# Journal of Secure Digital Technologies
# Volume 1 (№2) September 2026

ISSN (E): 3136-9340
*Sertificate № KZ19VPY00128988, date of issue: 25.12.2025*

*The journal is published 4 times a year*
*(March, June, September, December)*

**EDITORIAL TEAM**
**Editor-in-Chief**
Dr. Leila Rzayeva, Association prof., PhD, Digital government support center (Kazakhstan)

**Deputy Editor-in-Chief**
Dr. Ali Myrzatay, PhD, Digital government support center (Kazakhstan)

**Executive Secretary**
Dr. Abdenova Zhuldyz, PhD, Digital government support center (Kazakhstan)

**EDITORIAL BOARD MEMBERS (KAZAKHSTAN)**
**Seitkulov E.N.** – Candidate of Physical and Mathematical Sciences, Professor Computer Science, Computer Engineering and Control
**Abdenova G.A.** – PhD in Automation and Control, Candidate of Technical Sciences, Professor L.N. Gumilyov Eurasian National University
**Atanov S.K.** – Doctor of Technical Sciences, Professor.
**Sergazin G.K. –** PhD, Associate Professor Computer Science, Computer Engineering and Control
**Satybaldina D.Zh.** – Candidate of Physical and Mathematical Sciences, Professor Director, Research Institute of Information Security and Cryptology, L.N. Gumilyov Eurasian National University

**INTERNATIONAL EDITORIAL BOARD MEMBERS**
**Ibraheem Shayea** – PhD in Electrical and Electronic Engineering Istanbul Technical University (Turkey)
**Yadigar Nasiboglu Imamverdiyev** – DSc in Technical Sciences, Associate Professor Information Security, Computer Science and Computer Technology (Azerbaijan)
**Erdal Irmak** – PhD in Electrical and Electronics Engineering, Professor Gazi University (Turkey)
**Korhan Kayisli** – PhD in Electrical Electronics Engineering, Associate Professor Gazi University (Turkey)
**Yonghao Wang** – PhD in Electronic Engineering, Associate Professor Queen Mary, University of London (UK)
**Amir Mosavi** – PhD in Applied Informatics, Professor of AI Obuda University (Hungary)
**Gerald Feldman** – PhD in General Computing, Associate Professor Birmingham City University (UK)
**Mohamed Ragab** – PhD in Computer Science, Assistant Professor Birmingham City University (UK)
**Seref Sagiroglu** – PhD in System Engineering, Professor Gazi University (Turkey)

**MANAGING EDITOR**
**Dr. Ali Myrzatay**, PhD, Digital government support center (Kazakhstan)

**TECHNICAL EDITOR**
**Dr. Ali Myrzatay**, PhD, Digital government support center (Kazakhstan)

# CONTENTS

Cyber Incident Damage Assessment for Public Digital Systems and Critical Infrastructure .....3

Cyber Incident Damage Assessment for Public Digital Systems and Critical Infrastructure .... 37

Verifiable but Not Correct: Blockchain Trust Boundaries for Autonomous Agents ..... 71

Calibrated Digital Twins and Geometry-based Perception for Automated Driving in Heavy Rain and Snowfall: Evidence From Research Testing in China and Implications for Kazakhstan. ..... 88

Personal Data Protection in Kazakhstan's AI Enabled Public Services ..... 109

Asking Another Jurisdiction What It Holds Without Telling It What You Hold: Private Artifact Matching with a Ledger-Anchored Transcript ..... 131

On-Chain and Off-Chain Data Placement in Hybrid Blockchain Health Systems ..... 143

An Integrated Project Management Maturity Model for Digital Qazaqstan Initiatives: Embedding Cybersecurity, AI Governance and Forward-Looking Cyber-Risk Requirements into ST RK ISO 21502-2022..... 158

Appendix A. Indicators and evidence for the baseline practice groups ..... 181

Privacy Risks and Data Protection Readiness in AI Marketing in Kazakhstan..... 183

Supplementary Materials..... 205

**Kassen Aikunim,**
Master's student,
Academy of Public Administration under the President of the Republic of Kazakhstan
Astana, Kazakhstan

## PERSONAL DATA PROTECTION IN KAZAKHSTAN'S AI ENABLED PUBLIC SERVICES

**Abstract:** Public-sector artificial intelligence can influence access to services, inspections, eligibility, and other exercises of public authority, yet compliance cannot be demonstrated by citing legal safeguards or listing technical controls alone. This study develops the Kazakhstan AI Privacy and Accountability Governance Framework (KZ-AIPAG) to connect each legal requirement and public-service use case to a defined risk, control, evidence artefact, accountable actor, and lifecycle decision. The research combines normative legal analysis, a targeted structured literature review, comparative analysis of the Canadian Algorithmic Impact Assessment, the EU AI Act, General Data Protection Regulation Articles 22 and 35, and an end-to-end algorithmic-audit model, together with bounded applications to Smart Data Ukimet and eGov AI. The results show that Kazakhstan's principal gap is operational rather than a complete absence of regulation. KZ-AIPAG separates four use-case impact tiers from probability-by-impact hazard scoring, organises controls across six governance domains, and introduces an evidence chain, four decision gates, a unified role catalogue, and a RACI matrix that preserves audit independence. In the illustrative Smart Data Ukimet case, the framework yields a conditional no-go decision until the legal basis, data purpose, linkage and re-identification risk, access model, meaningful human review, and control-effectiveness evidence are documented. The framework is a proposed and analytically demonstrated instrument, not an empirically validated one. Its scientific contribution is a Kazakhstan-specific compliance-by-evidence model that makes accountability testable at service level; independent expert validation and operational pilots remain necessary.

**Keywords:** *artificial intelligence; personal data; digital public services; algorithmic decision-making; human oversight; explainability; accountability; risk governance; Kazakhstan*

## Защита персональных данных в государственных сервисах Казахстана с ИИ

**Аннотация:** Искусственный интеллект в государственном секторе может влиять на доступ к услугам, проверки, определение права на поддержку и иные формы осуществления публичной власти, однако соблюдение требований нельзя доказать лишь ссылкой на нормы права или перечнем технических мер. Цель исследования - разработать Казахстанскую модель управления конфиденциальностью и подотчётностью искусственного интеллекта KZ-AIPAG, которая связывает каждое правовое требование и сценарий государственной услуги с риском, мерой контроля, доказательным артефактом, ответственным субъектом и решением на этапе

жизненного цикла. Методология объединяет нормативно-правовой анализ, целевой структурированный обзор литературы, сопоставление канадской оценки алгоритмического воздействия, Акта ЕС об искусственном интеллекте, статей 22 и 35 Общего регламента по защите данных и сквозной модели алгоритмического аудита, а также ограниченные кейсы Smart Data Ukimet и eGov AI. Результаты показывают, что основной пробел в Казахстане имеет операционный характер и не сводится к отсутствию регулирования. KZ-AIPAG разделяет четыре уровня воздействия сценария и оценку отдельных угроз по формуле вероятности и последствий, объединяет шесть областей управления и вводит доказательную цепочку, четыре контрольные точки, единый каталог ролей и матрицу RACI, сохраняющую независимость внутреннего аудита. Для иллюстративного кейса Smart Data Ukimet получено условное решение об отказе в запуске до подтверждения правового основания, цели обработки, рисков связывания и повторной идентификации, модели доступа, содержательного человеческого контроля и эффективности мер. Научный вклад состоит в контекстной модели compliance-by-evidence, делающей подотчётность проверяемой на уровне услуги. Независимая экспертная валидация и пилотирование остаются необходимыми.

