

Technical Assignment

to provide consulting services for the analysis of existing data collection systems, the development of a new data management system and a consolidated supervisory data model, as well as the development of business and technical requirements for the **EDW** and **ADR** at the Central Bank of Uzbekistan

Tashkent – 2022

Contents

1. Terms, definitions and abbreviations	3
2. Introduction.....	4
3. Targets and goals.	4
4. Project components.	5
The project will consist of the following components:	5
4.1. Development of a detailed corporate data management structure.	5
4.2. Development of Integrated Supervisory Data Model.	6
4.3. Development of terms of reference for the implementation of EDW and ADR.....	7
5. Requirements imposed by the Client on the organization of the Project.	11
6. Requirements for the Contractor.....	11
7. Terms of service provision	12
8. Warranty period for the result of the services rendered	12

This Terms of Reference describes the objectives of the provision of services, the requirements imposed by the Central Bank of the Republic of Uzbekistan on the results and content of the Contractor's consulting services, as well as additional essential conditions for the provision of these services.

1. Terms, definitions and abbreviations

For all terms and definitions that do not have a legislative definition and are not defined by this Terms of Reference, the Client will provide written explanations at the request of the Contractor.

Term	Definition
Client	Central Bank of the Republic of Uzbekistan
Contractor	Organization providing services.
Information technology infrastructure of the Client (hereinafter referred to as IT infrastructure)	Organizational and technical association of software, computing and telecommunications facilities, links between them and operational personnel, ensuring the provision of information, computing and telecommunications resources, capabilities and services to employees (divisions) of the Client necessary for the implementation of his professional activities and the solution of relevant business problems.
Business glossary, glossary of terms	Glossary of specialized terms in any branch of knowledge with interpretation, comments and examples
Soft	Software
IT	Information Technology
EDW	Enterprise Data Warehouse / Data Lake и etc.
SupTech (Supervisory Technology)	Technologies used by regulators to improve the efficiency of control and supervision over the activities of financial market participants.
RegTech (Regulatory Technology)	Technologies used by financial institutions to improve the efficiency of compliance with the requirements of the regulator.
ADR (Automated Data Reporting)	Automated data reporting system, a solution that involves the automatic collection of data from various platforms and their integration into the software system for collecting and processing data.
BI (Business Intelligence)	Methods and tools for ensuring the translation of transactional information into a human-readable form suitable for business analysis, as well as means for mass work with such processed information.
Metadata	Data about data, descriptions of data. There are business metadata (glossary of terms, business rules), technical metadata (implemented data structures, BI reports), operational metadata (quantitative indicators of data, data quality level), organizational metadata (responsible, owners, producers, consumers of data).
Master data.	Data containing key information about customers, products, employees, technologies and materials.
DBMS	Database management system
RRI	Regulatory reference information.
ETL (Extract-Transform-Load)	Processes of data processing in the construction of EDW.
ERD (Entity Relationship Diagram)	Graphical representation that depicts relationships among people, objects, places, concepts or events within an information technology (IT) system.
GAP	GAP analysis is a technique that allows you to assess the gap between current and planned results.
Artificial Intelligence (AI)	Artificial intelligence is a field of computer science that enables intelligent computer programs to perform tasks such as problem solving, speech recognition, visual perception, decision making, and language translation.
Machine Learning	Machine learning is one of the subcategories of AI. The ability of computer systems to perform tasks based on many similar tasks and continuous output.
Artifact	Architectural Artifact - A detailed architectural work product that describes an architecture from a specific point of view that is important to various participants involved in strategic decision making and information systems implementation.

2. Introduction

The Client is initiating a program to introduce innovative regulatory and supervisory technologies (SupTech) that will improve the Client's supervisory capabilities and ensure effective compliance with regulated financial institutions. The Client plans to start the SupTech program by transforming its data collection, processing and data management capabilities.

This can only be achieved if there is a proper enterprise data architecture with all relevant factors that describe existing data, its structure and relationships. Since data is an important factor in the implementation of SupTech, data management is a key capability that must be mastered by the Client. By applying a well-established data management system to all of its data throughout its life cycle, the Client will be able to significantly increase the value extracted from the data, while reducing the associated costs and risks. Today, this project is unique in this region and successful implementation will be a good portfolio for the Contractor.

3. Targets and goals.

The main goals of this project are to ensure proper planning and preparation for the implementation of improved data management capabilities, by implementing a holistic approach to data management, implementing a modern data architecture and advanced technologies such as Big Data, advanced analytics, machine learning, etc., which will cover all the functionality necessary for the Client to effectively perform tasks related to the collection, processing, analysis and dissemination of data, in accordance with the strategic goals of the Client.

