BUILDING A COMPREHENSIVE DATA GOVERNANCE FRAMEWORK: A STRATEGIC INITIATIVE BY THE TURKISH STATISTICAL SYSTEM

Prepared by Bilal Kurban (Corresponding Author, bilal.kurban@tuik.gov.tr), Gaye Mat Çelik, Serhan Kökhan, Ali Osman Bilgin, İsmail Akgül, Tuğba Değirmenci, Turkish Statistical Institute

Abstract

Turkish Statistical Institute (TurkStat) is prepared to actively influence the data governance environment in Türkiye. To achieve this goal, TurkStat recognizes the need to establish a robust framework for official statistics environment. This strategic endeavor marks the initial step towards solidifying TurkStat's role as a leader in data governance within Türkiye. Turkstat has begun developing a robust data governance framework for the Turkish Statistical System to enhance the integrity, accessibility and utility of its data assets. The framework has key components that include policy, guidelines and legislation, data management, cultivation of a data-centric culture, technological infrastructure and fostering strategic partnerships.

Fundamentally, the framework emphasizes the formulation of clear policies, guidelines and regulations aligned with a corporate vision and strategic objectives. To ensure responsibility throughout the data lifecycle, organizational models are defined to specify data ownership, roles and responsibilities. Data security and privacy protocols are carefully outlined to safeguard against breaches and uphold confidentiality. Ethical considerations emphasize transparency, auditability and adherence to strict data quality standards. Continuous monitoring, evaluation and improvement mechanisms are integrated to refine processes iteratively.

The entire process of creating data, from identifying needs to evaluating it, is covered by meticulously planned data and metadata management procedures. Emphasis is placed on robust metadata management, standardization and effective database management techniques.

Cultivating a data-driven culture is prioritized through initiatives promoting data literacy, governance awareness and technological proficiency among stakeholders. The framework also addresses technological infrastructure requirements, including the deployment of advanced data analysis and management tools, corporate information systems and open-source solutions.

Collaboration and partnerships form a cornerstone of the framework, fostering national and international stakeholder cooperation, harmonization and integration with diverse data ecosystems.

This paper will summarize the comprehensive approach undertaken by the Turkish Statistical System to establish a resilient data governance framework, supported by statistical legislation, principles of integrity, transparency, collaboration and linguistic accessibility. A dedicated platform which is being developed to facilitate the sharing of the framework with partners and to observe gaps (such as legal issues around AI) in the data governance framework will be presented. Furthermore, a comprehensive data governance glossary in Turkish and English which is created to ensure clarity and consistency in terminology across all stakeholders will be shown.

I.   INTRODUCTION

1.   The Turkish Statistical Institute (TurkStat) has been the producer and coordinator of Türkiye's official statistics since 1926. Taking into account the needs and priorities of national and international users, TurkStat works within the framework of its mission to produce and make available high quality, up-to-date, reliable, consistent, impartial and internationally standardized statistics. Official statistics are of vital importance in terms of providing information on Türkiye's socioeconomic situation, supporting decision-making processes and policy-making. Recognizing this importance, TurkStat has developed different methodological frameworks to ensure that data are collected, processed and analysed in accordance with international standards through an effective data infrastructure equipped with modern data management systems and tools. This infrastructure enables TurkStat to efficiently manage and analyse large amounts of data. TurkStat has strict policies and procedures in place to safeguard against breaches in data security and confidentiality and to protect the rights of data subjects. Efforts are underway to improve data sharing policies and tools to make official statistics publicly available and accessible.

2.   In Türkiye, the programmatic period started in 2007, aiming to disseminate TurkStat's institutional experience in data production on a national basis. In this context, the Official Statistics Program (OSP), which is prepared for five-year periods based on the Turkish Statistics Law, is coordinated by TurkStat in order to determine the basic principles and standards for the production and publication of official statistics, and to ensure the production of up-to-date, reliable, timely, transparent and impartial data in areas of national and international need. The Statistical Council, as defined in the same Law, convenes annually under the chairmanship of the President of TurkStat, with the participation of senior executives of all Ministries and related institutions and representatives of non-governmental organizations.

