

Two Languages, Two Standards: AI and Linguistic Inequality in Kazakhstan's Public Services

Zhannat Bubekbayeva

Independent researcher based in Kazakhstan

Abstract

Kazakh-speaking users of e-government services are served poorly, and considerably worse so than users of Russian, despite the latter corresponding largely to a demographic minority in Kazakhstan. This fact stands in stark contrast to the country's otherwise strong performance in harnessing digital technology in the provision of government services to the public. Kazakhstan [ranks 24th globally](#) in the United Nations E-Government Development Index.

The current eGov Mobile app, which was [introduced in October 2019](#), has over 11.7 million registered users. Around 6.4 million people [use the app monthly](#). The platform deploys AI-powered assistants to help citizens navigate government services, from obtaining identity documents to registering businesses and applying for social benefits.

This paper tests whether that AI performs equally well in Kazakhstan's two most widely used languages. The results suggest clear disparities. A comparison of the eGov Mobile AI assistant shows that the Kazakh-language bot often omits practical information and produces text that reads as machine-translated rather than natively written. In one instance identified during this research, it returned a figure that exceeded the correct value by roughly 17 times.

The Russian-language bot performs better on average, but it is not reliable. It also produces serious errors, including references to documents and institutions from the Russian Federation that have no relevance to Kazakhstan.

These are not technical inconveniences. When a government AI assistant gives a citizen incorrect information about a fine, omits the cost of registering a property, or directs a pensioner to a document that does not exist in their country, the consequences are practical and immediate. And when these failures fall disproportionately on speakers of the state language, which is Kazakh, they raise uncomfortable questions about whose voices the state's digital infrastructure was built to serve.

01 What Was Tested and How

The findings of this paper are based on a series of test queries submitted to the eGov Mobile AI assistant in March 2026. The same questions were posed in Russian and then in Kazakh, covering routine government services.

In March 2026, seven paired queries were submitted to the eGov Mobile AI assistant—the same question in Russian, then in Kazakh—covering everyday services: obtaining an identity card, applying for maternity benefits, registering a sole proprietorship, replacing a lost passport, purchasing property, resolving a pension issue, and calculating fines based on the monthly calculation index (AEK in Kazakh, MRP in Russian).

Each pair was tested on the same device, within the same session window, with the chat cleared between queries. Switching between languages required changing the app's entire language setting. In a separate test, a Kazakh query was submitted through the Russian-language interface and a Russian query through the Kazakh-language interface. The bot responded in both cases in the language of the query rather than the interface, confirming that the disparities documented below do not stem from language routing but from differences in the underlying content.

The Kazakh queries were composed as natural, idiomatic Kazakh (not word-for-word translations of the Russian) and were assessed by a native Kazakh speaker. Where bot responses cited specific fees, deadlines, or procedures, they were cross-checked against official sources, including egov.kz and the 2026 budget law. When one query was retested the following day, the Kazakh bot returned a different response to the same question, meaning the system is not only capable of giving wrong answers, but of giving different wrong answers on different days.

02 What the Testing Found

Three categories of problems emerged: factual errors, information gaps, and language quality.

Wrong numbers, wrong country

The most striking finding involved the monthly calculation index (AEK/MRP), a basic fiscal reference point that determines the size of fines, fees, tax thresholds, and social payments across Kazakhstan. According to [the 2026-28 budget law](#) (Law No. 239-VIII), the correct value is 4,325 tenge (around \$9). The Russian bot stated 3,800 tenge, a figure that has never been the official AEK/MRP in any year. The Kazakh bot stated approximately 73,000 tenge, incorrectly named the indicator and effectively confused it with an entirely different economic metric. A citizen relying on the Russian bot would

underestimate their fines by about 12%. A citizen relying on the Kazakh bot would overestimate them by a factor of roughly 17.

The pension query produced a different kind of error. Asked a vague question about a pension problem, the Russian bot recommended that the citizen provide a Russian pension insurance number, a document used exclusively in the Russian Federation. It also directed the user to what appeared to be the Russian state services portal rather than Kazakhstan's eGov system. The Kazakh bot, by contrast, correctly named the Kazakhstani pension fund and referred to appropriate local institutions. In this case, the Russian bot imported guidance from the wrong country entirely.

When asked about replacing a lost passport, the Kazakh bot cited processing timelines of one to three working days. The official schedule, [verified on egov.kz](#), specifies that timelines vary by location and urgency category, ranging from one day in major cities to seven days in more remote areas. The Kazakh bot omitted the longer processing times. A citizen in a rural district following this advice would expect their passport in three days when the actual wait could be a week. The same Kazakh response also listed among the required documents a "1974 USSR passport, if available," a requirement absent from the Russian-language response. Whether or not this has a procedural basis (for instance, in relation to older citizens or stateless persons who may retain only Soviet-era documents), its presence in one language version and absence from the other reinforces the finding that the two bots draw on different source material.

Systematic information gaps

Across all seven queries, a consistent pattern emerged. Whenever the Russian bot provided a specific cost figure or processing timeline, the Kazakh bot either omitted it or offered a vague substitute. The Russian response on property registration, for instance, stated the exact state fee: 2,404.9 tenge plus a 100-tenge bank commission, payable online via Visa or MasterCard, with processing completed within two working days. The link it provided led to the [correct egov.kz service page](#), which confirmed every detail.

The Kazakh response to the same question stated that fees are "calculated according to the rates established in the tax code," cited a processing time of three working days, and provided a link which led not to property registration but to an unrelated service for issuing certificates confirming property ownership by minors. The three-day timeline corresponded to that unrelated service. The bot was not merely vague. It had confused two different government services and assembled its answer from the wrong one.

