On approval of the Rules for the collection, processing, storage, transfer of electronic information resources for the implementation of data analytics in order to implement functions by state bodies

Invalidated Unofficial translation

Order of the Minister of Digital Development, Innovation and Aerospace Industry of the Republic of Kazakhstan dated October 21, 2020 No. 394 / НҚ. Registered with the Ministry of Justice of the Republic of Kazakhstan on October 23, 2020 No. 21495. Abolished by the Order of the Minister of Digital Development, Innovation and Aerospace Industry of the Republic of Kazakhstan dated March 20, 2023 No. 99/NK

Unofficial translation

Footnote. Abolished by the Order of the Minister of Digital Development, Innovation and Aerospace Industry of the Republic of Kazakhstan dated March 20, 2023 No. 99/NK (effective from 01.04.2023).

In compliance with subparagraph 63-3) of article 7 of the Law of the Republic of Kazakhstan dated November 24, 2015 "On informatization", **ORDER**:

1. To approve the attached Rules for the collection, processing, storage, transfer of electronic information resources for data analytics in order to implement the functions of state bodies.

2. The Department of Digitalization of the Ministry of Digital Development, Innovation and Aerospace Industry of the Republic of Kazakhstan, in accordance with the established legislative procedure, shall ensure:

1) state registration of this order with the Ministry of Justice of the Republic of Kazakhstan;

2) posting this order on the Internet resource of the Ministry of Digital Development, Innovation and Aerospace Industry of the Republic of Kazakhstan after its official publication;

3) within ten working days after the state registration of this order with the Ministry of Justice of the Republic of Kazakhstan, submission to the Legal Department of the Ministry of Digital Development, Innovation and Aerospace Industry of the Republic of Kazakhstan information on the implementation of the measures provided for in subparagraphs 1) and 2) of this paragraph.

4. Control over the implementation of this order shall be entrusted to the supervising Vice-Minister of Digital Development, Innovation and Aerospace Industry of the Republic of Kazakhstan.

5. This order shall be enforced upon the expiration of ten calendar days after the day of its first official publication.

The Minister of Digital Development,
Innovation and Aerospace Industry
of the Republic of Kazakhstan B. Mussin

Rules for the collection, processing, storage, transfer of electronic information resources for

the implementation of data analytics in order to implement the functions of state bodies

Chapter 1. General provisions

1. These Rules for the collection, processing, storage, transfer of electronic information resources for the implementation of data analytics in order to implement the functions of state bodies (hereinafter referred to as the Rules), have been developed in accordance with subparagraph 63-3) of Article 7 of the Law of the Republic of Kazakhstan dated November 24, 2015 "On Informatization" (hereinafter referred to as the Law) and determine the procedure for the collection, processing, storage, transfer of electronic information resources for the implementation of data analytics in order to implement the functions of state bodies.

2. These Rules do not apply to:

1) relations arising from the implementation by the National Bank of the Republic of Kazakhstan and organizations that are part of its structure, work on the collection, processing, storage, transfer of electronic information resources for the implementation of data analytics in order to implement the functions of state bodies;

2) electronic information resources containing information constituting state secrets and secrets protected by law;

3) electronic information resources containing information about specific phenomena, processes and persons making up the data of the investigation and inquiry.

4) confidential information of limited distribution.

3. The following basic concepts and abbreviations are used in these Rules:

1) the authorized body in the field of informatization (hereinafter referred to as the authorized body) - the central executive body that carries out management and cross-sectoral coordination in the field of informatization and "electronic government";

2) information system (hereinafter - IS) - an organizationally ordered set of information and communication technologies, service personnel and technical

documentation that implement certain technological actions through information interaction and are designed to solve specific functional tasks;

3) information and communication service (hereinafter referred to as the IR service) - a service or a set of services for property lease (lease) and (or) placement of computing resources, provision of software, software products, service software products and hardware for use, including services links through which the functioning of these services is ensured;

4) extraction, transformation and loading (hereinafter - ETL) - the process of extracting data from the database of the information system of state bodies, their transformation and loading into a single data warehouse;

5) a unified data warehouse (hereinafter referred to as UDW) - a repository of large amounts of data, which are stored both in an unstructured and structured form, with specialized tools for accessing data;

6) depersonalization of personal data - actions, as a result of which the determination of the ownership of personal data by the subject of personal data is impossible;

7) database (hereinafter referred to as DB) is an organized structure designed for storing, changing and processing data;

8) data provider - legal entities entrusted with the obligation to provide data for analytics in accordance with the Law;

9) data analytics - the process of data processing in order to obtain information and conclusions for decision-making;

10) database of the information system of state bodies (hereinafter - DB IS SB) - a set of data from the information system of state bodies;

11) task - a situation requiring analysis and management decisions to optimize the activities of state bodies, as well as in order to implement strategic plans, state programs and development plans;

12) request - a document that contains information regarding the description of the statement of the task being implemented, the calculation methodology, information about the data used from the source systems, the frequency of data updating, the names of the graphs and their description, the definition of the expected result;

13) customer of analytics - a central executive body, a state body directly subordinate and accountable to the President of the Republic of Kazakhstan, a local executive body that needs processed data on a task in order to implement state functions;

