

**Language Parity as a
Public Service Standard:
Minimum Benchmarks
for State-Language AI
in Central Asian Public
Procurement**

by Elliott Ahn, PhD

SEPTEMBER
2026

© 2026 OSCE Academy in Bishkek. All rights reserved.

The views expressed and the conclusions reached in this report are exclusively those of the author and are not necessarily shared or endorsed by the OSCE Academy in Bishkek and the Norwegian Institute of International Affairs.

Extracts of this report may be quoted or reprinted without special permission for academic purposes provided that a standard source credit line is included. Academy Publication Guidelines are available on-line at https://www.osce-academy.net/upload/file/OSCE_AiB_Publication_guidelines.pdf

This report became possible with financial support from the Norwegian Ministry of Foreign Affairs via the Academy's Cooperation with the Norwegian Institute of International Affairs.

POLICY BRIEF

117, SEPTEMBER 2026

Language Parity as a Public Service Standard: Minimum Benchmarks for State-Language AI in Central Asian Public Procurement

by Elliott Ahn, PhD

Executive summary

Globalization concentrates digital life in a handful of big languages even as national languages hold formal primacy at home; every Central Asian state now manages that tension inside its public services. Central Asian public bodies are expanding digital services and beginning to use artificial intelligence (AI) in chatbots, search engines, document processing, and other citizen-facing functions. Yet language resources and evaluation infrastructure for Kazakh, Kyrgyz, Turkmen, and Uzbek remain much thinner than for English or Russian.¹ On a parallel reading benchmark, two models evaluated via the 2023 Belebele results scored 32–35 accuracy points in Kazakh and 79–88 in English.² By 2025, frontier systems had narrowed the Kazakh–Russian gap on matched school subjects to under five points, while leading open-weight models still averaged some 20 points lower — evidence that the gap closes only where buyers procure frontier capability.³ When an automated public service performs materially worse in a language through which citizens are expected to reach the state, the result is an (un)equal-access and public-trust risk. The response should not force a choice between state languages and languages of wider communication. This brief proposes

¹ Yana Veitsman and Mareike Hartmann, “Recent Advancements and Challenges of Turkic Central Asian Language Processing,” Proceedings of the First Workshop on Language Models for Low-Resource Languages (LoResLM 2025), pp. 309–324, <https://aclanthology.org/2025.loreslm-1.25/>

² Lucas Bandarkar, et al., “The Belebele Benchmark,” arXiv:2308.16884 (2023), reported in Akylbek Maxutov, Ayan Myrzakhmet and Pavel Braslavski, “Do LLMs Speak Kazakh? A Pilot Evaluation of Seven Models,” Proceedings of the First Workshop on Natural Language Processing for Turkic Languages (SIGTURK 2024), pp. 81–91, <https://aclanthology.org/2024.sigturk-1.8/>

³ Mukhammed Togmanov, et al., “KazMMLU: Evaluating Language Models on Kazakh, Russian, and Regional Knowledge of Kazakhstan,” Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics, Vol. 1 (2025), pp. 14403–14416, <https://aclanthology.org/2025.acl-long.701/>

a language-portfolio clause for citizen-facing AI contracts. For every covered language, the buyer would define task-specific metrics, an absolute service floor, and a maximum gap from the best-performing service language; test these before acceptance; withhold final payment if they are missed; and retest after material model changes. As an initial default, the permitted gap would be no more than ten percentage points, tightening to five within 24 months, while safety-critical scenarios would have to pass in every language. These figures are proposed starting points for market testing, not universal constants. The European Charter for Regional or Minority Languages (ECRML) offers a non-binding comparative methodology: graduated commitments, population-sensitive coverage, and monitoring. The clause can be piloted through tender specifications, subject to national procurement and equality-law review.

Elliott Ahn, PhD, is a Distinguished Professor at the Graduate School of AI Policy and Strategy, Gwangju Institute of Science and Technology (GIST). He serves in an unpaid advisory capacity on the Expert Board of Azerbaijan's ASAN AI Hub and is Executive Director of Antler IBEX, where he heads Korea and Japan deep tech. Neither organization funded, commissioned, or reviewed this article. He holds an LLM from Georgetown University Law Center and advises government innovation programmes in Korea and Japan.