The main tasks of the Contractor within the framework of this Project will be to develop appropriate methodological, architectural and conceptual documents (details of services and results are given in clause 4 of this Terms of Reference), within the framework of which the Contractor must carry out:

- Examination of the Client's data collection system (for all types of channels), the degree of automation of the Client's data management processes necessary to build an effective and high-tech supervision system as part of the program for the implementation of SupTech solutions;
- Development of the Client's corporate data management structure in accordance with the best practices in the field of data management;
- Drawing up a roadmap describing the necessary data management activities to increase the degree of automation and improve the level of data quality, as well as the implementation of ADR within SupTech;
- Analyze the current regulatory framework that establishes reporting obligations to identify issues that need to be improved to support the supervisory data model;
- Preparation of detailed recommendations: on the implementation of the presentation for the Client of data sets in the most detailed, in which data are generated in information systems and databases of banks and carry out their life cycle; to achieve such a provision of the Client with data and tools that will replace the provision of current standard reports containing only aggregated data;
- Development of the Client's Integrated Supervisory Data Model;
- Analysis of the connection of all basic information systems of the Client in a single system of accumulation, verification and data analytics;
- Development of a model of a single information space in order to organize systematic work on corporate data management and data quality control;
- Development of Terms of Reference for the implementation of a data management solution,

including EDW and ADR, and the components of metadata management, data quality management and regulatory reference information management;

- Development of functional requirements for systems for automated construction of supervisory reporting and business analysis (BI).

4. Project components.

The project will consist of the following components:

4.1. Development of a detailed corporate data management structure.

The overall goal of this component is to define and implement a mature and powerful data management system for the Client based on the best international practices developed by the Data Management Association (DAMA).

Key goal of the component:

Data is a key factor for the main business processes and functions of the Client. By applying a well-established data management system for all data throughout the life cycle, the Client will be able to significantly increase the value extracted from the data, while reducing the associated costs and risks.

For this component, the Contractor will develop and provide recommendations on the appropriate organizational structure, skills, policies and procedures, tools and any other capabilities necessary for the Client to effectively manage its information assets.

The main activities that need to be performed to achieve the objectives of the component:

Under this component, the Contractor is expected to provide and perform the following key actions:

- Assessing the current state of the Client's data management practices to understand current and future data management needs;
- Defining the desired data management structure, including policies, procedures, processes, tools, organizational structures, roles, responsibilities and skills;
- Establish standards and guidelines for documentation and data architecture development. Based on industry best practices, the Contractor will use various architectural artifacts to properly document the data architecture, such as:
 - Catalogs: data taxonomy, data dictionary, data structures. Data catalogs should contain information about data categories and subcategories, their meaning, structure, and specific integrity rules.
 - Diagrams: graphical representation of data objects and their relationships (for example, ERD).
 - Matrices: used to define relationships between data objects and other types of objects, such as data and applications that store/process data; users/structural units and CRUD (Create, Read, Update, Delete)-relation to data.
- Conducting an inventory and documentation, description of the Client's high-level corporate data architecture;
- Gap assessment (GAP) between the current state and the desired state for data management at the Client;
- Development of a transformation plan (Roadmap) for the implementation of an improved data management system, for the organization and development of internal processes, procedures, organizational structures, technologies, etc. for data management;

- Training of key individuals who will be involved in data management processes (data owners, data custodians, data architects, data analysts);

Main expected results:

Under this component, the Contractor is expected to provide at least the following set of results:

- Report on the current state of data management;
- Consolidated and structured vision of the Client in relation to data management and big data architecture (To Be - the desired state);
- Detailed data management structure;
- Data Management Transformation Plan;
- Key processes and procedures of the data management system, key roles and responsibilities;
- Curriculum and trainings for key persons on data management;
- High-level enterprise data architecture artifacts;
- Data model artifacts specific to particular areas in the field of supervision (prudential, statistical, financial and permissive regulatory reporting).

4.2. Development of Integrated Supervisory Data Model.

This component is intended to focus on the assessment of the Client's current supervisory data needs and the development of integrated supervisory data model using the application defined in Component 4.1 - Data Management Framework.