The maternity benefits query followed the same pattern: the Russian bot noted that the service is free and takes four working days online; the Kazakh bot mentioned neither.

On identity card issuance, the Russian bot listed three fee tiers and two processing timelines; the Kazakh bot omitted one fee tier and all timelines.

These are not random omissions. They point to a structural disparity in the depth and completeness of the Kazakh-language knowledge base underpinning the bot. In one additional test, both versions were asked about the documents needed to confirm a citizen's registered place of residence. The Russian bot correctly noted that [address certificates were abolished](#) in 2019 and explained how citizens can now verify their registration. The Kazakh bot answered a different question entirely, providing guidance on a service used to clarify or correct address details in property records.

The two language versions do not appear to be translations of each other. They draw on different underlying content, with different details included and excluded, suggesting that the Kazakh track was either built from different source material or subjected to less rigorous quality control.

Language that reads like a machine wrote it

Every Kazakh response received was comprehensible but unnatural. Sentences ran longer than idiomatic Kazakh would warrant, linking words and transitions felt mechanical, and the overall register did not match how a Kazakh-speaking government official would write. In one response, the bot used the Russian abbreviation IIN (индивидуальный идентификационный номер, the individual identification number used for official records) instead of the Kazakh equivalent ZhSN (жеке сәйкестендіру нөмірі), a small detail, but one that reveals the text was derived from Russian source material rather than authored in Kazakh. The Russian responses, by contrast, read as natural and appropriately formal.

This gap in language quality is not about comprehension; it is about whether the state takes its own language seriously in digital spaces. In one case, a Kazakh-language response included a link to a guideline document that turned out to be written entirely in Russian. The Kazakh layer, in other words, is sometimes little more than a surface over Russian-language infrastructure. In other cases, the Kazakh bot draws on entirely different (and often less complete or incorrectly mapped) source material.

The system does not fail in one consistent way. It fails in several ways at once, which makes it harder for citizens to anticipate where the gaps are and harder for the government to fix them.

03 What This Means

It would be tempting to treat these findings as growing pains: an early product that will improve with feedback, as the bot's own disclaimer promises. That framing is too generous. Several of these queries were first tested three weeks before the structured evaluation described here. The responses were identical. Whatever feedback loop exists, it is not producing visible improvements on a timeframe that matters to citizens.

Kazakhstan's government chose to deploy this AI assistant into a live, citizen-facing service used by millions. It chose to offer it in both Kazakh and Russian. And in doing so, it made an implicit promise: that citizens would receive equal service regardless of which language they use. The evidence shows that promise is not being kept.

The problem is not that AI is imperfect. The problem is that imperfection is unevenly distributed. Kazakh-speaking citizens systematically receive less complete, less accurate, and less actionable information from the same government platform. In a country where the Kazakh language is constitutionally designated as the state language, and where language policy is intertwined with sovereignty, national identity, and the state's post-independence legitimacy, this carries implications that extend beyond a technical shortcoming.

It is also a procurement failure. The disparities documented here suggest that language equality was not treated as a requirement in the design, commissioning, or acceptance of this system. If no one checked whether the Kazakh bot gives the same fee as the Russian bot for the same service, then no one was asked to check.

This matters beyond Kazakhstan. Across Central Asia, governments are integrating AI into public services at pace. The question is not whether these tools will be deployed (they already are) but whether procurement standards will demand that they work equally for all citizens, in all the languages a given state claims to be serving. Without that demand, AI risks becoming a mechanism that presents itself as modernization while in practice reproducing and deepening existing linguistic hierarchies.

The Signpost

First, the government should establish minimum language parity standards for any AI system deployed in public services. These standards should require that all language versions draw on the same verified knowledge base, return equivalent information for equivalent queries, and undergo structured testing by native speakers before

launch. In Wales, a constituent nation of the United Kingdom with its own distinct language, Welsh, public-sector AI systems are required to treat Welsh no less favorably than English. Kazakhstan should do the same for Kazakh.

Second, procurement processes for AI-enabled government services should include language parity as an explicit evaluation criterion. Contracts should specify bilingual testing protocols, require documented evidence of parity, and include provisions for independent audit after deployment.

Third, the current eGov AI assistant should be subjected to a systematic bilingual audit and the results made public. The bot already invites user feedback. A transparent quality assessment would demonstrate that the government takes that feedback loop seriously and is willing to hold its own tools to account.

These are minimum standards of care for a state that has committed to digital governance and claims to serve its citizens equally in both languages. The 2026 republican budget [allocates](#) 177 billion tenge (\$372 million) to the Artificial Intelligence and Digital Development Ministry. A fraction of that investment in bilingual quality assurance would have caught every error documented in this paper before a single citizen encountered it. The tools exist. The frameworks exist. What has been missing is the expectation that language parity is not optional.

Methodology Note

Testing was conducted in March 2026 using the eGov Mobile application (iOS). All responses were documented via screenshot and cross-referenced against official government sources. The author is a native Kazakh speaker. The raw data and full evaluation logs are available on request.

CAPS Unlock

The views expressed in this paper are those of the author alone and do not represent the institutional position of CAPS Unlock or any organization with which the author is affiliated. CAPS Unlock publishes this paper as a contribution to informed public debate. Publication does not imply endorsement of the analysis or conclusions.

We welcome responses, critiques, and alternative perspectives, particularly from researchers and practitioners with direct experience of the issues discussed here. If you would like to engage with this paper, propose a reply, or submit a contribution to the Insight Series, please contact us at editorial@capsunlock.org or visit capsunlock.org.