Email: seokhyun.ahn@gmail.com

Introduction

Globalization has an ambivalent relationship with language. Shared infrastructure lowers the cost of communication across borders, but it also rewards scale: A few widely used languages dominate science, commerce, entertainment, and the data used to train artificial intelligence. At the same time, national languages hold a privileged legal and symbolic position within their states as languages of administration and education and as markers of sovereignty and identity. Formal primacy, however, does not guarantee practical reach. A language can be constitutionally privileged at home yet be eclipsed online where thinner datasets and weaker tools make its speakers dependent on a more widely used language.

International frameworks protect linguistic identity and access in different, overlapping ways. The Council of Europe's European Charter for Regional or Minority Languages (ECRML) creates treaty commitments for ratifying states; Article 22 of the EU Charter of Fundamental Rights requires the Union to respect linguistic diversity; the OSCE High Commissioner on National Minorities links language access to effective participation in public life; and UNESCO promotes multilingualism in cyberspace. These instruments do not create a single universal entitlement to receive every service in every mother tongue. Together, however, they establish that linguistic diversity, non-discrimination, and meaningful access are legitimate public interests.⁴

Global homogenization therefore brings both gains and risks. Widely used languages lower transaction costs, widen access to research and markets, and make systems easier to connect with. Less widely spoken languages carry identity, local knowledge, democratic participation, and trust; but protection can become costly or inward-looking if it is pursued by exclusion. The policy problem is not to choose one process over the other, it is to keep global languages as bridges while ensuring that state and minority languages remain usable channels for public institutions. How can states achieve this without either obstructing globalization or

⁴ Council of Europe, European Charter for Regional or Minority Languages, European Treaty Series No. 148, (Strasbourg: Council of Europe, 1992); Charter of Fundamental Rights of the European Union, Official Journal of the European Union, C 326 (2012), Article 22; OSCE High Commissioner on National Minorities, Recommendations on the Effective Participation of National Minorities in Social and Economic Life and Explanatory Note, (The Hague: OSCE HCNM, 2023), <https://hcnm.osce.org/sites/default/files/f/documents/2/3/553783.pdf>; UNESCO, Recommendation concerning the Promotion and Use of Multilingualism and Universal Access to Cyberspace, (Paris: UNESCO, 2003).

pursuing it at the expense of its own citizens, language and culture? This brief argues for an institutional answer: additive multilingualism, translated into concrete and testable requirements for the systems governments actually operate.

The Turkic language family brings the dilemma into sharp focus. That its speakers number well over 150 million is often cited, but the figure matters less than what it conceals.⁵ The languages are related, not interchangeable: shared roots and varying degrees of affinity across borders do not mean that a Kyrgyz speaker simply reads Uzbek, or that a service built for one language serves another. The region is also divided by script. Kyrgyzstan retains the Cyrillic alphabet; Kazakhstan has been preparing a move from Cyrillic to Latin for years without a final alphabet yet in official use; Uzbekistan adopted a Latin alphabet after independence while Cyrillic remains in wide everyday use; Turkmenistan, Azerbaijan, and Türkiye write in Latin. For a well-educated professional these differences are an inconvenience. For a less qualified citizen dealing with an administrative platform, they can be an insurmountable barrier: a system trained on one script may fail on the other, and a person who writes in the script learned at school may receive no usable answer at all.⁶ Central Asian Turkic speakers are therefore connected to a large transnational family yet semi-isolated in practical terms: no single member language belongs to the small group that dominates global digital infrastructure, resources remain thin relative to the family's demographic weight, and the family's internal divisions of script and grammar mean that no one language can stand in for the others. The choice between the mother tongue and the wider world is nonetheless a false one. Public institutions can keep both open, and the most practical place to begin is the point at which governments buy the technology through which they speak to citizens.

⁵ Rustem Yeshpanov, Saida Mussakhoyeva and Yerbolat Khassanov, "Multilingual Text-to-Speech Synthesis for Turkic Languages Using Transliteration," *Proceedings of Interspeech 2023* (2023), pp. 5521–5525, <https://doi.org/10.21437/Interspeech.2023-249>. On the scale and reach of the family as a whole, see Lars Johanson, *Turkic*, (Cambridge: Cambridge University Press, 2021).