For the purposes of defining a data model, it is proposed to create a single unique data dictionary using one common dictionary to support all reporting needs. In line with the goals of integrating data from different departments, the option of using a single unique dictionary is expected to be more efficient than maintaining two or more dictionaries and the associated efforts to duplicate or collate definitions and harmonize the rules of different dictionaries and methodologies.

Key goal of the component:

This project is intended to initiate the development of supervisory data model using the new data management framework to be developed by those identified in component 4.1.

For this component, the specific objectives are to develop a supervisory data model and develop a regulatory reporting model - a unified, normalized, universal data catalog and standardized integrated regulatory logic for the purposes of prudential, statistical, financial and permissive regulatory reporting (Integrated Reporting Structure and Vocabulary).

The main activities that need to be performed to achieve the objectives of the component:

For this component, Contractor:

- Conduct an inventory of all submitted data with the Client and document the high-level “Current State” data architecture for the Client;
- Together with the Client, will review data sets collected from supervised organizations. They are currently template-based, which means that redundant data is requested and sent;
- Review the regulatory framework that establishes reporting obligations to identify issues that need to be improved to support the Integrated-Supervisory Data Model;
- Develop an Integrated Supervisory Data Model and a Regulatory Reporting Model. The revised

supervisory data model should contain only primary raw data related to supervised entities and requested only once (with some exceptions). The initial data will be additionally processed to extract the necessary information in accordance with the individual needs of various structural divisions of the Client;

- Develop a model of regulatory reporting, taking into account the possibility of using specialized software solutions for ADR.

Main expected results:

Under this component, the Contractor is expected to provide the following set of results:

- Report on the current inventory of submitted data by the Client, containing detailed recommendations for increasing the number of detailed data sets by replacing the current standard reports containing only aggregated data;
- Integrated model of the Client's supervisory data, including:
 - A wide unified taxonomy of the Client's data (a data catalog that defines all objects, their grouping, semantic and syntactic definitions, etc.).
 - Relationship diagrams (ERDs) that define relationships between data objects.
 - Validation rules for each defined data object.
- Recommendations on RegTech, in particular on setting requirements for the collection and processing of data from supervised organizations;
- Model of the test environment for approbation of methodologies for the formation of supervisory reporting;
- Other elements and artifacts as required coming from the data management structure in accordance with clause 4.1.

The work will be carried out mainly through on-site interviews, workshops with the Client's supervisory and technical staff, review of relevant detailed documentation. The Contractor will describe the methodology and tools used in the analysis and design phase and provide a sample of the results.

4.3. Development of terms of reference for the implementation of EDW and ADR.

As part of this project, it is also important to create the right context for defining business requirements for software solutions that play a key role in data management throughout the entire life cycle, from collection to processing, storage, use and archiving.

The key goal of this component is the development of terms of reference for the implementation of EDW and ADR, which will be the next key step in the implementation of the Client's strategic goals.

Key goal of the component:

As part of this component of the project, the Contractor will determine the applicable business and technological requirements for software solutions for the development of technical specifications for the implementation of EDW and ADR, which will be the next key step in the implementation of the Client's strategic goals.

The main activities that need to be performed to achieve the objectives of the component:

The main activities under this component should focus on defining the current and future requirements for EDW and ADR solutions.

An established data management framework, defined in accordance with Section 4.1., and a unified supervisory data model, defined in accordance with Section 4.2., should be key inputs to this activity.

Main expected results:

The main expected output for this component is the terms of reference for EDW and ADR, which should contain at least the following:

- Description of the general goals and objectives to be achieved within the framework of the implementation of EDW and ADR;
- Description of the Client's vision of the big data architecture to be achieved as part of the implementation of EDW and ADR;
- A description of the Client's expected information architecture to be supported by EDW and ADR, including business architecture, application layer architecture, data layer architecture, and technology architecture;
- Detailed functional, technical and technological requirements that each of the solutions must correspond;
- Implementation service requirements, including description of key activities and deliverables, identification of key eligibility criteria, etc;
- Requirements for maintaining the system after its implementation, including warranty service.

It is also expected that the Contractor will develop a set of recommendations on criteria for qualification of potential Participants and criteria for evaluating proposals for terms of reference for the implementation of EDW and ADR.