**Ключевые слова:** *искусственный интеллект; персональные данные; цифровые государственные услуги; алгоритмическое принятие решений; человеческий контроль; объяснимость; подотчётность; управление рисками; Казахстан*

## Қазақстанның ЖИ қолданатын мемлекеттік қызметтеріндегі дербес деректерді қорғау

**Аңдатпа:** Мемлекеттік сектордағы жасанды интеллект қызметтерге қол жеткізуге, тексерулерге, қолдау алу құқығын айқындауға және жария билікті жүзеге асырудың өзге де нысандарына ықпал етуі мүмкін, алайда талаптардың сақталуын тек құқықтық нормаларға сілтеме жасау немесе техникалық шараларды тізбелеу арқылы дәлелдеу жеткіліксіз. Зерттеудің мақсаты - әрбір құқықтық талап пен мемлекеттік қызмет сценарийін тәуекелмен, бақылау шарасымен, дәлелдік артефактпен, жауапты тұлғамен және өмірлік цикл шешімімен байланыстыратын Қазақстанның KZ-AIPAG жасанды интеллект құпиялылығы мен есептілігін басқару моделін әзірлеу. Әдіснама нормативтік-құқықтық талдауды, нысаналы құрылымдалған әдебиет шолуын, Канаданың алгоритмдік әсерді бағалау құралын, Еуропалық Одақтың Жасанды интеллект туралы актісін, Деректерді қорғаудың жалпы регламентінің 22 және 35-баптарын және алгоритмдік аудиттің толық циклдік моделін салыстыруды, сондай-ақ Smart Data Ukimet пен eGov AI бойынша шектелген кейстерді қамтиды. Нәтижелер Қазақстандағы негізгі олқылық реттеудің толық болмауынан емес, талаптарды қызмет деңгейінде іске асыру мен дәлелдеудің жеткіліксіздігінен туындайтынын көрсетті. KZ-AIPAG пайдалану сценарийінің төрт әсер деңгейін жекелеген қауіптерді ықтималдық пен салдар бойынша бағалаудан бөледі, басқарудың алты саласын

біріктіреді және дәлелдер тізбегін, төрт шешім қақпасын, бірыңғай рөлдер каталогын әрі ішкі аудит тәуелсіздігін сақтайтын RACI матрицасын ұсынады. Smart Data Ukimet иллюстрациялық кейсінде құқықтық негіз, өңдеу мақсаты, деректерді байланыстыру және қайта сәйкестендіру тәуекелі, қол жеткізу моделі, мазмұнды адам бақылауы мен бақылау тиімділігі дәлелденгенге дейін шартты түрде іске қоспау шешімі алынды. Ғылыми үлес - есептілікті қызмет деңгейінде тексерілетін ететін Қазақстанға бейімделген compliance-by-evidence моделі. Тәуелсіз сараптамалық валидация және практикалық пилоттау қажет.

**Түйін сөздер:** жасанды интеллект; дербес деректер; цифрлық мемлекеттік қызметтер; алгоритмдік шешім қабылдау; адам бақылауы; түсіндірмелілік; есептілік; тәуекелдерді басқару; Қазақстан

## 1. Introduction (Literature Review)

A public-sector AI service differs from a private application because its output may be embedded in the exercise of public authority: it may prioritise an inspection, propose a support measure, prepare a recommendation for a public official, or affect access to a service. The scale and inter-agency integration of data increase not only the value of analytics but also the consequences of error, secondary use, excessive access, and opaque delegation of decision-making. Personal data protection therefore cannot be reduced to encryption or consent; it also encompasses the legality of the purpose, data quality, user competence, reviewability, and institutional accountability.

Kazakhstan's current legal framework contains substantive safeguards. Article 21 of the Constitution guarantees protection against the unlawful collection, processing, storage, and use of personal data, including through digital technologies [1]. Article 19-1 of the Law on Personal Data and Their Protection governs automated processing that affects rights and legitimate interests, requires an explanation, permits an objection, and establishes a three-working-day period for its consideration [2]. The Law on Artificial Intelligence allocates duties to system owners and operators and provides for risk management, audit, and user information [3]. Article 43 of the Digital Code establishes non-discrimination, information rights, an explanation of key factors, and human review of consequential fully automated decisions [4]. The proposition that Kazakhstan faces a complete regulatory vacuum is therefore not supported.

The article's central thesis is that demonstrable personal data protection does not arise merely from the existence of an individual rule or technical control. It arises from a traceable chain in which the service purpose determines permissible data; the type of impact determines the risk tier; the risk determines mandatory controls and human oversight; each control has an owner, an artefact, and an acceptance criterion; and any material change to the model, data, or purpose triggers reassessment.

The study addresses four research questions. RQ1: Which regulatory, governance, privacy, and human-factor gaps can be substantiated across the public-sector AI lifecycle? RQ2: How should AI use cases be classified according to their effects on rights? RQ3: What

architecture should a context-specific, multi-layer governance framework adopt? RQ4: How can the framework be implemented and its effectiveness measured? The object of the research is the use of AI in the delivery of digital public services. The subject is the set of legal, organisational, technological, and managerial mechanisms used to protect personal data in such services.

### 1.1 Conceptual framework

Artificial intelligence is understood in the functional sense adopted by the Law on Artificial Intelligence: a system capable of generating outputs, for stated objectives, that influence a physical or virtual environment [3]. A public-sector AI service is an organisational and technical service, or a component of a process, in which a confirmed AI system produces an answer, classification, forecast, recommendation, or decision for a public authority or citizen. A digital public service is broader and may be entirely rule-based without using AI.

Conventional automation executes predetermined rules without statistical inference. Algorithmic automation encompasses any digital system that makes or influences a decision through automated data processing; under Article 43 of the Digital Code, this category includes, but is not limited to, AI [4]. AI-assisted decision-making leaves the legally consequential determination to an authorised official, whereas a fully automated decision is made without human participation in assessing the circumstances or approving the outcome. A confirmation button does not constitute meaningful human oversight if the official lacks sufficient time, information, competence, or authority to reject the recommendation.

Profiling means using data to evaluate or predict characteristics of an individual; personalisation means adapting an interface or offer; a recommender system produces a ranked option; a predictive model forecasts an event; and automated assessment assigns a category or score. In this study, explainability means providing the affected person with an intelligible account of the key factors, limits, and consequences of an output. Accountability means the organisation's capacity to identify the responsible actor and produce evidence that the relevant obligation has been fulfilled.

The scope is limited to confirmed AI or algorithmic decision systems in public services; ordinary rule-based digitisation is excluded. Kazakhstan's Data Management Requirements, OECD and OECD/UNESCO instruments, the NIST AI RMF and de-identification guidance, and ISO/IEC standards on AI, privacy, risk, and information security provide implementation benchmarks rather than binding Kazakhstan law [5]–[15].

### 1.2 Literature synthesis and research gap

The core peer-reviewed corpus covers organisational AI governance, systematic governance synthesis, privacy- and security-aware ethical assessment, public-service privacy behaviour, automated administrative responsibility, interaction-based governance, fairness assessment in public organisations, and Kazakhstan profiling law [16]–[23].