⁶ On the parallel functioning of Cyrillic and Latin in Uzbekistan and the generational gap identified as an obstacle to script transition in Kyrgyzstan, see M. Dzhusupov, "The Transition from Cyrillic into Latin Alphabet and Linguographic Interference in the Russian Speech of the Turkophone," *Russian Language Studies*, Vol. 20, No. 3 (2022), pp. 312–329; G. Toktogulova, et al., "Prospects of Kyrgyz Language Transition to the Latin Script Based on Experience of Kazakhstan and Uzbekistan," *Language Problems and Language Planning* (2025). On script variation and its consequences for language technology, see Veitsman and Hartmann, "Recent Advancements and Challenges of Turkic Central Asian Language Processing," pp. 309–324.

The Policy Challenge

Evidence of an Uneven Language Baseline

Research on Central Asian Turkic language processing consistently identifies scarcity of corpora, labelled data, tools, and evaluation sets.⁷ Direct model evaluations make the operational issue clearer: in Kazakh, large language models perform below their English results on parallel tasks and struggle most with generative work.⁸ Newer benchmarks — including Kardeş-NLU, TUMLU, and KazMMLU — cover different subsets of the region and show marked variation by language, task, and model; coverage nevertheless remains uneven.⁹ The most recent bilingual evidence confirms the pattern with current systems: on KazMMLU, built from Kazakhstan’s own school curriculum, models generally score better in Russian than in Kazakh on matched subjects, the best Kazakh result (76.9 percent, GPT-4o) trails the best Russian result (81.4 percent, DeepSeek V3), and the strongest open-weight models average only 56–57 percent overall.¹⁰ That a gap of under five points is already attainable at the frontier matters for this brief: it suggests the phased cap proposed below is ambitious but procurable. The narrowing at the frontier is not self-executing: without a contractual requirement and acceptance test, buyers cannot assume that an ordinary tender will deliver frontier-level performance. The clause converts a best-case observation into a contractual floor. Together, these studies support measuring performance separately by language and task. They do not prove that any particular deployed government system fails, because few public bodies publish matched service-level tests. That evidence gap is itself the procurement problem: without an acceptance test, neither the buyer nor the public can know whether a multilingual interface provides an equivalent service.

⁷ Veitsman and Hartmann, “Recent Advancements and Challenges of Turkic Central Asian Language Processing”, pp. 309–324.

⁸ Maxutov, Myrzakhmet and Braslavski, “Do LLMs Speak Kazakh?”, pp. 84–88.

⁹ Lütfi Kerem Şenel, et al., “Kardeş-NLU: Transfer to Low-Resource Languages with the Help of a High-Resource Cousin – A Benchmark and Evaluation for Turkic Languages,” Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics, Vol. 1 (2024), pp. 1672–1688, <https://aclanthology.org/2024.eacl-long.100/>; Jafar Isbarov, et al., “TUMLU: A Unified and Native Language Understanding Benchmark for Turkic Languages,” Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics, Vol. 1 (2025), pp. 22816–22838, <https://aclanthology.org/2025.acl-long.1112/>; Togmanov, et al., “KazMMLU”, pp. 14403–14416.

¹⁰ Togmanov, et al., “KazMMLU”, Table 4 and Figure 6.

Table 1: Reported Accuracy Gaps between Kazakh and Higher-Resource Languages

<i>Benchmark (year)</i>	<i>Model</i>	<i>Kazakh</i>	<i>Comparator</i>	<i>Gap (points)</i>
Belebele parallel reading (2023)	GPT-3.5-turbo	35.0	87.7 (English)	52.7
Belebele parallel reading (2023)	LLaMA 2-Chat 70B	32.4	78.8 (English)	46.4
KazMMLU school subjects (2025)	Best model per language	76.9 (GPT-4o)	81.4 (Russian, DeepSeek V3)	4.5

Notes or Source: Rows one and two: Bandarkar, et al. (2023), as reported in Maxutov, et al. (2024); random guessing yields 25. Row three: Togmanov, et al. (2025), high-school subjects only, best-performing model shown per language.