Additional requirements for expected results:

Among the more specific goals and expected results that will need to be taken into account and detailed as part of the development of the terms of reference, the following can be distinguished:

General business goals:

- Ensuring that large amounts of data are processed in a fast and efficient manner, and made available to supervisory departments as close to real time as possible and without any unreasonable delays, in order to improve the Client's reacting time;
- Consolidate data from multiple sources, clean and convert it into a modeled format suitable for storage and further use by supervisory departments;
- Providing flexible and timely automation of any complex data transformation scenarios so that the Client can quickly adapt to changing surveillance needs;
- The ability to supplement data during processing with additional attributes extracted from public registers in order to increase their usability and make them more valuable for supervisory decision-making;
- Availability of a log of changes and data transformations for the purposes of analysis and audit;
- View and access all data (both structured and unstructured) from a centralized and secure EDW based on clear permissions granted by data owners;
- Provide fast access to data without any performance or limitation due to the large size of the datasets;
- Enable virtual integration of any additional data from other surveillance systems to gain end-to-end visibility into existing data;
- The ability to establish secure ways of exchanging data between departments and other external parties without compromising data security;

- The ability to efficiently automate analytical tasks for the various data aggregations needed to assess risk factors within risk profiling methods and feed the output to real-time dashboards and other visualization tools, and allow supervisory departments to focus more on assessment tasks;
- Provide supervisory departments with access to tools that would enable them to perform advanced data analysis and forecasting;
 - Enable data exploration with simple search capabilities and data mining tools;
 - Use advanced AI/ML algorithms that can better support processing existing data and detecting anomalies or obtaining predictive information;
 - Use data visualization technology to analyze and visualize information in a practical and useful way for supervisor departments to extract useful information from massive datasets;
- The ability to easily create and customize intuitive analytic dashboards to suit the needs of each supervisory department, without special technical knowledge or without the involvement of an IT specialist;
- Improve the forecasting capabilities of supervisor units using powerful visualization techniques such as heat maps, histograms, etc.;
- Simplify and streamline data collection processes through greater automation, shorter data submission cycles, automated feedback loops, minimal need for manual operations;
- Increasing efficiency by eliminating repetitive, reproducible and routine tasks, standardizing procedures in processes and achieving a greater degree of consistency and quality of processes;
- Providing clear and unified instructions for supervised organizations. Providing common terminology with precise definitions that will be used in different structural departments of the Client;
- Ensuring the provision of quality data to the Client in order to effectively support supervisory goals;
- Providing a strong security environment for handling sensitive data in a controlled and secure manner.

In the part of the Enterprise Data Warehouse (EDW):

The target EDW Architecture should comply with the general accepted requirements for its construction, and a description of the EDW architecture should be developed, taking into account the principles of its division into layers, the approach to the data model and the ETL architecture:

- on data collection;
- on data processing and extraction;
- on providing access to data.

On data collection:

- EDW should provide the ability to download data from different sources (from external and internal information systems), has the ability to download data from formats (XML, JSON, Excel, CVS, etc.);
- EDW should provide the ability to track the status of the ETL process, detailing the identity and number of data items retrieved from the source systems involved in the ETL process;
- EDW must ensure the availability of control procedures and reports confirming the correctness of the uploaded data, compliance with the business logic and data model;
- Transformations of data from source systems to the level of the data model must be performed using transformation tables and parameterized procedures. It is not allowed to link elements of directories and

analytics directly in the program code of transformation procedures. EDW should, as far as possible, exclude redundant and complex data transformations while adapting the corresponding source systems to the reporting requirements;

- The system should be documented at all levels, including the structure of data storage, description of algorithms for integration with data sources, building data marts, etc.

On data processing and extraction;

- EDW should provide the ability to extract data as of a given date and/or for a certain period (data slice extraction);

- EDW should provide the ability to extract a portion of data that has been changed (including those marked as "deleted") or added since the last successful data extraction (incremental data extraction);

- EDW should provide the ability to notify users defined by the system administrator when certain events occur within the ETL process;

- EDW should be open to adjacent systems and should support the ability to export/import data to adjacent systems through software interfaces or data files of various formats (XML, CSV, JSON, XLS, TXT, etc.) and have detailed data export/import documentation.

On providing access to data.

- Each of the systems included in the EDW must provide remote connection of users (operators, administrators, developers, etc.);

- EDW should provide differentiation of powers and role-based access;

- EDW must comply with the required level of information security and protection against unauthorized access;

- Updates of EDW software and hardware should not be tied to updates of software/hardware releases of source systems;

- To be able to view and access all data (both structured and unstructured) from a centralized and secure EDW based on clear permissions granted by data owners;

- Be able to easily create and customize intuitive analytics dashboards to suit the needs of each supervisory department, without special technical knowledge or without the involvement of IT;

- Provide fast access to data without any performance limitations or limitations due to the large size of the datasets;

- Provide supervisory departments with access to tools that would enable them to perform advanced ad hoc data analysis;

- Use data visualization technology to analyze and visualize information in a practical and useful way for supervisory departments to extract useful information from massive datasets;

- The list of components on which the EDW will be built must be agreed with the Client and be part of the architecture description.