The literature converges on lifecycle governance, distributed responsibility, privacy- and security-aware assessment, meaningful human oversight, and contestability [16]–[22].

It also shows that convenience does not prove rights compliance, that general principles are reshaped in implementation, and that public-sector AI must be treated as a sociotechnical service rather than a model in isolation [19]–[22], [28]–[36]. The Kazakhstan study on profiling is used as historical context because it predates the 2025–2026 reforms [23].

The research gap lies in the absence, within the reviewed corpus, of an integrated model for Kazakhstan's updated legal and institutional context that translates rights and general risk principles into service-level evidence, distinguishes AI from other forms of automation, accounts for inter-agency architecture, assigns an accountable actor, and links reassessment to material system changes. The review does not establish global priority for the proposed model; novelty is claimed as a context-specific design and applied analytical instrument.

The objective is to develop and analytically demonstrate a Kazakhstan-specific governance framework that converts legal and governance requirements into testable service-level accountability.

The scientific novelty lies in the compliance-by-evidence unit of analysis: rule → use case → risk → control → evidence artefact → accountable owner → gate decision. Unlike a general checklist, the model separates the governance tier of a public-authority use case from the score of an individual hazard and makes reassessment compulsory after a material change.

The study contributes: (1) a four-tier public-sector AI use-case classification; (2) an AI-privacy risk matrix that distinguishes inherent, target, and verified residual risk; (3) the six-domain KZ-AIPAG architecture with four decision gates; and (4) a role and evidence model that preserves internal-audit independence and operationalises redress.

The remainder of the paper is organised as follows: Section 2 explains the research design and materials; Section 3 reports the legal, comparative, framework, and case results; Section 4 answers the research questions and discusses implications and limitations; and Section 5 concludes the study.

## 2. Methods and Materials

### 2.1 Research design and unit of analysis

The study uses qualitative normative legal analysis, a targeted structured literature review, comparative framework analysis, and two bounded case studies. The unit of analysis is the chain 'rule or principle → use case → risk to a right → control → evidence artefact → role → lifecycle gate'. The comparative legal analysis is retained because Table 1 applies common comparison dimensions to Canada, the European Union, GDPR, and Raji et al., and then identifies the distinct function of KZ-AIPAG.

### 2.2 Sources, search, and selection criteria

The literature search was updated on 29 September 2026. Crossref metadata and the publisher platforms ACM Digital Library, SpringerLink, ScienceDirect, and Oxford Academic were searched, with backward citation checking and separate searches of official

legal and policy portals. Four English-language search strings were used: 'artificial intelligence public sector governance'; 'algorithmic impact assessment public administration'; 'algorithmic auditing accountability public services'; and 'Kazakhstan artificial intelligence personal data'. The first 50 Crossref results ranked for each query were screened (200 records); DOI/title deduplication left 195 unique records. A purposive set of 18 peer-reviewed works was retained after relevance and metadata verification. The legal corpus comprised the operative official texts and official implementation materials cited in the references.

Inclusion required a peer-reviewed work or primary official source, a direct connection to public-sector AI, personal data, impact assessment, algorithmic audit, human oversight, accountability, or Kazakhstan law, and verifiable bibliographic metadata. Exclusion criteria were duplicate records, purely technical model-performance papers without governance relevance, marketing materials, unsourced commentary, and outdated legal text. This is a targeted structured review, not a PRISMA systematic review; Crossref's relevance-ranked sample and the English-language search create selection limitations.

[ТРЕБУЕТСЯ ДОПОЛНИТЬ: приложить журнал поиска или экспорт результатов Crossref/издательских платформ, подтверждающий дату поиска, 200 просмотренных записей, 195 уникальных записей и 18 включённых работ. До подтверждения эти числа следует рассматривать как заявленный, а не независимо проверенный протокол.]

### 2.3 Coding and analytical procedure

Table 1. Coding framework

| Code family | Recorded fields | Decision rule | Output |
|---|---|---|---|
| Legal rule | Instrument, article, effective date, addressee, duty, exception | Use current official text; distinguish binding rule from guidance | Legal map |
| Use case | Purpose, data, output, user, affected person, decision effect | Classify as AI only where the source or technical evidence confirms AI | System card and scope statement |
| Tier | Legal effect, material influence, sensitivity, scale, reversibility, review | Assign the highest Tier criterion triggered | Tier decision with rationale |
| Risk | Hazard, cause, affected right, P, I, evidence, uncertainty | Score inherent risk; record proposed target after controls | Risk register |
| Control and evidence | Preventive/detective/corrective control, owner, test, artefact | A control is effective only after design and operating-effectiveness testing | Evidence pack and gate result |

| Code family | Recorded fields | Decision rule | Output |
|---|---|---|---|
| Role | R, A, C, I, independence conflict | Exactly one A; internal audit cannot own an operational process it audits | RACI matrix |

Risk scores were assigned by the lead author through structured analytical judgement, not by an expert panel. For each of eight risks, the author applied the published 1-5 anchors, recorded the legal and operational assumptions, and performed a second consistency pass against the tier rule and control logic. No independent human raters scored the cases; consequently, inter-rater reliability is not reported. Section 13 specifies the independent validation required before empirical claims can be made.

### 2.4 Case boundaries and validation status

The study's evidence is limited to publicly available material. Non-public contracts, data-flow diagrams, model cards, logs, impact assessments, access reviews, and audit results were unavailable. The case application is therefore an analytical demonstration of the framework, not an audit finding about the actual operating effectiveness of Smart Data Ukimet or eGov AI.

KZ-AIPAG has undergone analytical consistency checking by the lead author but has not been independently validated. A preregistered panel of 10–15 specialists in Kazakhstan law, AI/model validation, cybersecurity, audit, and public administration should independently classify 8–12 standardised scenarios, score probability and impact, assess control sufficiency, and identify blocking conditions. Fleiss' kappa should assess categorical agreement and Kendall's W should assess ordinal ratings. Until that study and an operational pilot are completed, the framework and target scores remain analytically demonstrated hypotheses rather than validated measurements.

## 3. Results

### 3.1 Legal and institutional findings

The comparison produces three principal findings. First, substantive legal safeguards already exist: purpose specification and proportionality, data protection, conditions governing automated processing, explanation, objection, and review [2], [4]. Second, the Law on Artificial Intelligence adds obligations concerning system ownership, risk management, audit, and transparency [3]. Third, an implementation gap remains: legislation does not automatically convert each obligation into a coherent set of service-level records, roles, and tests. This gap may be addressed through secondary regulation, departmental governance, and system architecture. Legislative amendment is warranted only after establishing that lower-level mechanisms cannot deliver the required outcome.

The principal conclusion is that the most defensible gap is a governance-and-evidence gap. High-impact use cases would additionally benefit from consistent requirements for AI impact assessment, a public register or system card, overseer

competence, and minimum audit-trail content. Before proposing legislative amendments, policymakers should determine whether these requirements can be introduced through sectoral rules, digital-system requirements, and procurement contracts.

Temporal applicability is stated explicitly. The Constitution adopted at the referendum of 15 March 2026 entered into force on 1 July 2026 [1]. The Law on Artificial Intelligence No. 230-VIII of 17 November 2025 was first officially published on 18 November 2025 and entered into force on 18 January 2026 [3]. Article 19-1 of the Personal Data Law was introduced by Law No. 231-VIII of 17 November 2025; that amending law also entered into force after sixty calendar days, on 18 January 2026 [2]. The Digital Code No. 255-VIII of 9 January 2026 entered into force on 12 July 2026 [4]. The EU AI Act entered into force on 1 August 2024 and generally applies from 2 August 2026, subject to staged application [25]; GDPR has applied since 25 May 2018 [26].

