

Два языка, два стандарта: ИИ и языковое неравенство в государственных услугах Казахстана

Жаннат Бубекбаева

Независимая исследовательница, Казахстан

АННОТАЦИЯ

Казахоязычные пользователи электронных государственных услуг обслуживаются хуже, чем русскоязычные и этот разрыв весьма существенный несмотря на то, что русскоязычное население составляет демографическое меньшинство страны. Этот факт разительно контрастирует с высокими достижениями Казахстана в области цифровизации государственных услуг: в Индексе развития электронного правительства ООН страна занимает 24-е место в мире.

Приложение eGov Mobile, запущенное в октябре 2019 года, насчитывает более 11,7 миллиона зарегистрированных пользователей, из которых около 6,4 миллиона обращаются к нему ежемесячно. Платформа использует ИИ-ассистентов, помогающих гражданам получать государственные услуги — от оформления удостоверяющих документов до регистрации бизнеса и оформления социальных льгот.

В этой статье исследуется, насколько равнозначно этот ИИ функционирует на двух наиболее распространённых языках Казахстана. Результаты свидетельствуют о явном неравенстве. Сравнительный анализ ИИ-ассистента eGov Mobile показал, что казахоязычный бот нередко опускает важную практическую информацию и выдаёт тексты, больше похожие на машинный перевод, а не живой язык. В одном из выявленных в ходе исследования случаев бот назвал сумму, превышающую верное значение примерно в 17 раз.

Русскоязычный бот в среднем работает лучше, однако и он не отличается надёжностью. Ему также свойственны серьёзные ошибки: в частности, он делает отсылки к документам и учреждениям Российской Федерации, не имеющим никакого отношения к Казахстану.

Всё это — не просто технические неудобства. Когда государственный ИИ-ассистент сообщает гражданину неверный размер штрафа, умалчивает о стоимости регистрации недвижимости или направляет пенсионера к несуществующему в его стране документу, последствия оказываются вполне существенными. А когда подобные сбои непропорционально часто затрагивают носителей государственного языка, —

возникают неудобные вопросы: интересы каких граждан изначально закладывались в основу цифровой инфраструктуры государства?

01 Что и как проверялось

Результаты, представленные в этой статье, основаны на серии тестовых запросов, направленных ИИ-ассистенту eGov Mobile в марте 2026 года. Одни и те же вопросы о типичных государственных услугах задавались сначала на русском, затем на казахском языке.

В марте 2026 года ИИ-ассистенту eGov Mobile было направлено семь парных запросов на русском и казахском языках, касающихся повседневных услуг: получения удостоверения личности, оформления пособия по беременности и родам, регистрации индивидуального предпринимательства, восстановления утраченного паспорта, приобретения недвижимости, решения пенсионного вопроса и расчёта штрафов на основе месячного расчётного показателя (АЕК на казахском, МРП на русском).

Каждая пара запросов тестировалась на одном устройстве, в рамках одной сессии, с очисткой чата между вопросами. Смена языка требовала изменения языковых настроек всего приложения. В рамках отдельного теста казахский запрос был направлен через русскоязычный интерфейс, а русский — через казахоязычный. В обоих случаях бот отвечал на языке запроса, а не интерфейса, что подтверждает: выявленные расхождения обусловлены не маршрутизацией по языковому признаку, а различиями в базовом контенте.

Казахские запросы составлялись как естественные, идиоматические тексты (а не дословные переводы с русского) и прошли проверку у носителя казахского языка. Там, где ответы бота содержали конкретные суммы, сроки или порядок действий, они сверялись с официальными источниками, в том числе с порталом egov.kz и законом о бюджете на 2026 год. При повторном тестировании одного из запросов на следующий день казахский бот дал иной ответ на тот же вопрос. Это означает, что система способна не только давать неверные ответы, но и давать разные неверные ответы в разные дни.

02 Что показало тестирование

Выявленные проблемы можно поделить на три: фактические ошибки, информационные пробелы и качество языка.

Неверные цифры, не та страна

Наиболее показательной оказалась ошибка, связанная с месячным расчётным показателем (АЕК/МРП) — базовым бюджетным ориентиром, определяющим размер штрафов, сборов, налоговых порогов и социальных выплат по всему Казахстану. Согласно закону о бюджете на 2026–2028 годы (Закон № 239-VIII), актуальное значение составляет 4 325 тенге (около 9 долларов США).

Русскоязычный бот назвал сумму в 3 800 тенге — значение, которое никогда не являлось официальным АЕК/МРП ни в одном году. Казахоязычный бот указал около 73 000 тенге, при этом неверно обозначил сам показатель и фактически перепутал его с другим экономическим понятием. Гражданин, воспользовавшийся данными русскоязычного бота, занизил бы размер своих штрафов примерно на 12%. Гражданин, доверившийся казахоязычному боту, завысил бы их примерно в 17 раз.