In terms of the Automated Reporting System (ADR).

- The Contractor will have to develop and agree with the Client on new principles of work and processes for providing data (as well as preparing a presentation for supervised organizations on upcoming changes);

- Develop detailed requirements for processes and procedures for providing data, including supported data formats, minimum validation requirements and other rules for checking the integrity and quality of data, etc.;

- Develop minimum portal requirements for supervised entities, as well as all functionalities to support planning and monitoring reporting obligations and provide reliable feedback;

- Develop minimum requirements regarding the non-functional aspects of the software solution (technological and technical part of the solution) based on all factors that may affect the performance, correctness and reliability of reporting processes;

- Develop minimum requirements for integration with the EWD solution, including a description of the interdependencies between these projects and how the Client will manage them;

- Develop a test environment for testing new forms of calculation of standards planned for implementation in the course of the Client's subsequent activities;

- Develop requirements for training IT users and supervisory users on the use of ADR (training on principles and best practices for building data management systems).

5. Requirements imposed by the Client on the organization of the Project.

- The Contractor must organize a project team to provide services, and also appoint a team leader who directly manages the team and is authorized to solve all emerging issues. Changes in the composition of the team are possible only by written agreement with the Client.
- Meetings of the Project Team from the Client and the Project Team from the Contractor to determine the status of the Project and solve operational issues should be held on a regular basis, at least once a week on the territory of the Client or via telecommunication facilities. At least 50% (fifty) of the Contractor's project team must work on the territory of the Client. The Contractor agrees the date and time of the meeting with the Client by phone or by e-mail specified in the contract.
- After each meeting with the representatives of the Client, the team leader on the part of the Contractor, no later than the next working day, must draw up a meeting protocol and agree with the head of the Project team on behalf of the Client via e-mail. The Client reviews the protocol within 5 (five) business days and sends a response to the Contractor. If there are comments to the protocol, the Contractor eliminates them no later than 3 (three) business days and sends them to the Client for re-approval.
- The team leader from the Contractor must prepare a weekly report that includes a list of tasks completed during the week, planned tasks, existing risks and problems, and the status of services.

6. Requirements for the Contractor

- The Contractor is obliged to prepare a general presentation, with the presentation of intermediate results for the components (clauses 4.1, 4.2, 4.3) of the Project for the Client and to discuss the project;

- Submission of intermediate results for the components (clauses 4.1, 4.2, 4.3) of the Project specified in clause 4 of this Terms of Reference is carried out by the Contractor after agreeing the documents developed by him with the Client;

- To agree on the intermediate results by components (clauses 4.1, 4.2, 4.3) of the Project, the Contractor sends them by e-mail to the Client. The Client, within 15 (fifteen) working days, reviews the received documents and sends the Contractor either a notice of approval of the documents, or comments on them. The Contractor within 15 (fifteen) working days eliminates the comments and sends the documents

for re-approval to the Client. The approval period is included in the total period for the provision of services;

- Not later than the next working day after the end of the general term for the provision of services, the Contractor forms and sends the results of the Project to the Client for approval using e-mail. The Client, using e-mail, approves the results of the Project no later than 2 (two) working days;

- The final versions of all documents are accepted according to the Transfer and Acceptance Certificate, drawn up in any form separately for each component. (clauses 4.1, 4.2, 4.3);

- Completion of the project is formalized by the certificate of completion work on the whole project.

7. Terms of service provision

Beginning of the provision of services - no later than 20 (twenty) calendar days from the date of conclusion of the Contract;

The recommended period for the end of the provision of services is no more than 120 (one hundred and twenty) working days from the date of commencement of the provision of services.

8. Warranty period for the result of the services rendered

The warranty period is valid from the date of signing by the parties of the act of completed work and the subsequent implementation of the Terms of Reference, but not more than 18 (eighteen) months.

Guarantees apply to all constituent elements of the result of services provided by the Contractor under this Terms of Reference.

The contractor is responsible for defects found within the warranty period.

In case of detection of shortcomings in the documents submitted by the Contractor, the Contractor undertakes to eliminate them within 10 (ten) business days from the date of discovery or in other terms agreed by both parties, but not more than 30 (thirty) business days.