### 3.2 Comparative position and Kazakhstan-specific contribution

The Canadian Directive on Automated Decision-Making operationalises four impact levels and scales notice, explanation, peer review, human intervention, and recourse requirements to the level [24]. The EU AI Act uses a different architecture: prohibited practices, high-risk classification and obligations, transparency duties for specified systems, and a residual category rather than a single four-level scoring matrix [25]. GDPR Article 22 protects individuals against solely automated consequential decisions and requires safeguards in defined exceptions, while Article 35 requires a data protection impact assessment where processing is likely to create high risk [26]. Raji et al. organise internal algorithmic auditing across the system lifecycle and treat documentation produced at each stage as part of an integrated audit report [27].

Table 2. Comparative position and Kazakhstan-specific contribution

| **Instrument** | **Classification or trigger** | **Core mechanism** | **What KZ-AIPAG retains** | **What KZ-AIPAG adds for Kazakhstan** |
|---|---|---|---|---|
| Canada AIA and Directive | Four impact levels I-IV based on decision effects and reversibility | Mandatory questionnaire and proportionate requirements | Four escalating use-case tiers and proportional controls | Direct mapping to Kazakhstan legal duties, public-body roles, evidence artefacts, and gates |
| EU AI Act | Prohibitions, high-risk systems, specified transparency duties, other systems | Provider/deployer obligations, registration, conformity and oversight | Risk-based escalation, transparency, documentation, monitoring | Service-level evidence chain compatible with Kazakhstan institutions; no claim of EU legal equivalence |

| Instrument | Classification or trigger | Core mechanism | What KZ-AIPAG retains | What KZ-AIPAG adds for Kazakhstan |
|---|---|---|---|---|
| GDPR Arts. 22 and 35 | Solely automated significant decisions; processing likely to create high risk | Safeguards, human intervention, contestation, DPIA | Human review, objection, impact assessment, redress | Links Art. 19-1 of Kazakhstan's Personal Data Law to a three-working-day review workflow |
| Raji et al. 2020 | Lifecycle audit of organisational decisions and artefacts | End-to-end internal algorithmic audit | Lifecycle evidence and independent challenge | Single accountable gate owner, national legal map, and procurement/agency evidence pack |
| KZ-AIPAG | Highest triggered Tier 1-4 criterion; separate P × I hazard scoring | Six governance domains, evidence chain, gates, RACI, monitoring | Integrated operational model | Kazakhstan-specific bridge from law to service decision, including the Digital Code and AI Law |

The novelty is therefore limited but identifiable. KZ-AIPAG does not claim a new universal theory of AI risk. Its contribution is a jurisdiction-specific compliance-by-evidence design: it joins Kazakhstan's rules on personal data, automated processing, algorithmic systems, AI risk management, audit, transparency, ownership duties, and liability to operational artefacts and public-sector decision authority. This resolves a gap left by frameworks that are either jurisdiction-neutral, focused on providers, or not connected to a public service's RACI and legal review clock.

### 3.3 Use-case classification and risk assessment

Tier assignment is an ex ante classification of the use case, not the arithmetic result of P × I. The assessor tests six dimensions: legal or similarly significant effect; material influence on a public official; personal-data sensitivity; population scale; reversibility; and meaningful review. The highest criterion triggered determines the tier. Controls may reduce a hazard score, but they do not automatically lower the tier because the underlying public-authority function remains consequential.

The relationship to P × I is two-dimensional. Tier sets the governance floor for the service. P × I (1-25) ranks individual hazards and treatment urgency within that service. A Tier 4 service cannot be treated as low governance merely because a proposed control yields a target score of 4; conversely, a Tier 1 service with a critical cybersecurity hazard must treat the hazard before release. A tier may be reduced only when the use case itself is redesigned

so that the triggering effect, data, scale, irreversibility, or review deficit is removed and documented.

Inherent risk is assessed before additional controls as probability P × impact I. P uses five evidence-anchored categories: 1 rare under exceptional conditions; 2 unlikely in ordinary operation; 3 possible; 4 likely because of scale, exposure, or repeated use; and 5 expected or recurring. I equals 1 negligible, 2 minor and readily reversible, 3 material but reversible, 4 serious or difficult to reverse, and 5 severe, large-scale, or legally consequential.

Scores are interpreted as 1-4 low, 5-9 moderate, 10-16 high, and 17-25 critical. The revised matrix reports target residual risk, not observed residual risk, because the proposed controls have not been tested. A target becomes a residual score only after evidence demonstrates both control design and operating effectiveness. Each P and I value must be accompanied by a scenario, data source, control assumption, assessor, and date; unsupported numerical precision is prohibited.

Acceptance rules are: critical inherent risk blocks deployment pending treatment; high target risk requires senior written acceptance and independent challenge; Tier 4 requires senior approval and effective redress regardless of arithmetic score; and any unresolved legality, discriminatory effect, or absence of meaningful human review is a blocking condition rather than a risk that can be averaged away.

Table 3. Tier assignment rule and minimum governance

| **Tier** | **Binding assignment rule** | **Illustrative use** | **Minimum governance** |
|---|---|---|---|
| 1 Informational | No individual legal effect; no material influence; no sensitive profiling; error readily reversible. | Generic search, FAQ, non-personal drafting. | Owner, source register, notice, basic security, error channel. |
| 2 Assisted processing | Processes personal data or structures work, but does not materially determine a consequential outcome. | Document extraction, routing, personalised navigation. | Data map, purpose register, quality test, access control, human correction. |
| 3 Material decision support | Output materially influences exercise of public authority, access, inspection, eligibility, or risk treatment; meaningful review remains available. | Risk flag, recommendation, fraud/anomaly triage. | Combined AI/privacy impact assessment, validation, subgroup testing, reviewer protocol, independent challenge, monitoring. |
| 4 Determinative or difficult-to-reverse | Solely automated or effectively determinative consequential outcome, or high-scale sensitive profiling with ineffective or delayed review. | Automated eligibility, status, sanction, or similarly significant classification. | Explicit legal basis, senior approval, independent assurance, effective human intervention, explanation and redress, enhanced audit and publication. |

The eight scenarios below are analytical risk hypotheses. P = 2 is not an empirical frequency estimate; it is a target of 'unlikely' conditional on the stated control package. The justification differs for each risk and must be replaced by observed evidence after operation.