A Public-Service and Security Risk

The OSCE High Commissioner on National Minorities states that equal access to public goods and services is a precondition for effective participation, and that people should be able to access them in a language they understand and, preferably, their own.¹¹ Automated systems do not sit outside that standard. A chatbot that misroutes a benefits application, invents a deadline, or fails to recognise an emergency request creates the same practical exclusion as a poorly staffed counter. Uneven performance may also weaken confidence in official channels, especially in health, disaster, and civil-protection communication. This link should be treated as a plausible policy risk rather than a quantified causal effect until Central Asian service data are available.

The Other Side of the Coin

Global digital markets can create homogenizing pressure, but outcomes are not deterministic. UNESCO documents both the exclusion of most languages and technology's potential to strengthen their digital

¹¹ OSCE High Commissioner on National Minorities, Recommendations on the Effective Participation of National Minorities in Social and Economic Life, Recommendations 15 and 17.

presence.¹² OSCE guidance similarly treats multilingual education, learning state or official languages, and maintaining one's own language as complementary. Central Asian policy should avoid a false choice: improve state-language capability while retaining Russian or other established services where required. Additional international languages can widen access to science, trade, and education, but which language matters is context-specific; English and Chinese are options, not universal substitutes for Russian. Kazakhstan's English-medium reforms show both competitiveness in aims and unequal implementation where capacity is concentrated.¹³ The Council of Europe's own implementation study reaches a parallel conclusion: machine translation and related language technologies can serve as practical bridges between smaller and more widely used languages where data and applications are deliberately developed.¹⁴ The objective is additive multilingualism.

Why Procurement Is the Near-Term Lever

AI suppliers optimize to the requirements on which they are evaluated and paid. Public-sector guidance already recommends defining acceptable model performance, testing under a range of conditions, and evaluating systems across the lifecycle; the European Union's model contractual AI clauses show that buyers can translate AI governance into reusable contract language.¹⁵ Neither supplies a Central Asian language-parity metric; the buyer must add one. A model clause can begin inside existing tender documents where national law already permits functional performance requirements. Each jurisdiction should nonetheless review procurement, equality, language, data-protection, and administrative-law constraints before scaling it.

¹² UNESCO, Recommendation concerning the Promotion and Use of Multilingualism and Universal Access to Cyberspace; UNESCO, Global Roadmap for Multilingualism in the Digital Era, (Paris: UNESCO, 2025).

¹³ Laura Karabassova, "English-Medium Education Reform in Kazakhstan," Current Issues in Language Planning, Vol. 22, No. 5 (2021), pp. 553–573; Bridget Goodman and Assel Kambatyrova, "Both Necessary and Irrelevant," International Journal of the Sociology of Language, No. 277 (2022), pp. 77–94.

¹⁴ Miriam Gerken, Facilitating the Implementation of the European Charter for Regional or Minority Languages through Artificial Intelligence, (Strasbourg: Council of Europe, 2022), pp. 13–16 and 22, <https://rm.coe.int/min-lang-2022-4-ai-and-ecrml-en/1680a657c5>; Committee of Experts of the ECRML, Statement on the Promotion of Regional or Minority Languages through Artificial Intelligence, (Strasbourg: Council of Europe, March 16, 2022).

¹⁵ United Kingdom Office for Artificial Intelligence, Guidelines for AI Procurement, (London: Office for Artificial Intelligence, 2020), <https://www.gov.uk/government/publications/guidelines-for-ai-procurement/guidelines-for-ai-procurement>; European Commission Public Buyers Community, Updated EU AI Model Contractual Clauses, (Brussels: European Commission, 2025).

Policy Change: A Language-Portfolio Clause

1. Define Coverage without Forcing Substitution

The clause should apply to each state language offered by the system, every language that law or established practice requires, and significant regional or minority languages where demand and speaker concentration justify coverage. It must not permit compliance by degrading Russian, English, or another well-performing channel. Any global-language bridge should be optional and user controlled, and should use the same verified service content rather than directing citizens to an unaccountable external model.

2. Measure the Service, Not the Model Brand

For each material function, the tender should name a service outcome and a metric: intent-recognition F1, transaction completion, retrieval accuracy against the agency's authorised knowledge base, factual-error rate, summarization fidelity, character error rate, or word error rate. Generic research benchmarks can screen market readiness, but they cannot replace a test built from the forms, terminology, scripts, dialects, and user journeys of the actual service.

