyktar men mumkindikter // Arts Academy. 2024. T.10. №2. S.65–73. <https://doi.org/10.56032/2523-4684.2024.2.10.64>.

5. Abduova B.S., Tolesheva G.K. Kazak tilinin martebesi – ult bolashagyna bagdar. M.Kozybaev atyndagy Soltustik Kazakstan universitetinin Khabarshysy. 2022. № 1 (53). p.53–58. <https://doi.org/10.54596/2309-6977-2022-1-53-58>.

Авторлар туралы мәліметтер:

Ирубаева Айгуль Танабуговна – заң ғылымдарының кандидаты, Қазақстан Республикасы Заңнама және құқықтық ақпарат институты Лингвистика орталығының жетекшісі, Жеңіс даңғылы, 15А, 010000, Астана, Қазақстан.

ORCID: <https://orcid.org/0009-0006-4205-9416>;

Scopus ID: 57195918216;

E-mail: iruaitan@mail.ru.

Нургазиев Танат Есенбаевич – хат-хабарларға арналған автор, PhD докторы, Қазақстан Республикасы Заңнама және құқықтық ақпарат институты Лингвистика орталығының аға ғылыми қызметкері, Жеңіс даңғылы, 15А, 010000, Астана, Қазақстан.

ORCID: <https://orcid.org/0000-0003-0470-7962>;

Scopus Author ID: 57191033650;

E-mail: tanat.nurgaziev@mail.ru.

Алғыс. Авторлар сарапшылық пікірі мен сындарлы көзқарасы үшін рецензенттерге алғыс білдіреді.

Дәйексөз келтіру үшін. Ирубаева А.Т., Нургазиев Т.Е. Құқықтық мәтіндерді сараптауда және аударуда жасанды интеллект пен сандық корпустарды пайдалану мүмкіндігі мен тәуекелдері // Қазақстан Республикасының Заңнама және құқықтық ақпарат институтының Жаршысы. Ғылыми-құқықтық журнал. 2026;81(1): 313-321. DOI – https://doi.org/10.52026/2788-5291_2026_81_1_313.

Авторлардың қосқан үлесі:

Ирубаева А.Т. – зерттеудің тұжырымдамалық негізін қалыптастырып, деректерді жинау, жүйелеу және бастапқы талдау жұмыстарын жүргізді. Сондай-ақ жұмыстың негізгі әдіснамалық бөлімін әзірлеп, мақаланың негізгі тұжырымдарын ғылыми-талдамалық тұрғыда негіздеді.

Нургазиев Т.Е. – жекелеген нәтижелерге аналитикалық өңдеу жүргізіп, библиографиялық дереккөздердің бір бөлігін іріктеу және рәсімдеу жұмыстарын орындады. Сонымен қатар кіріспе мен қорытынды бөлімдерді дайындап, мәтінді жан-жақты ғылыми редакциялауды және түйінді тұжырымдарды қалыптастыруды жүзеге асырды.

Мүдделер қақтығысы туралы ақпарат. Авторлар мүдделер қақтығысының жоқтығын туралы мәлімдейді.

Қаржыландыру көзі. Авторлар зерттеу жүргізу кезінде қаржыландырудың жоқтығы туралы мәлімдейді.

Мақала редакцияға келіп түсті: 29.01.2026; рецензиялаудан кейін келіп түсті: 26.02.2026; басып шығаруға қабылданды: 31.03.2026.

Авторлар қолжазбаның соңғы нұсқасын оқып, мақұлдады.

Сведения об авторах:

Ирубаева Айгуль Танабуговна – кандидат юридических наук, руководитель Лингвистического центра Института законодательства и правовой информации Республики Казахстан, проспект Женис, 15А, 010000, Астана, Казахстан.

ORCID: <https://orcid.org/0009-0006-4205-9416>;

Scopus ID: 57195918216;

E-mail: iruaitan@mail.ru.

Нургазиев Танат Есенбаевич – автор для корреспонденции, доктор PhD, старший научный сотрудник Лингвистического центра Института законодательства и правовой информации Республики Казахстан, проспект Женис, 15А, 010000, Астана, Казахстан.

ORCID: <https://orcid.org/0000-0003-0470-7962>;

Scopus Author ID: 57191033650;

E-mail: tanat.nurgaziev@mail.ru.

Благодарности. Авторы выражают благодарность рецензентам за их экспертное мнение и конструктивные замечания.

Для цитирования. Ирубайева А.Т., Нургазиев Т.Е. Возможности и риски использования искусственного интеллекта, цифровых корпусов при экспертизе и переводе правовых текстов // Вестник Института законодательства и правовой информации Республики Казахстан. Научно-правовой журнал. 2026;81(1): 313-321. DOI – https://doi.org/10.52026/2788-5291_2026_81_1_313.

Вклад авторов.

Ирубайева А.Т. – сформировала концептуальную основу исследования, провела сбор, систематизацию и первичный анализ данных. Разработала основную методологическую часть работы и обосновала ключевые положения статьи в научно-аналитическом формате.

Нургазиев Т.Е. – выполнил аналитическую обработку отдельных результатов, осуществил отбор и оформление части библиографических источников. Подготовил разделы введения и заключения, провёл комплексную научную редактуру текста и сформировал основные итоговые положения.

Информация о конфликте интересов. Авторы заявляют об отсутствии конфликта интересов.

Источник финансирования. Авторы сообщают об отсутствии финансирования при проведении исследования.

Статья поступила в редакцию: 29.01.2026; поступила после рецензирования: 26.02.2026; принята в печать: 31.03.2026.

Авторы прочитали и одобрили окончательный вариант рукописи.

Information about the authors:

Irubayeva Aigul Tanabugovna – Candidate of Legal Sciences, Head of the Linguistics Center at the Institute of Legislation and Legal Information of the Republic of Kazakhstan, 15A, Zhenis Avenue, 010000, Astana, Kazakhstan.

ORCID: <https://orcid.org/0009-0006-4205-9416>;

Scopus ID: 57195918216;

E-mail: iruaitan@mail.ru.

Nurgaziyev Tanat Esenbayevich – corresponding author, PhD, Senior Researcher at the Linguistics Center of the Institute of Legislation and Legal Information of the Republic of Kazakhstan, 15A, Zhenis Avenue, 010000, Astana, Kazakhstan.

ORCID: <https://orcid.org/0000-0003-0470-7962>;

Scopus Author ID: 57191033650;

E-mail: tanat.nurgaziyev@mail.ru.

Acknowledgements. The authors express their gratitude to the reviewers for their expert insights and constructive comments.

For citation. Irubayeva A.T., Nurgaziyev T.E. Potentials and risks of using artificial intelligence and digital corpora in the expert analysis and translation of legal texts // Bulletin of Institute of Legislation and Legal Information of the Republic of Kazakhstan. Scientific and legal journal. 2026;81(1): 313-321. DOI – https://doi.org/10.52026/2788-5291_2026_81_1_313.

Contribution of the authors:

Irubayeva A.T. – developed the conceptual framework of the study, conducted the collection, systematization and initial analysis of the data. She also designed the main methodological section of the work and substantiated the key propositions of the article in a scholarly analytical format.

Nurgaziyev T.E. – performed the analytical processing of selected results, carried out the selection and formatting of part of the bibliographic sources. He prepared the introduction and conclusion sections, conducted comprehensive scientific editing of the text, and formulated the main final statements.

Conflict of interest statement. The authors declares the absence of any conflict of interest.

Funding. The authors reports that the study received no financial support.

Received: 29.01.2026; revised: 26.02.2026; accepted for publication: 31.03.2026

The authors has read and approved the final manuscript.