Table 4. AI privacy risks and target residual scores

| **Risk and inherent score** | **Control package** | **Why target P = 2 is plausible** | **Target P × I** | **Verification needed** |
|---|---|---|---|---|
| R1 Prompt leakage 3 × 3 = 9 | Constrained interface, data-flow map, no-training default, DLP tests, short retention. | Exposure should require a configuration failure, bypass, or exceptional input rather than normal use. | 2 × 2 = 4 | Prompt sampling, vendor configuration evidence, leakage test. |
| R2 Extraction error 3 × 3 = 9 | Gold dataset, confidence threshold, sampling, correction route. | Routine errors should be detected or corrected; undetected material error remains possible in atypical documents. | 2 × 2 = 4 | Field-level error rates by document type and subgroup. |
| R3 Secondary use 4 × 4 = 16 | Field-purpose register, compatibility review, separated access, change approval. | Unauthorised reuse should require an approved-purpose bypass or governance failure. | 2 × 3 = 6 | Purpose audit and access-log sample. |
| R4 False positive 3 × 4 = 12 | Calibrated thresholds, subgroup tests, proportionate investigation, human corroboration. | A false signal remains possible, but adverse action should not follow without corroboration. | 2 × 3 = 6 | Prospective false-positive and outcome analysis. |
| R5 Automation bias 4 × 4 = 16 | Reviewer authority, adequate time, contrary evidence, reason record, override monitoring. | Rubber-stamping should become exceptional if reviewers must evidence independent reasoning and overrides are reviewed. | 2 × 3 = 6 | Decision-file sampling and reviewer observation. |
| R6 Inadequate explanation 4 × 5 = 20 | Reason template, factor trace, specialist review, three-working-day objection route. | A deficient explanation should be uncommon when generated reasons are checked against the decision record. | 2 × 4 = 8 | Comprehension test and appeal-file audit. |

| Risk and inherent score | Control package | Why target P = 2 is plausible | Target P × I | Verification needed |
|---|---|---|---|---|
| R7 Re-identification 3 × 5 = 15 | Controlled environment, linkage testing, output checking, data minimisation. | Re-identification should require an atypical linkage or control failure, not routine analyst access. | 2 × 4 = 8 | Independent linkage and singling-out test. |
| R8 Excess privilege 4 × 5 = 20 | RBAC/ABAC, quarterly recertification, privileged monitoring, alerts, revocation SLA. | Persistent excess access should be unlikely if entitlement changes and dormant privileges are detected promptly. | 2 × 4 = 8 | Access recertification sample and alert-response evidence. |

The common target probability does not imply a common causal mechanism. It represents the same ordinal end-state – unlikely in ordinary operation – reached through different controls and tested through different evidence. If any verification result fails, the target score must be withdrawn and the risk re-estimated.

### 3.4 KZ-AIPAG architecture and lifecycle

KZ-AIPAG is reorganised into three connected views. The decision view moves from mandate and scope to tier, risks, controls, evidence, independent challenge, and a gate decision. The domain view assigns six concurrent governance domains: legal and rights; data; AI and model; cybersecurity; human and organisational; and assurance, audit, and redress. The lifecycle view ensures that approval is revisited after a material change, incident, complaint, drift signal, or new data source.

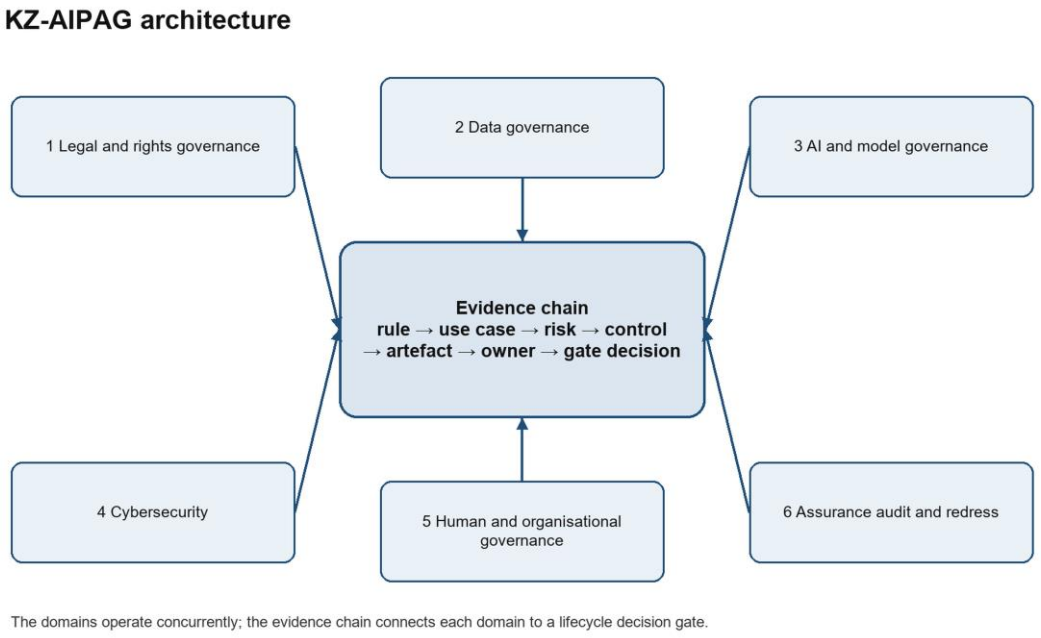



Fig. 1. Revised KZ-AIPAG architecture.

The central evidence chain is the model's unit of accountability: rule → use case → risk → control → artefact → owner → gate decision. A gap at any link prevents a reviewer from determining whether the legal duty was discharged. The six domains do not operate as a sequence; each contributes evidence to the same decision file.

Four gates make the structure operational. Gate 1 approves scope and tier. Gate 2 approves design and planned treatments. Gate 3 confirms verification and readiness. Gate 4 decides continued operation after monitoring or a trigger event. The accountable actor signs a pass, conditional pass, restriction, or no-go decision; conditions must identify an owner and due date.

This structure incorporates continuous risk management from NIST and ISO, impact assessment from Canada and GDPR, transparency and high-risk governance from the EU AI Act, and lifecycle audit documentation from Raji et al. [8], [11], [12], [24]–[27]. Its jurisdiction-specific element is the Kazakhstan legal map and the service-level evidence and decision authority attached to it.

The eighteen implementation activities are consolidated into seven stages: mandate and purpose; legal basis and tier; data mapping and impact assessment; access, oversight, and security design; testing and independent validation; approval and deployment; and monitoring, incident response, reassessment, or retirement. Each stage produces a named artefact and has a transition criterion.

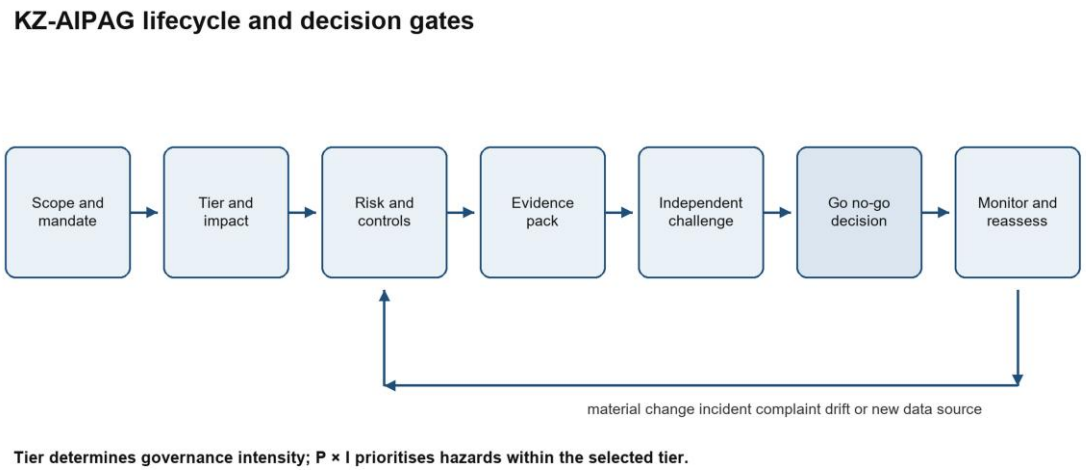



Fig. 2. KZ-AIPAG lifecycle and decision gates.