3. Use Matched, Independent Acceptance Testing

The authority should commission semantically matched test cases in every covered language, validated by native-language specialists and representative users. The suite should include both Cyrillic and Latin scripts where used, common code-switching, rural and older users' phrasing, accessibility cases, and pre-designated high-consequence scenarios. A public sample lets vendors prepare; a held-back set prevents benchmark gaming. An evaluator independent of the supplier should run the frozen system, report each language and task separately, and publish the method, combining human review with automated scoring against known ground truth.¹⁶

¹⁶ Elham Tabassi, Artificial Intelligence Risk Management Framework (AI RMF 1.0), NIST AI 100-1, (Gaithersburg: National Institute of Standards and Technology, 2023).

4. Apply a Floor, a Gap Cap, and a Confidence Rule

For a covered language, define the parity gap as the best-performing service-language score minus that language's score on the same test. Acceptance requires both the absolute minimum set for the service and a gap no larger than the contractual cap — preventing a vendor from achieving parity by making every language equally poor. The pilot default should cap the gap at ten percentage points for core transactions, tightening to five after 24 months or two procurement cycles. The test plan should pre-register sample sizes and require the upper bound of a 95 percent confidence interval to remain within the cap. High-consequence scenarios, such as emergency instructions or eligibility deadlines, should use mandatory pass criteria in every language rather than an averaged score.

5. Put the Obligation into Operative Contract Language

A tender schedule can state: 'For each Covered Language and Material Function, the Contractor shall achieve the Absolute Service Floor and a Parity Gap no greater than the applicable cap in Schedule X. Compliance shall be determined at acceptance by the Authority or its independent evaluator using the approved held-back test set. Failure constitutes non-acceptance of the affected function and triggers the remedies in Schedule Y.' The schedules then identify languages, functions, metrics, floors, gap cap, evidence, test owner, cure period, and retest events. This lets an ordinary procurement team administer the obligation without judging model architecture.

6. Link the Result to Payment and Lifecycle Control

The contract should make acceptance a condition of final payment, identify the tested model and knowledge-base versions, require notice of material changes, and trigger re-testing after updates or annually. Persistent failure should activate a cure period, service restriction, payment holdback, or an exit right proportionate to risk. Results should be published by system, language, task, version, and date, with security-sensitive details withheld only where necessary.

ECRML

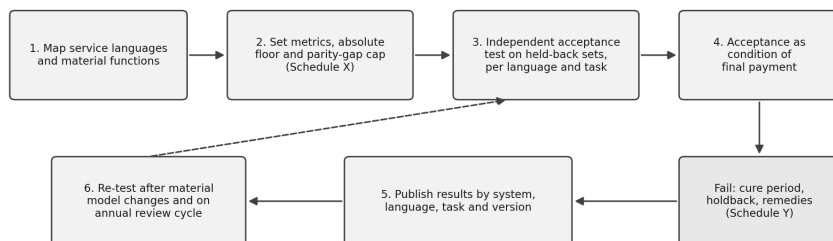


Figure 1: The language-portfolio clause in the procurement lifecycle.

7. Use the ECRML as a Comparative Reference

No Central Asian state is among the 25 parties to the ECRML.¹⁷ Article 1 ordinarily defines regional or minority languages as different from a state's official languages, while Article 3 allows a Party to designate a less widely used official language for selected Part III undertakings. The Charter is therefore a non-binding comparison, not transplanted law. Its preamble rejects zero-sum framing: protection should not detract from official languages or the need to learn them. Useful features include graduated undertakings, sensitivity to speaker numbers and location, consultation, and monitoring. A Council of Europe Secretariat study identifies chatbots, smart search, speech tools, and machine translation as ways to support language use in administration and public services. In Central Asia these ideas can inform phased coverage, while OSCE recommendations provide the closer normative anchor.

¹⁷ Council of Europe, European Charter for Regional or Minority Languages, Preamble and Articles 1, 3 and 10; chart of signatures and ratifications.

Recommendations

1. Issue a Model Clause. National procurement authorities and digital-development ministries should jointly publish, within six months, a customizable language-portfolio schedule for citizen-facing AI, defining covered systems and languages, requiring an absolute floor plus a parity cap, specifying evidence and remedies, and including a legal-review checklist.