Запрос по пенсионной теме обнаружил иного рода ошибку. На расплывчатый вопрос о пенсионной проблеме русскоязычный бот рекомендовал предоставить российский страховой номер индивидуального лицевого счёта — документ, применяемый исключительно в Российской Федерации. Кроме того, он направил пользователя на то, что выглядело как российский портал государственных услуг, а не на казахстанскую систему eGov. Казахоязычный бот, напротив, корректно назвал казахстанский пенсионный фонд и сослался на надлежащие местные учреждения. В данном случае именно русскоязычный бот воспроизвёл рекомендации, относящиеся к другой стране.

На вопрос о восстановлении утраченного паспорта казахоязычный бот указал сроки оформления от одного до трёх рабочих дней. Официальный регламент, проверенный на egov.kz, предусматривает сроки, варьирующиеся в зависимости от места обращения и категории срочности: от одного дня в крупных городах до семи дней в отдалённых районах. казахоязычный бот умолчал о более длительных сроках. Житель сельского района, следуя такому совету, ожидал бы паспорт через три дня, тогда как в реальности это могло растянуться до недели. В том же казахоязычном ответе среди необходимых документов фигурировал «паспорт СССР 1974 года, при наличии» — требование, отсутствующее в русскоязычном ответе. Независимо от того, имеет ли оно процедурное обоснование (например, применительно к пожилым гражданам или лицам без гражданства, у которых могут сохраниться только советские документы), само по себе присутствие этого пункта в одной языковой версии и его отсутствие в другой

подтверждает вывод о том, что оба бота опираются на разные исходные материалы.

Системные информационные пробелы

По всем семи запросам прослеживалась устойчивая закономерность: если русскоязычный бот приводил конкретную сумму или срок, казахоязычный либо вовсе опускал эти сведения, либо заменял их размытыми формулировками. Так, ответ на русском языке о регистрации недвижимости содержал точный размер государственной пошлины — 2 404,9 тенге плюс банковская комиссия в 100 тенге, уплачиваемые онлайн через Visa или MasterCard, со сроком обработки два рабочих дня. Приведённая ссылка вела на соответствующую страницу услуги на egov.kz, где подтверждалась каждая из этих деталей.

Казахоязычный ответ на тот же вопрос сообщал лишь, что размер сборов «рассчитывается согласно ставкам, установленным налоговым кодексом», называл срок в три рабочих дня и содержал ссылку, ведущую не на регистрацию недвижимости, а на постороннюю услугу по выдаче справок, подтверждающих право собственности несовершеннолетних на имущество. Указанный трёхдневный срок относился именно к этой посторонней услуге. Бот не просто давал расплывчатый ответ: он перепутал две разные государственные услуги и сформировал ответ на основе неверной.

С запросом о пособии по беременности и родам произошла похожая история: русскоязычный бот указал, что услуга бесплатна и оформляется онлайн за четыре рабочих дня; казахоязычный не упомянул ни того, ни другого. По вопросу выдачи удостоверения личности русскоязычный бот перечислил три ценовых тарифа и два варианта срока оформления; казахоязычный опустил один тариф и все сроки.

Это не случайные упущения. Они свидетельствуют о структурном неравенстве в глубине и полноте казахоязычной базы знаний, лежащей в основе бота. В дополнительном тесте обе версии получили вопрос о документах, необходимых для подтверждения зарегистрированного места жительства. Русскоязычный бот корректно сообщил, что адресные справки были упразднены в 2019 году, и объяснил, как граждане могут теперь подтвердить регистрацию. Бот на казахском языке ответил на совершенно другой вопрос, предоставив инструкцию по услуге уточнения адреса недвижимости.

Две языковые версии, судя по всему, не являются переводами друг друга. Они опираются на разные исходные материалы с различным набором включённых и исключённых сведений, что указывает либо на формирование казахоязычного

раздела на основе иных источников, либо на менее строгий контроль качества при его создании.

Язык, выдающий машинное происхождение

Все полученные казахоязычные ответы были понятны, но звучали неестественно. Предложения оказывались длиннее, чем это принято в идиоматическом казахском, связующие слова и переходы ощущались механическими, а стиль не соответствовал тому, как писал бы казахоязычный государственный служащий. В одном из ответов бот использовал русскоязычную аббревиатуру ИИН (индивидуальный идентификационный номер) вместо казахского эквивалента ЖСН (жеке сәйкестендіру нөмірі) — деталь незначительная, но красноречивая: она обнажает, что текст был производным от русскоязычных источников, а не создавался изначально на казахском языке. Русскоязычные ответы, напротив, читались естественно и в соответствующем формальном регистре.

Этот разрыв в языковом качестве не касается понимания — речь идёт о том, воспринимает ли государство собственный язык всерьёз в цифровом пространстве. В одном случае казахоязычный ответ содержал ссылку на документ с инструкцией, который оказался написан целиком на русском языке. Казахский уровень, иными словами, порой представляет собой лишь внешнюю оболочку над русскоязычной инфраструктурой. В других случаях казахоязычный бот обращается к принципиально иным и нередко менее полным или ошибочно соотнесённым исходным материалам.

Система даёт сбой не в одном стиле. Она ошибается сразу несколькими способами и именно это делает её уязвимости непредсказуемыми для граждан и трудноустраняемыми для государства.