### 3.5 Responsibility allocation

The authoritative role catalogue is: agency head or delegated senior risk owner; service owner; legal and privacy function; data owner; AI/model owner; AI and data specialists; IT operations; cybersecurity function; procurement/vendor-management function; redress officer; independent validator; internal audit; and vendor. The same labels are used in the lifecycle, RACI, risk register, and evidence templates. R performs the work, A is the single operationally accountable approver, C is consulted, and I is informed.

Table 5. RACI for core KZ-AIPAG processes

| **Process** | **A** | **R** | **C** | **I** |
|---|---|---|---|---|
| Purpose and legal basis | Agency head | Service owner; legal/privacy | Data owner; AI owner | Internal audit; vendor |
| Tier and impact assessment | Service owner | AI owner; legal/privacy | Data owner; cybersecurity; affected-function representative | Agency head; internal audit |
| Data mapping and quality | Data owner | AI/data specialists | Service owner; legal/privacy; cybersecurity | AI owner; internal audit |
| Model validation | Independent validator | AI/data specialists | AI owner; service owner; legal/privacy | Agency head; internal audit |

| Process | A | R | C | I |
|---|---|---|---|---|
| Access and cybersecurity | Cybersecurity lead | IT operations | Data owner; service owner; legal/privacy | Agency head; internal audit |
| Human oversight | Service owner | Authorised human reviewers | Legal/privacy; AI owner; HR/learning function | Agency head; internal audit |
| Redress and correction | Service owner | Redress officer; legal/privacy | Data owner; AI owner; IT | Agency head; internal audit |
| Incident response | Service owner | Cybersecurity; IT operations | Legal/privacy; data owner; vendor | Agency head; internal audit |
| Independent audit | Head of internal audit | Internal audit team | Service owner; legal/privacy; cybersecurity | Agency head; oversight body |

Internal audit is not accountable for redress, deployment, control operation, or risk acceptance. Its accountability is limited to the independent-audit engagement and its quality. Operational ownership of redress rests with the service owner, while the redress officer and legal/privacy function perform review and correction. This preserves the organisational independence needed for subsequent assurance.

### 3.6 Bounded case-study results

Smart Data Ukimet is publicly described as an inter-agency platform for collecting, storing, and comprehensively analysing public-sector data, with a data lake, hashing, and depersonalisation functions [37]. Public sources do not establish every model, dataset, decision path, or control. The application below therefore uses a bounded scenario: linked inter-agency data are analysed to produce a risk signal that may influence an individual public-service review. If evidence shows that the output is only aggregate and cannot materially influence an individual outcome, the tier must be reduced and documented.

Table 6. End-to-end KZ-AIPAG application to Smart Data Ukimet

| Step | Assessment | Required measures | Evidence or document | Gate decision |
|---|---|---|---|---|
| Scope and tier | Linked data and a risk signal may materially influence an individual review: Tier 3. Escalate to Tier 4 if the output is determinative or review is ineffective. | Confirm legal mandate, decision effect, affected groups, and reviewer authority. | Purpose statement; system card; tier rationale; legal map. | Gate 1 scope: conditional pass only if the individual decision path is documented. |
| Risks | R3 secondary use; R4 false positive; R5 automation bias; R7 re-identification; R8 excessive privilege. | Field-purpose mapping; subgroup/error testing; reviewer protocol; linkage test; least-privilege design. | Risk register; feature-purpose register; validation report; re-identification report; access matrix. | Gate 2 design: no-go while any high legal or re-identification risk lacks treatment. |
| Control verification | Proposed controls reduce target risk, but operating effectiveness is unknown. | Sample decisions, test access recertification, challenge linkage, test correction and override paths. | Test scripts; samples; access review; issue log; signed remediation evidence. | Gate 3 assurance: pass only after blocking findings close. |
| Deployment decision | Public value does not override unresolved rights risk. | Record accepting authority, conditions, rollback, monitoring thresholds, and complaint channel. | Signed go/no-go record; release and rollback plan; public notice; redress procedure. | Current analytical outcome: conditional no-go pending the evidence listed above. |
| Operation | New datasets, changed purposes, drift, complaints, or incidents can invalidate the assessment. | Monitor false positives, overrides, access anomalies, complaints, and linkage changes; reassess on trigger. | Monitoring dashboard; decision log; complaint register; reassessment record. | Gate 4 continuation: suspend or restrict use when a threshold is exceeded. |

The decision is deliberately conservative. It does not assert that Smart Data Ukimet is unlawful or unsafe; it shows that a deployment decision cannot be justified from public descriptions alone. The missing evidence is itself a governance finding under KZ-AIPAG because the gate owner cannot verify legality, proportionality, re-identification resistance, meaningful review, or access control.

An official announcement describes eGov AI as an intelligent search function in eGov Mobile that helps users locate public services [38]. On the public description, the baseline classification is Tier 1 if no identifier, profile, or consequential decision is involved, and Tier 2 if queries are retained, personalised, or routed through components that process

personal data. Minimum evidence includes a prompt data-flow map, retention and training configuration, authoritative source register, multilingual quality tests, privacy notice, and human hand-off. Any later use to recommend eligibility, prioritise a case, or influence an outcome requires reassessment to Tier 2 or 3.

### 3.7 Human factors and performance indicators

A technically secure system remains institutionally vulnerable when a public servant does not understand the model's limits, holds excessive privileges, or merely confirms its recommendation. Automation bias can turn a human-in-the-loop arrangement into rubber-stamping; insufficient data literacy impedes recognition of poor data quality; weak cyber literacy increases phishing and improper-sharing risks; and diffuse responsibility allows the process owner to shift the substantive decision to the developer.

KZ-AIPAG requires a role-based competence matrix. End users must understand the system's purpose, uncertainty, and override procedure; data owners, provenance and quality; legal and privacy staff, applicable rules and redress; AI specialists, validation and drift; and agency heads, residual risk. For high-risk systems, the right to override, minimum review time, access to source factors, and protection against adverse consequences for a justified rejection of the recommendation must be procedurally guaranteed. Training completion is not a sufficient performance indicator without scenario-based competence testing.

Performance should be measured through complementary indicators: the proportion of registered systems with an approved tier; current impact assessments for high-risk systems; sampled decisions with accessible review records; justified override rates; expired or excessive privileges per 1,000 accounts; redress timeliness; consequential errors per 1,000 outputs; the distribution of verified residual risk; scenario-based staff competence; and timely audit closure. Each measure requires its source register and interpretive limitation; document existence, course attendance, zero overrides, or an average risk score cannot establish effectiveness on their own.

## 4. Discussion

### 4.1 Direct answers to the research questions

RQ1. The analysis does not support a regulatory-vacuum thesis: the applicable instruments establish purpose limitation and proportionality, data protection, conditions for consequential automated processing, explanation, objection, non-discrimination, and human review. The substantiated gap primarily concerns operationalisation, evidence, coordination of inter-agency roles, and control of material changes. Any sector-specific legal gaps must be established at the level of a particular service.