3.   In today's world, where data is becoming more important every day, but has turned into institutional silos and piles, data governance is of great importance in order to take full advantage of the ongoing "data revolution". TurkStat has started to work on transforming the Turkish Statistical System (TSS) into a data governance ecosystem to manage it in a more organized, internationally standardized and high value-added manner. As a matter of fact, in the 2024-2028 Strategic Plan, TurkStat's vision is defined as "to be a pioneering institution in data governance". Within the framework of this vision, TurkStat's century-long history and nearly 20 years of experience in the official statistics program, as well as the activities of the Statistical Council and OSP working groups, are considered to expand towards a common understanding of national data governance in Türkiye. In order to contribute to this expansion, and recognizing the importance of developing a national data governance policy, TurkStat has started to develop a robust data governance framework for the TSS with the aim of becoming a leading institution in the field of data governance.

II.   DATA GOVERNANCE FRAMEWORK

4.   With the increasing value of data today, it is of utmost importance to manage this data in an accurate, secure and efficient manner. With this awareness, TurkStat attaches great importance to strengthening national data governance and bringing the TSS in line with modern standards.

5.   Today, data management and data governance are considered as important management perspectives. While data management covers activities related to the collection, storage, protection and use of data, data governance sets out general principles and rules on how data should be used. These principles include issues such as data security, data privacy and ethics.

6.   TurkStat started its efforts to establish a data governance framework by creating an inclusive data governance and framework definitions in the light of the basic principles of the organization. These definitions are

7.   **Data governance:** It is the discipline of ensuring ownership, responsibility, access, metadata, quality, privacy, security, data management and culture within the framework of legislation, policies and regulations throughout the life cycle of data, and establishing a harmonized infrastructure with national and international cooperation in this process.

8.   **Data governance framework:** A set of roles, principles, standards and disciplines that covers the entire life cycle of data with the participation of all stakeholders, provides high data quality, trust and efficiency with an up-to-date technological infrastructure, ensures data privacy and security, promotes data integration, develops data culture and literacy, and enables national and international cooperation.

9.   It is of great importance that the experience covering every process of data management while collecting, managing, distributing and using data evolves into national data governance. Recognizing the importance of developing a national data governance policy with the goal of becoming a leading organization in the field of data governance, TurkStat is developing a robust data governance framework for the TSS.

10.   Basically, this framework includes:

- Highlight policies, guidelines and regulatory needs aligned with the institutional vision and strategic objectives, and define organizational models for determining data ownership, roles and responsibilities throughout the data lifecycle.

- It is emphasized that data security and privacy protocols should be based on ethical considerations such as transparency, auditability and data quality standards.

- In order to ensure sustainability, monitoring, evaluation and improvement mechanisms are integrated with the philosophy of continuous improvement.

- All data creation processes, from the identification of needs to their evaluation, are meticulously planned according to effective metadata - database management principles and standardizations.

- Priority is given to fostering a data-driven culture that encourages data literacy, governance awareness and technological competence among stakeholders.

- The framework also addresses technological infrastructure requirements, including the deployment of advanced data analysis and management tools, enterprise information systems and open source solutions.

- Collaboration and partnerships encourage national and international stakeholder cooperation, harmonization and integration with various data ecosystems.

A.   Data Governance Framework Components

11.   A data governance framework is the backbone of any organization's approach to managing and using its data assets effectively and responsibly. In the TSS, this framework is designed to ensure the integrity, accessibility and utility of data assets, thereby enhancing the organization's ability to fulfil its mission and serve its stakeholders.

12.   In order to identify gaps in the current functioning of the data governance framework for the TSS, first, national and international best practices were reviewed to identify the ideal data governance

components. This section summarizes the key components of the data governance framework and their importance in achieving these objectives. The components of this framework are shown in the diagram below.