2. Build National Test Suites. Digital ministries should commission universities and native-language research communities to develop versioned, task-specific datasets for common public-service journeys, reserve held-back items, require coverage of relevant scripts and user groups, and assign acceptance testing to an independent evaluator.

3. Pilot before High-Consequence Deployment. Buying agencies should begin with two high-volume, lower-risk services and use market sounding to test whether the ten-point cap is ambitious but procurable, then publish disaggregated results before extending the method to benefits, health, or emergency services. Where no compliant bid emerges for a covered language, the authority should contract with a published remediation schedule and interim thresholds rather than withhold the service altogether.

4. Protect the Full Language Portfolio. Service-owning ministries should map statutory, established, and high-demand service languages before drafting the tender, add state-language parity without weakening an existing Russian or minority-language channel, and provide a user-controlled bridge to an international language where national policy calls for it.

5. Coordinate Regionally. OSCE field operations and Academy-linked networks should convene procurement bodies, language communities, and researchers around a voluntary regional testing forum sharing metric definitions, retired test items, and evaluator training, while preserving national ownership.

Conclusion

The problem described in this brief is created by technology, but it will not be solved by technology alone. Left to the market, language models improve first where data are concentrated, and the languages of Central Asia are neither. Progress at the frontier is real, yet it reaches a citizen only if the system a government actually buys performs in that citizen's language and script. Nothing in the technology guarantees that outcome; only the buyer can.

That is why the matter requires the deliberate attention of governments rather than patience with the market. Access to information and to administrative platforms in one's own language is not a convenience to be traded off against efficiency. Where the state moves its services onto systems that work in some languages and fail in others, it narrows in practice a right that remains intact on paper, and it does so first for the citizens least able to route around the failure: the less qualified, the rural, the older, and those who write in the script they were taught.¹⁸ The instrument proposed here is modest by design. It creates no new institution and can begin within existing procurement rules wherever those rules already permit functional performance requirements; it asks only that language performance be specified, tested, and paid for like any other requirement in a public contract, with a floor for the state language, continuity for established service languages, and a bridge to the wider world for those who want one. Coupled with public reporting and regional cooperation on test suites, that is enough to turn a stated commitment to linguistic access into a measurable one. The technology will keep improving. The task of governments is to make sure that the improvement arrives in every language through which their citizens are entitled to reach the state.

¹⁸ Ulla Buchert, et al., "Is Digitalisation of Public Health and Social Welfare Services Reinforcing Social Exclusion? The Case of Russian-Speaking Older Migrants in Finland," *Critical Social Policy* (2022); Vikram Singh, et al., "Digital Divide: Barriers to Accessing Online Government Services in Canada," *Administrative Sciences*, Vol. 12, No. 3 (2022).

References

Bandarkar, Lucas, et al. “The Belebele Benchmark: A Parallel Reading Comprehension Dataset in 122 Language Variants.” arXiv preprint arXiv:2308.16884, 2023. <https://arxiv.org/abs/2308.16884>

Buchert, Ulla, et al. “Is Digitalisation of Public Health and Social Welfare Services Reinforcing Social Exclusion? The Case of Russian-Speaking Older Migrants in Finland.” Critical Social Policy (2022). <https://doi.org/10.1177/02610183221105035>

Committee of Experts of the European Charter for Regional or Minority Languages. Statement on the Promotion of Regional or Minority Languages through Artificial Intelligence. Strasbourg: Council of Europe, March 16, 2022.

Council of Europe. European Charter for Regional or Minority Languages. European Treaty Series No. 148. Strasbourg: Council of Europe, 1992.

Dzhusupov, M. “The Transition from Cyrillic into Latin Alphabet and Linguographic Interference in the Russian Speech of the Turkophone.” Russian Language Studies 20, no. 3 (2022): 312–329. <https://doi.org/10.22363/2618-8163-2022-20-3-312-329>

European Commission Public Buyers Community. Updated EU AI Model Contractual Clauses. Brussels: European Commission, March 5, 2025. <https://public-buyers-community.ec.europa.eu/communities/procurement-ai/resources/updated-eu-ai-model-contractual-clauses>

European Union. Charter of Fundamental Rights of the European Union. Official Journal of the European Union C 326 (2012), Article 22.