RQ2. The study develops a four-tier classification in which the tier depends on decision effect, material influence, data sensitivity, scale, reversibility, and meaningful review. Tier sets the governance floor; P × I prioritises individual hazards within that tier. The revised AI-Privacy Risk Matrix connects the risk source, affected right, legal rule,

control, evidence, owner, inherent risk, and target residual risk. Because control effectiveness was not tested, the target scores must not be reported as observed residual risk.

RQ3. KZ-AIPAG comprises Legal and Rights Governance, Data Governance, AI Governance, Cybersecurity, Human and Organisational Governance, and Assurance, Audit, and Redress. Unlike a compilation of NIST, OECD, and ISO recommendations, it introduces a Kazakhstan-specific evidence chain and service gates, under which transition is permitted only when a verified artefact and an accountable actor are present.

RQ4. The framework is implemented through eighteen stages from public-purpose definition to retirement. Its effectiveness is measured through coverage, quality, oversight, redress, access, drift, residual risk, competence, and audit closure. Potential applications include departmental regulation, procurement control, project appraisal, and internal audit. The principal limitations are the absence of non-public documentation and operational validation. Future research should pilot the framework in a specific public service, validate the scoring method, compare language groups, and test whether the evidence chain reduces actual errors and the time required to restore affected rights.

### 4.2 Interpretation and relation to prior research

The findings support prior research on lifecycle governance and distributed responsibility [16], [17], [22], while refining it for Kazakhstan: substantive rights are already established, so the scientific and practical task shifts from declaring principles to constructing an evidence chain. This is consistent with NIST, where governance informs mapping, measurement, and management [8]. KZ-AIPAG nevertheless makes the administrative consequence and the citizen's practical ability to secure review explicit objects of assurance.

The case studies reveal a distinction between data-centric and interaction-centric risks. For Smart Data Ukimet, the material issues are linkage, purpose, inter-agency access, and re-identification. For eGov AI, they are prompt privacy, output reliability, multilingual quality, and safe hand-off. A single checklist is therefore inappropriate. Nor does the analysis support the proposition that a large volume of data is inherently unlawful: risk arises from disproportionate purpose, opaque secondary use, poor quality, excessive access, or the absence of effective redress.

KZ-AIPAG differs from the compared instruments in its unit of accountability and its national legal linkage. Canada provides a mature four-level impact tool; the EU AI Act supplies a binding risk-and-obligations architecture; GDPR provides individual-decision safeguards and DPIA; and Raji et al. provide an end-to-end audit logic [24]–[27]. KZ-AIPAG combines these functions only where compatible and adds a Kazakhstan-specific chain linking Articles 15, 20, 21, and 30 of the AI Law, Article 19-1 of the Personal Data Law, and the Digital Code to a named public-service owner, evidence artefact, and gate decision. This is an implementation contribution, not evidence that the framework outperforms its comparators.

The theoretical contribution is a compliance-by-evidence model for public-sector AI. Human oversight is treated as an institutional capacity to change an outcome, and accountability as a link between the A role, an evidence artefact, and a lifecycle gate. The practical implication is that implementation should begin not with the acquisition of controls, but with an inventory of use cases and one end-to-end path from data collection to redress.

An alternative explanation is that the observed implementation gap reflects limited public documentation rather than deficient internal governance. The framework therefore treats missing public evidence as a condition that prevents an external assurance conclusion, not as proof that a public body lacks the relevant control.

### 4.3 Practical and policy implications

Implementation should begin with a service evidence pack and one accountable actor for every gate. Dataset-specific linkage tests are required after new inter-agency combinations; meaningful human review must record authority, review time, reasons, and overrides; procurement contracts should secure audit rights, change notification, and limits on vendor data use; and internal audit should test operating effectiveness through sampled decisions rather than document existence. These measures can be introduced through departmental rules, system requirements, and procurement controls before legislative amendment is proposed. Their performance indicators are evidence-pack completeness, gates with an assigned accountable role, reassessments after linkage changes, quality-sampled override records, compliant contracts, and timely closure of audit findings.

### 4.4 Limitations

The study is limited by an English-language, relevance-ranked search; reliance on public legal and service descriptions; author-assigned ordinal risk scores; and the absence of contracts, system cards, data-flow diagrams, logs, impact assessments, access reviews, audit findings, and independent raters. The bounded cases demonstrate analytical use but do not establish operating effectiveness, causal impact, or comparative superiority. These limitations restrict generalisation beyond the selected Kazakhstan context.

## 5. Conclusion

The study finds that the central personal-data challenge in Kazakhstan's AI-enabled public services is not a complete regulatory vacuum but the absence of a consistently verifiable bridge between a legal duty and a service-level decision. KZ-AIPAG supplies that bridge by linking the use-case tier, individual hazards, controls, evidence, accountable roles, and lifecycle gates.

The scientific novelty is the Kazakhstan-specific compliance-by-evidence model and its separation of governance intensity from hazard prioritisation. The comparative, legal, and bounded case analyses support the internal coherence and practical relevance of the model, but they do not constitute independent validation. Its practical value lies in structuring impact assessment, procurement, approval, monitoring, audit, and redress around evidence that can be tested.

Future research should preregister and conduct the proposed expert study, pilot the framework in a defined public service, test the scoring anchors and gate thresholds, compare outcomes across language groups, and measure whether the evidence chain reduces consequential errors and the time required to restore affected rights.

## Data Availability Statement

Data availability. The study uses public legal instruments, official service descriptions, standards, and publications cited in the reference list. No personal data or confidential operational datasets were processed. The coding schema is reproduced in Table 2; no separate research dataset was created.

## Use of Artificial Intelligence

Use of artificial intelligence. Generative AI tools were used to support language translation, editorial restructuring, bibliographic cross-checking, and document formatting. The author remains responsible for source verification, legal interpretation, risk judgements, and the final text. No claim of expert-panel validation was generated or inferred by AI.