Components of the data governance framework

1. Policy, Guidelines and Regulations	2. Data Management	3. Data Culture	4. Technology and Infrastructure	5. Interoperability and Communication
1.1. Vision and Mission	**2.1. Management of Data**	**3.1. Data Literacy**	**4.1. Hardware**	**5.1. User Engagement**
1.1.1. Strategic Plan and Goals	2.1.1. Data Storage Management	3.1.1 Digital Competence	**4.2. Software**	**5.2. Interoperability and Reusability**
1.2. Roles and Responsibilities	2.1.2. Security Management	3.1.2 Statistical Literacy	4.2.1. Open Source Structures	**5.3. Partnerships**
1.2.1. Boards	**2.2. Metadata Management**	**3.2. Data Governance Awareness**	4.2.2. Licensed	5.3.1. National
1.2.2. Data Authority	2.2.1. Data and Metadata Catalogue Management		4.2.3. In-house Solutions	5.3.2. International
1.2.3. Institutions	2.2.2. Code System and Classifications		**4.3. Networks and Connections**	
1.3. Data Ethics	2.2.3. Data Standardization			
1.3.1. Transparency	**2.3. Data Quality Management**			
1.3.2. Audit	**2.4. Statistical Business Process Management**			
1.4. Data Security and Privacy	2.4.1. Specify Needs			
1.4.1. Privacy and Security Policies and Procedures	2.4.2. Design			
1.5. Data Access and Sharing	2.4.3. Build			
1.5.1. National Data Access and Sharing	2.4.4. Collect			
1.5.2. International Data Access and Sharing	2.4.5. Process			
1.5.3. microdata	2.4.6. Analyse			

1.6. Quality	2.4.7. Disseminate
1.6.1. Standards and Metrics	2.4.8. Evaluate
1.7. Monitoring, Evaluation and Continuous Improvement	
1.8. Risk management	
1.9. Finance and Budget	

1. **Policy, Guidelines and Regulations**: This component encompasses the establishment of overarching policies, guidelines and regulations that govern the management, use and sharing of data within the TSS. These policies provide a framework for decision-making, ensuring consistency, accountability and alignment with organizational objectives. In addition to guiding stakeholders in their roles and responsibilities, clear policies and guidelines help to define the organization's vision, mission and strategic goals. Regulations also promote transparency, integrity and trust in the data governance process by ensuring compliance with legal requirements and ethical standards.

   1.1. **Vision and Mission:** The foundation of the data governance framework lies in the organization's vision and mission, which provide overarching direction and purpose.

   1.1.1. **Strategic Plan and Goals:** Strategic planning sets out the organization's long-term objectives and defines specific goals to be achieved within designated timeframes.

   1.2. **Roles and Responsibilities:** Clear delineation of roles and responsibilities ensures accountability and effective decision-making throughout the data lifecycle.

   1.2.1. **Boards:** Governance boards oversee the implementation of policies and ensure alignment with organizational objectives.

   1.2.2. **Data Authority:** A designated data authority is responsible for enforcing policies, resolving disputes and overseeing compliance.

   1.2.3. **Institutions:** Various institutions within the organization play specific roles in data governance that contribute to its overall effectiveness.

   1.3. **Data Ethics:** Ethical considerations guide the responsible use of data, emphasizing transparency, accountability and adherence to established standards.

   1.3.1. **Transparency:** Openness about data practices and decision-making processes fosters trust and promotes accountability.

   1.3.2. **Audit:** Regular audits ensure compliance with ethical guidelines and identify areas for improvement.

   1.4. **Data Security and Privacy:** Robust security measures safeguard confidentiality and privacy by protecting against unauthorized access, breaches and data misuse.

   1.4.1. **Privacy and Security Policies and Procedures:** Comprehensive policies and procedures govern data handling, storage and access to mitigate security risks.

   1.5. **Data Access and Sharing:** Facilitating data access and sharing promotes collaboration, innovation and informed decision-making while balancing privacy and confidentiality concerns.

1.5.1. National Data Access and Sharing: Frameworks for national data sharing enable stakeholders to access and utilize data for various purposes.

1.5.2. International Data Access and Sharing: Engaging in international data sharing initiatives enhances collaboration and enriches data resources.

1.5.3. Microdata: Access to microdata enables detailed analysis and research while maintaining confidentiality and privacy.

1.6. Quality: Ensuring data quality requires adhering to established standards and metrics in order to maintain trust and confidence in statistical outputs.