03 Что это означает

Легко можно объяснить эти выводы с естественным этапом роста молодого продукта, который улучшится с обратной связью. Это обещает и собственный дисклеймер бота. Но такая трактовка слишком снисходительна. Часть этих запросов впервые была протестирована за три недели до нашей структурированной оценки. Ответы оказались идентичными. Какой бы механизм обратной связи не существовал, он не даёт видимых улучшений в сроки, которые имеют значение для граждан.

Правительство Казахстана приняло решение внедрить этого ИИ-ассистента в действующий государственный сервис, которым пользуются миллионы людей.

Оно предложило его на казахском и русском языках. И тем самым дало косвенное обещание: граждане будут получать равноценный сервис вне зависимости от того, на каком языке они обращаются. Факты свидетельствуют о том, что это обещание не выполняется. Проблема не в том, что ИИ несовершенен.

Проблема в том, что несовершенство распределено неравномерно. Государственная платформа систематически предоставляет казахоязычным гражданам менее полную, менее точную и менее практичную информацию с той же государственной платформы. В стране, где казахский язык конституционно закреплён как государственный, а языковая политика неразрывно связана с суверенитетом, национальной идентичностью и легитимностью государства после получения независимости, это влечёт последствия, выходящие далеко за рамки технического недостатка.

Это также показывает провал в сфере государственных закупок. Задokumentированные различия говорят о том, что языковое равенство не рассматривалось как требование при проектировании, заказе или приёмке этой системы. Если никто не проверял, выдаёт ли казахский бот ту же сумму сбора, что и русский, за одну и ту же услугу, — значит, никто и не был обязан это проверять.

Сказанное актуально не только для Казахстана. По всей Центральной Азии правительства стремительно интегрируют ИИ в государственные услуги. Вопрос уже не в том, будут ли эти инструменты внедряться (они уже внедряются), а в том, потребуют ли стандарты государственных закупок, чтобы они работали одинаково для всех граждан на всех языках, которые то или иное государство декларирует как языки обслуживания. Без такого требования ИИ рискует превратиться в механизм, который позиционирует себя как модернизацию, а на практике воспроизводит и углубляет существующие языковые иерархии.

Рекомендации

1. Правительству следует установить минимальные стандарты соответствия языков для любых систем искусственного интеллекта, внедряемых в сфере государственных услуг. Эти стандарты должны требовать, чтобы все языковые версии опирались на одну и ту же проверенную базу знаний, возвращали равнозначную информацию на равнозначные запросы и проходили структурированное тестирование носителями языка перед запуском. В Уэльсе, составной нации Соединённого Королевства с собственным языком, валлийским, системы искусственного

интеллекта в государственном секторе обязаны относиться к валлийскому не хуже английского. Казахстан должен сделать то же самое в отношении казахского языка.

2. Процедуры закупок систем государственных услуг с использованием ИИ должны включать языковой паритет в качестве явного критерия оценки. В контрактах следует прописывать протоколы двуязычного тестирования, требовать документально подтверждённых свидетельств паритета и предусматривать возможность независимого аудита после внедрения. В-третьих, действующий ИИ-ассистент портала eGov должен быть подвергнут систематическому двуязычному аудиту, а его результаты — преданы огласке. Бот уже предлагает пользователям оставлять обратную связь. Прозрачная оценка качества продемонстрирует, что правительство воспринимает этот механизм всерьёз и готово предъявлять к собственным инструментам реальные требования.

Это минимальные стандарты ответственности для государства, взявшего курс на цифровое управление и декларирующего равное обслуживание граждан на обоих языках. Республиканский бюджет на 2026 год выделяет 177 миллиардов тенге (372 миллиона долларов) Министерству искусственного интеллекта и цифрового развития. Ничтожная доля этих инвестиций, направленная на обеспечение двуязычного контроля качества, позволила бы выявить каждую ошибку, задокументированную в настоящей статье, ещё до того, как с ней столкнулся бы хоть один гражданин. Инструменты существуют. Механизмы существуют. Не хватало понимания важности языкового паритета.

Примечание о методологии

Тестирование проводилось в марте 2026 года с использованием мобильного приложения eGov (iOS). Все ответы были задокументированы в виде скриншотов и сверены с официальными государственными источниками. Автор является носителем казахского языка. Исходные данные и полные журналы оценки предоставляются по запросу

CAPS Unlock

Мнения, изложенные в настоящей статье, отражают исключительно позицию автора и не представляют институциональную позицию CAPS Unlock или какой-либо организации, с которой аффилирован автор. CAPS Unlock публикует данную статью как вклад в информированную общественную дискуссию. Факт публикации не означает поддержки анализа или выводов, представленных в ней.

Мы приветствуем отклики, критику и альтернативные точки зрения — в особенности от исследователей и практиков, имеющих непосредственный опыт работы с рассматриваемыми вопросами. Если вы хотите вступить в диалог по материалам статьи, предложить ответный материал или направить свой вклад в раздел «Аргумент», свяжитесь с нами по почте editorial@capsunlock.org или через наш сайт capsunlock.org.