## References

**[1]** Republic of Kazakhstan. (2026). Konstitutsiya Respubliki Kazakhstan [Constitution of the Republic of Kazakhstan] (adopted by national referendum on 15 March 2026). https://old.adilet.zan.kz/rus/docs/K2600000000 (In Russian).
**[2]** Republic of Kazakhstan. (2013, as amended 2025). Zakon Respubliki Kazakhstan No. 94-V 'O personal'nykh dannykh i ikh zashchite' [Law of the Republic of Kazakhstan No. 94-V on personal data and their protection]. https://old.adilet.zan.kz/rus/docs/Z1300000094 (In Russian).
**[3]** Republic of Kazakhstan. (2025). Zakon Respubliki Kazakhstan No. 230-VIII 'Ob iskusstvennom intellekte' [Law of the Republic of Kazakhstan No. 230-VIII on artificial intelligence]. https://old.adilet.zan.kz/rus/docs/Z2500000230 (In Russian).
**[4]** Republic of Kazakhstan. (2026). Tsifrovoy kodeks Respubliki Kazakhstan No. 255-VIII [Digital Code of the Republic of Kazakhstan No. 255-VIII]. https://old.adilet.zan.kz/rus/docs/K2600000255 (In Russian).
**[5]** Government of the Republic of Kazakhstan. (2024). Postanovlenie No. 925 'Ob utverzhdenii trebovaniy po upravleniyu dannymi' [Resolution No. 925 on approval of data management requirements]. https://adilet.zan.kz/rus/docs/P2400000925 (In Russian).
**[6]** OECD. (2024). AI, data governance and privacy (OECD Artificial Intelligence Papers No. 22). OECD Publishing. https://doi.org/10.1787/2476b1a4-en
**[7]** OECD & UNESCO. (2024). G7 toolkit for artificial intelligence in the public sector. OECD Publishing. https://doi.org/10.1787/421c1244-en
**[8]** Tabassi, E. (2023). Artificial intelligence risk management framework (AI RMF 1.0) (NIST AI 100-1). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.AI.100-1
**[9]** National Institute of Standards and Technology. (2024). Artificial intelligence risk management framework: Generative artificial intelligence profile (NIST AI 600-1). https://doi.org/10.6028/NIST.AI.600-1
**[10]** Garfinkel, S., Near, J., Dajani, A., Singer, P., & Guttman, B. (2023). De-identifying government datasets: Techniques and governance (NIST SP 800-188). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-188
**[11]** International Organization for Standardization. (2023). ISO/IEC 42001:2023 Information technology – Artificial intelligence – Management system.
**[12]** International Organization for Standardization. (2023). ISO/IEC 23894:2023 Information technology – Artificial intelligence – Guidance on risk management.
**[13]** International Organization for Standardization. (2025). ISO/IEC 42005:2025 Information technology – Artificial intelligence – AI system impact assessment.
**[14]** International Organization for Standardization. (2025). ISO/IEC 27701:2025 Information security, cybersecurity and privacy protection – Privacy information management systems – Requirements and guidance.
**[15]** International Organization for Standardization. (2022). ISO/IEC 27001:2022 Information security, cybersecurity and privacy protection – Information security management systems – Requirements.
**[16]** Mäntymäki, M., Minkkinen, M., Birkstedt, T., & Viljanen, M. (2022). Defining organizational AI governance. AI and Ethics, 2(4), 603-609. https://doi.org/10.1007/s43681-022-00143-x

**[17]** Batool, A., Zowghi, D., & Bano, M. (2025). AI governance: A systematic literature review. AI and Ethics, 5(3), 3265-3279. https://doi.org/10.1007/s43681-024-00653-w
**[18]** Korobenko, D., Nikiforova, A., & Sharma, R. (2024). Towards a privacy and security-aware framework for ethical AI: Guiding the development and assessment of AI systems. In Proceedings of the 25th Annual International Conference on Digital Government Research (pp. 740-753). ACM. https://doi.org/10.1145/3657054.3657141
**[19]** Willems, J., Schmid, M. J., Vanderelst, D., Vogel, D., & Ebinger, F. (2023). AI-driven public services and the privacy paradox: Do citizens really care about their privacy? Public Management Review, 25(11), 2116-2134. https://doi.org/10.1080/14719037.2022.2063934
**[20]** Langer, C. (2024). Decision-making power and responsibility in an automated administration. Discover Artificial Intelligence, 4, Article 59. https://doi.org/10.1007/s44163-024-00152-1
**[21]** Papyshev, G. (2025). Governing AI through interaction: Situated actions as an informal mechanism for AI regulation. AI and Ethics, 5(2), 1109-1120. https://doi.org/10.1007/s43681-024-00446-1
**[22]** Dankloff, M., Skoric, V., Sileno, G., Ghebreab, S., van Ossenbruggen, J., & Beauxis-Aussalet, E. (2025). Analysing and organising human communications for AI fairness assessment. AI & Society, 40(4), 2347-2367. https://doi.org/10.1007/s00146-024-01974-4
**[23]** Akhmetova, S. B., Kassymzhanova, A. A., & Ibrayeva, A. S. (2025). Modern development of artificial intelligence technologies and problems of legal regulation of profiling and targeted advertising in Kazakhstan. Bulletin of L. N. Gumilyov Eurasian National University. Law Series, 150(1), 77-97. https://doi.org/10.32523/2616-6844-2025-150-1-77-97
**[24]** Treasury Board of Canada Secretariat. (n.d.). Directive on automated decision-making. https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32592
**[25]** European Parliament & Council of the European Union. (2024). Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence. Official Journal of the European Union. https://eur-lex.europa.eu/eli/reg/2024/1689/oj
**[26]** European Parliament & Council of the European Union. (2016). Regulation (EU) 2016/679 (General Data Protection Regulation). Official Journal of the European Union. https://eur-lex.europa.eu/eli/reg/2016/679/oj
**[27]** Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., Smith-Loud, J., Theron, D., & Barnes, P. (2020). Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing. In Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (pp. 33-44). ACM. https://doi.org/10.1145/3351095.3372873
**[28]** Alon-Barkat, S., & Busuioc, M. (2023). Human-AI interactions in public sector decision making: Automation bias and selective adherence to algorithmic advice. Journal of Public Administration Research and Theory, 33(1), 153-169. https://doi.org/10.1093/jopart/muac007
**[29]** Zuiderwijk, A., Chen, Y.-C., & Salem, F. (2021). Implications of the use of artificial intelligence in public governance: A systematic literature review and a research agenda. Government Information Quarterly, 38(3), 101577. https://doi.org/10.1016/j.giq.2021.101577
**[30]** Fatima, S., Desouza, K. C., Buck, C., & Fielt, E. (2022). Public AI canvas for AI-enabled public value: A design science approach. Government Information Quarterly, 39(4), 101722. https://doi.org/10.1016/j.giq.2022.101722
**[31]** van Noordt, C., & Misuraca, G. (2022). Artificial intelligence for the public sector: Results of landscaping the use of AI in government across the European Union. Government Information Quarterly, 39(3), 101714. https://doi.org/10.1016/j.giq.2022.101714
**[32]** de Bruijn, H., Warnier, M., & Janssen, M. (2022). The perils and pitfalls of explainable AI: Strategies for explaining algorithmic decision-making. Government Information Quarterly, 39(2), 101666. https://doi.org/10.1016/j.giq.2021.101666
**[33]** Delfos, J., Zuiderwijk, A. M. G., van Cranenburgh, S., Chorus, C. G., & Dobbe, R. I. J. (2024). Integral system safety for machine learning in the public sector: An empirical account. Government Information Quarterly, 41(3), 101963. https://doi.org/10.1016/j.giq.2024.101963
**[34]** Busuioc, M. (2021). Accountable artificial intelligence: Holding algorithms to account. Public Administration Review, 81(5), 825-836. https://doi.org/10.1111/puar.13293
**[35]** Wirtz, B. W., Weyerer, J. C., & Geyer, C. (2019). Artificial intelligence and the public sector – Applications and challenges. International Journal of Public Administration, 42(7), 596-615. https://doi.org/10.1080/01900692.2018.1498103
**[36]** Janssen, M., Brous, P., Estevez, E., Barbosa, L. S., & Janowski, T. (2020). Data governance: Organizing data for trustworthy artificial intelligence. Government Information Quarterly, 37(3), 101493. https://doi.org/10.1016/j.giq.2020.101493
**[37]** International Telecommunication Union. (n.d.). Smart Data Ukimet: Kazakhstan. https://www.itu.int/en/ITU-D/Regional-Presence/CIS/Pages/GovStack/SDU.aspx
**[38]** Ministerstvo iskusstvennogo intellekta i tsifrovogo razvitiya Respubliki Kazakhstan. (2026). Itogi 2025 g.: kazakhstantsy vse chashche poluchayut gosuslugi cherez smartfon [2025 results: Kazakhstanis increasingly access

public services via smartphones]. https://www.gov.kz/memleket/entities/maidd/press/news/details/1134705?lang=ru (In Russian).