14) historical data - data contained in the database since its creation;

15) interaction regulations - a document that establishes the rules, procedure and basic procedures for interaction between the operator and the data owner in terms of determining the responsibility of the participants in ensuring interaction, the list of

information objects, the schedule and methods of organizing interaction related to the processes of receiving and transmitting information;

16) electronic information resources - information provided in electronic digital form and contained on an electronic medium, an Internet resource and (or) in an information system;

17) operator of the information and communication infrastructure of "electronic government" (hereinafter referred to as the operator) is a legal entity determined by the Government of the Republic of Kazakhstan, which is entrusted with ensuring the functioning of the information and communication infrastructure of "electronic government" assigned to it;

18) ER-diagram - a diagram showing the structure of a database, indicating relationships between various entities (database tables);

19) information and analytical system "Smart Data Ukimet" - an informatization object located on the information and communication platform of "electronic government" and designed to form a single data space for the purpose of providing analytical information on the activities of the Government of the Republic of Kazakhstan.

4. When processing, storing, transferring electronic information resources for the implementation of data analytics in order to implement functions, state bodies take into account uniform requirements in the field of information and communication technologies and information security.

Chapter 2. The procedure for collecting electronic information resources

Paragraph 1. Initiating a request

5. In the event of situations requiring analysis in order to make management decisions to optimize the activities of state bodies, as well as in order to implement strategic plans, government programs and development plans, the analytics customer sends a request for data analytics to the authorized body.

6. The authorized body, within a period of not more than 5 (five) working days, considers the received request from the analyst's customer and, if the decision is positive, gives an order for execution to the operator.

In the event of a negative decision, the authorized body sends the analytics customer a reasoned justification about the impossibility of conducting analytics.

7. Reasons for refusal are:

1) lack of digitized data for analytics;

2) lack of social / economic effects;

3) the need for analytics of confidential data, information constituting a state secret;

3) the lack of a calculation methodology for conducting analytics;

4) lack of logical architecture of the database (description of fields, tables and links in the database) required for analytics.

8. The operator, within 10 (ten) business days from the date of receipt of the order, determines the availability of data in UDW, which is necessary for the implementation of data analytics as requested. In the absence of the necessary data for the implementation of analytics, the operator sends a request to the authorized body to connect DB IS SB to UDW.

9. In the absence of DB IS SB in UDW, the authorized body, together with the operator, within a period of not more than 10 (ten) working days, coordinates with the data provider information regarding the data available in the IS, the physical and logical structure of DB IS SB, the connection method, the list of transmitted and anonymized data, deadlines.

Paragraph 2. Development and approval of interaction regulations

10. The operator, within 10 (ten) working days, requests from the data provider a questionnaire to collect general information about the software and telecommunications environment of the state body in accordance with Appendix 1 to these Rules and technical documentation (ER-diagram, logical and physical data model , data architecture) ...

11. The operator, within 20 (twenty) working days, develops the rules for interaction of the information-analytical system "Smart Data Ukimet" with IS SB together with the data provider. To develop the interaction regulations, the data provider transfers to the operator technical information in accordance with the requirements for the description of data sources in accordance with Appendix 2 to these Rules.

12. The operator within 7 (seven) working days agrees, approves and sends to the data provider the rules of interaction.

13. The data provider within 5 (five) working days agrees and approves the interaction regulations.

14. If it is necessary to make changes to the interaction regulations, the data provider sends an official letter to the operator justifying the necessary changes.

15. The operator within 10 (ten) working days agrees, approves and sends to the data provider the rules of interaction with the changes made.

16. The data provider within 3 (three) working days agrees and approves the interaction regulations with the changes made.

Paragraph 3. Anonymization of data

17. In accordance with article 17 of the Law of the Republic of Kazakhstan dated May 21, 2013 "On personal data and their protection", when transferring data, their mandatory depersonalization takes place in the manner determined by the Rules for the collection and processing of personal data.

18. In order to depersonalize data, instructions are used to depersonalize personal data and a one-way hashing function without the possibility of reverse recovery, which ensures the legality, integrity and safety of data quality.

19. Data depersonalization is performed by the operator if it is impossible to carry out the data depersonalization procedure independently by the data provider. To depersonalize data, the operator takes the necessary measures in accordance with the interaction regulations.

20. Reconciliation, depersonalization and transfer of historical personal data occurs as follows:

1) the operator, together with the data provider, determines and agrees the list of data to be depersonalized;

2) the operator provides the data provider with the developed algorithms (sequence of actions) for one-way hashing without the possibility of reverse recovery, by database types and instructions for anonymizing personal data;

3) the data provider performs anonymization and provides the operator with historical data;

4) if the data provider is unable to depersonalize the data, the operator sets up the depersonalization procedure through the ETL tool. At the same time, the data is stored in the UDW in an impersonal form;

5) the operator checks the correctness of historical data, the correctness of data depersonalization and loads them into the UDW.