1.6.1. Standards and Metrics: Data quality is improved by defined standards and metrics, which facilitate consistent data collection, processing and analysis.

1.7. Monitoring, Evaluation and Continuous Improvement: Ongoing monitoring and evaluation mechanisms enable organizations to identify and address challenges, refine processes and improve performance iteratively.

1.8. Risk Management: Proactive risk management strategies safeguard against adverse impacts by mitigating potential threats to data integrity, security and privacy.

1.9. Finance and Budget: Adequate financial resources are essential for implementing and sustaining data governance initiatives, supporting infrastructure and capacity building.

2. **Data Management:** Data management involves the systematic management of data throughout its lifecycle, from acquisition and storage to analysis and dissemination. This component focuses on implementing best practices and procedures for managing data effectively, ensuring its quality, security and accessibility. Effective data management includes processes for data storage, security, metadata management, quality assurance and statistical business process management. By establishing robust data management practices, organizations can optimize the value of their data assets and support evidence-based decision-making**.**

2.1. Management of Data: Effective data management encompasses data storage, security and lifecycle management to ensure data availability, integrity and usability.

2.1.1. Data Storage Management: Proper storage infrastructure and practices optimize data accessibility, reliability and scalability.

2.1.2. Security Management: Implementing robust security measures safeguards against unauthorized access, breaches and data loss.

2.2. Metadata Management: Managing metadata supports informed decision-making and analysis by facilitating data discovery, understanding and usability.

2.2.1. Data and Metadata Catalogue Management: Centralized catalogue management enables efficient discovery and retrieval of data and associated metadata.

2.2.2. Code System and Classifications: Standardized code systems and classifications enhance data consistency, comparability and interoperability.

2.2.3. Data Standardization: Establishing data standards ensures consistency and compatibility across datasets, facilitating integration and analysis.

2.3. Data Quality Management: Implementing quality management practices enhances the reliability, accuracy, and relevance of statistical outputs, which also ensures data fitness for purpose.

2.4. Statistical Business Process Management: Defined business processes improves efficiency and consistency by streamlining data collection, processing, analysis and dissemination.

2.4.1. Specify Needs: Identifying user requirements and data needs informs the design and development of statistical processes and products.

2.4.2. Design: Designing data collection instruments, methodologies and protocols ensures data quality and relevance.

2.4.3. Build: Developing data collection systems and infrastructure supports efficient and effective data capture and storage.

2.4.4. Collect: Collecting data according to established protocols and standards ensures consistency and reliability.

2.4.5. Process: Processing collected data involves cleaning, validating and transforming raw data into usable information.

2.4.6. Analyse: Analysing data using appropriate methodologies and tools generates insights and supports evidence-based decision-making.

2.4.7. Disseminate: Disseminating statistical outputs and analysis products to stakeholders facilitates access to information and promotes transparency.

2.4.8. Evaluate: Evaluating data quality, relevance and impact informs continuous improvement efforts and enhances organizational effectiveness.

3. **Data Culture:** Data culture refers to the attitudes, beliefs and behaviours within an organization regarding the use and management of data. Developing a data-driven culture is essential for promoting data literacy, governance awareness and accountability among stakeholders. This component emphasizes the importance of building digital and statistical competence among staff, promoting a culture of data-driven decision-making, and raising awareness of data governance principles and practices. A strong data culture facilitates collaboration, innovation and continuous improvement in data management and use.

3.1. Data Literacy: Promoting data literacy improves stakeholders' ability to understand, interpret and effectively use data in decision-making and problem-solving processes.

3.1.1. Digital Competence: Developing digital skills enables stakeholders to leverage technology for data access, analysis and communication.

3.1.2. Statistical Literacy: Developing statistical literacy improves stakeholders' understanding of statistical concepts and methodologies, data interpretation and communication.

3.2. Data Governance Awareness: Raising awareness of data governance principles and practices promotes a culture of accountability, transparency and responsibility.

4. **Technology and Infrastructure:** Technology and infrastructure play a crucial role in supporting data governance initiatives in the TSS. This component focuses on the hardware, software, networks and connectivity needed to effectively manage and analyse data. Adequate investment in technology infrastructure ensures that organizations have the tools and resources to efficiently collect, store,

process and analyse data. Whether using open source or licensed software solutions, organizations should ensure that their technology infrastructure aligns with their data governance objectives and supports their operational needs.