Gerken, Miriam. Facilitating the Implementation of the European Charter for Regional or Minority Languages through Artificial Intelligence. Strasbourg: Secretariat of the European Charter for Regional or Minority Languages, Council of Europe, 2022. <https://rm.coe.int/min-lang-2022-4-ai-and-eocrml-en/1680a657c5>

Goodman, Bridget, and Assel Kambatyrova. “Both Necessary and Irrelevant: Political Economy and Linguistic Injustice of English in Higher Education in Kazakhstan.” International Journal of the Sociology of Language, no. 277 (2022): 77–94.

Isbarov, Jafar, et al. "TUMLU: A Unified and Native Language Understanding Benchmark for Turkic Languages." In Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 22816–22838. Vienna: Association for Computational Linguistics, 2025. <https://aclanthology.org/2025.acl-long.1112/>

Johanson, Lars. Turkic. Cambridge: Cambridge University Press, 2021. <https://doi.org/10.1017/9781139016704>

Karabassova, Laura. "English-Medium Education Reform in Kazakhstan: Comparative Study of Educational Change across Two Contexts in One Country." Current Issues in Language Planning 22, no. 5 (2021): 553–573.

Maxutov, Akylbek, Ayan Myrzakhmet, and Pavel Braslavski. "Do LLMs Speak Kazakh? A Pilot Evaluation of Seven Models." In Proceedings of the First Workshop on Natural Language Processing for Turkic Languages (SIGTURK 2024), 81–91. Bangkok: Association for Computational Linguistics, 2024. <https://aclanthology.org/2024.sigturk-1.8/>

OSCE High Commissioner on National Minorities. Recommendations on the Effective Participation of National Minorities in Social and Economic Life and Explanatory Note. The Hague: OSCE HCNM, October 2023. <https://hcnm.osce.org/sites/default/files/f/documents/2/3/553783.pdf>

Şenel, Lütü Kerem, et al. "Kardeş-NLU: Transfer to Low-Resource Languages with the Help of a High-Resource Cousin – A Benchmark and Evaluation for Turkic Languages." In Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics (Volume 1: Long Papers), 1672–1688. St. Julian's: Association for Computational Linguistics, 2024. <https://aclanthology.org/2024.eacl-long.100/>

Singh, Vikram, et al. "Digital Divide: Barriers to Accessing Online Government Services in Canada." Administrative Sciences 12, no. 3 (2022): 112. <https://doi.org/10.3390/admsci12030112>

Tabassi, Elham. Artificial Intelligence Risk Management Framework (AI RMF 1.0). NIST AI 100-1. Gaithersburg, MD: National Institute of Standards and Technology, 2023.

Togmanov, Mukhammed, et al. “KazMLU: Evaluating Language Models on Kazakh, Russian, and Regional Knowledge of Kazakhstan.” In Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 14403–14416. Vienna: Association for Computational Linguistics, 2025. <https://aclanthology.org/2025.acl-long.701/>

Toktogulova, G., et al. “Prospects of Kyrgyz Language Transition to the Latin Script Based on Experience of Kazakhstan and Uzbekistan.” Language Problems and Language Planning (2025). <https://doi.org/10.1075/lplp.25003.tok>

UNESCO. Global Roadmap for Multilingualism in the Digital Era. Paris: UNESCO, 2025.

UNESCO. Recommendation concerning the Promotion and Use of Multilingualism and Universal Access to Cyberspace. Paris: UNESCO, 2003.

United Kingdom Office for Artificial Intelligence. Guidelines for AI Procurement. London: Office for Artificial Intelligence, June 8, 2020. <https://www.gov.uk/government/publications/guidelines-for-ai-procurement/guidelines-for-ai-procurement>

Veitsman, Yana, and Mareike Hartmann. “Recent Advancements and Challenges of Turkic Central Asian Language Processing.” In Proceedings of the First Workshop on Language Models for Low-Resource Languages (LoResLM 2025), 309–324. Abu Dhabi: Association for Computational Linguistics, 2025. <https://aclanthology.org/2025.loreslm-1.25/>

Yeshpanov, Rustem, Saida Mussakhojayeva, and Yerbolat Khassanov. “Multilingual Text-to-Speech Synthesis for Turkic Languages Using Transliteration.” In Proceedings of Interspeech 2023, 5521–5525. Dublin: ISCA, 2023. <https://doi.org/10.21437/Interspeech.2023-249>