Paragraph 4. Transfer, download of historical data and setup of automatic routine data update

21. The data provider transfers historical data to the operator on a physical medium or in another agreed way. The operator, after receiving historical data from the data provider, through the ETL tool performs work on connecting, converting formats and encodings, identifying and removing errors and inconsistencies in the data in order to improve their quality.

22. To unload data from the database, the data provider configures network access in accordance with the interaction rules.

23. The data provider creates a user account in the database for reading certain data , opens access on the server for the servers of the Smart Data Ukimet Information and Analytical System on certain ports.

24. The operator examines the database structure and the transmitted historical data from the data provider, forms the logic for calculating updated data, creates ETL processes for their depersonalization and loading, and tests the scheduled data loading.

Chapter 3. Procedure for processing electronic information resources

25. The operator agrees with the authorized body the terms and volume of the provision of data received from various DB IS SB and combined with each other, in accordance with the request sent by the customer of the analytics to the authorized body.

26. The Operator, within 45 (forty-five) working days, provides data obtained from various DB IS SB and combined with each other, according to the request sent by the analytics customer to the authorized body for subsequent analytics.

Chapter 4. The procedure for storing electronic information resources

27. Collected data (detailed data, processed data, results of analytical solutions) are subject to storage and updating in the operator's UDW.

28. Updating of data is carried out by the operator in automatic mode in accordance with the interaction regulations.

Chapter 5. The procedure for the transfer and implementation of analytics of electronic information resources

29. The operator provides the analytics customer with access to the prepared processed data to check for compliance with the requested data.

30. The customer of analytics within 5 (five) working days provides information to the operator about the results of the check. In case of negative results, including the absence and (or) insufficiency of the necessary data, the operator re-processes the data in accordance with paragraphs 25 and 26 of these Rules.

31. The positive results of checking the compliance of the requested data contribute to the conduct of data analytics, which is carried out by the customer of the analytics independently or by attracting a supplier, at the expense of his own funds.

Appendix 1 to the Rules for
collection, processing, storage,
transmission of electronic information
resources to implement
data analytics for
implementation of functions
government bodies

Questionnaire for collecting general information about the software and telecommunications environment of a state body

Introduction

The document is a questionnaire for collecting general information about the software and telecommunications environment for the purpose of preliminary examination of the possibility of connecting to the IC - service "Information and analytical system" Smart Data Ukimet ".

Contacts of the responsible persons of the customer

This section contains the contacts of the responsible persons who can be contacted for clarifying questions when analyzing the answers to questions in this Document.

№ п.п.	First and Last name of employee	Position	Contact information (tel., e - mail)	Responsibility area
1.				
2.				

Information about the government agency

Name of company	
Full name of the person in charge who filled out the questionnaire	
Responsible person	
Contact phone number of the responsible person	
E-mail address of the person in charge	
Internet site address	

Information about the information systems implemented and operated.

For each system, fill out the questionnaire:

Name, year of introduction?	
The database management system used (hereinafter referred to as the DBMS)	
Is the DBMS located within the perimeter of the unified transport environment of government agencies?	
Where is the DBMS physically located (data center of National Information Technologies JSC, commercial data center, own server rooms)?	
From what year is history being conducted?	
Approximate number of tables?	
Number of users?	
Does the System contain data representing state or commercial secrets?	
The amount of raw data (excluding system fields, indexes, partitions, etc.)	
The total size of the database occupied on disk	
How the backup works	
How often the data is updated, constantly / hourly / daily / monthly / yearly.	
What is the approximate increase in data per day / month / quarter / year?	

What types of data manipulation (DML operators) are embedded in the information system (IS):	
1) Data is only replenished (insert), that is, data is always inserted into the IS;	
2) The data is replenished (insert) and / or updated (update), that is, data is inserted and updated in the IS;	
3) The data is replenished (insert), updated (update) and deleted (delete), that is, all types of data manipulation are performed in the IS.	
Is there a minimum load time period for the DBMS (for example, from 02:00 to 05:00)?	
Who maintains the system?	
Are there specialists who are able to create data samples using the DBMS (SQL language, etc.)?	

Requirements for describing data sources

1. Data architecture:
- Conceptual data model - displays the subject area of the state body and the objects used in the work.
- A Logical Data Model is a detailed view of the data requirements and business rules that govern the quality of the data, usually with a focus on a specific context of use. The logical model expands the conceptual one by defining their attributes, descriptions and constraints for entities, specifies the composition of entities and the relationship between them.
- Physical data model - the physical data model implements detailed data requirements, taking into account the limitations of the technology used by the applications, performance requirements and modeling standards. At this stage, relational databases are designed taking into account the specific capabilities of the system. Description of data objects (data object definition) defines the composition and formats of business and technical fields, their purpose, links to check tables, describes the rules of validation (for example, mandatory filling). Provides data quality requirements management, data change management, integration and migration.
- The data life cycle diagram defines the stages of data change / addition / update in the steps of business processes, including data behavior. Allows you to define the life cycle of records, set the rules: who and under what events / circumstances changes the status of the record and in what sequence.

- A data integration diagram is a data architecture artifact that displays integration flows of data between information systems. The value of the artifact lies in the fact that with the help of it is possible to quickly identify the place of origin of the data object and the path of data distribution between systems.