4.1. Hardware: Infrastructure investment in hardware supports data processing, storage and analysis, enabling efficient and effective data management.

4.2. Software: The use of appropriate software solutions supports evidence-based decision-making by improving data processing, analysis and visualization capabilities.

4.2.1. Open Source Structures: Utilizing open source software promotes flexibility, innovation and cost-effectiveness in data management and analysis.

4.2.2. Licensed: Using licensed software provides access to specialized tools and features that support complex data analysis and management requirements.

4.2.3. In-house Solutions: Developing in-house software solutions enables customization and alignment with organizational needs and goals.

4.3. Networks and Connections: Establishing robust networks and connections facilitates data exchange, collaboration and interoperability with internal and external stakeholders.

5. Interoperability and Communication: Interoperability and communication are key components of an effective data governance framework that enables seamless data exchange and collaboration among stakeholders. This component emphasizes the importance of user engagement, interoperability and partnerships in promoting data sharing and integration. By engaging with internal and external stakeholders, organizations can ensure that data governance practices are aligned with user needs and preferences. Building partnerships at national and international level facilitates knowledge exchange, capacity building and alignment with global best practices and standards.

5.1. User Engagement: Involving stakeholders in data governance processes fosters collaboration, increases participation and ensures alignment with user needs and preferences.

5.2. Interoperability and Reusability: Ensuring interoperability and reusability of data and systems increases efficiency, facilitates data integration and supports innovation.

5.3. Partnerships: Collaborating with national and international partners strengthens data governance frameworks, promotes data sharing and improves data quality and relevance.

5.3.1. National: Partnering with national stakeholders promotes coordination, harmonization and integration of data governance efforts across sectors and organizations.

5.3.2. International: Participating in international partnerships facilitates knowledge exchange, capacity building and alignment with global best practices and standards.

13. Effective implementation of a comprehensive data governance framework is essential to ensure the integrity, accessibility and utility of data assets within the TSS. By addressing the key components outlined in this framework, TurkStat can improve its data governance practices, strengthen stakeholder trust, and fulfil its mandate to provide accurate and reliable statistical information to support informed decision-making and policy development in Türkiye.

B.   Data Governance Framework Platform

14.   The development of a data governance platform provides multiple benefits to the TSS and its stakeholders. First, it facilitates effective decision-making and compliance monitoring by providing a central hub to access and manage data governance policies, guidelines and resources. Furthermore, the platform promotes transparency and accountability by enabling stakeholders to track data governance activities, initiatives and outcomes. It also encourages collaboration and information sharing by facilitating communication and engagement among stakeholders, thereby improving data quality, integrity and usability. For these purposes, a data governance platform is being designed to serve as a vital tool for strengthening data governance practices, increasing organizational efficiency and achieving strategic objectives in the TSS.

15.   The platform is also envisioned as an opportunity to facilitate the sharing of the framework with partners and to observe gaps in the data governance framework (such as legal issues related to AI).

Data governance platform screenshot

C.   Data Governance Glossary

16.   Creating a data governance glossary in Turkish offers several benefits to stakeholders in the TSS and beyond. First, it promotes clarity and consistency in communication by providing a standardized set of terms and definitions related to data governance practices. This ensures that all stakeholders, regardless of their level of expertise, can correctly understand and interpret key concepts. A glossary also helps to overcome language barriers and facilitate collaboration between different stakeholders by providing English translations and explanations of technical terms. Furthermore, having a centralized source for data governance terminology enhances internal learning and knowledge sharing, enabling stakeholders to effectively implement and adhere to data governance principles and practices. For these purposes, a data governance glossary in Turkish and English is being created to serve as a valuable tool to promote transparency, compliance and accountability in data governance efforts.

17.   In order to create this glossary, an inventory of data governance terms was first created by reviewing all national and international literature. Then, a hierarchy of priorities is established, prioritizing international and official ones, and the glossary terms are individualized according to this hierarchy.

Data Governance Glossary app
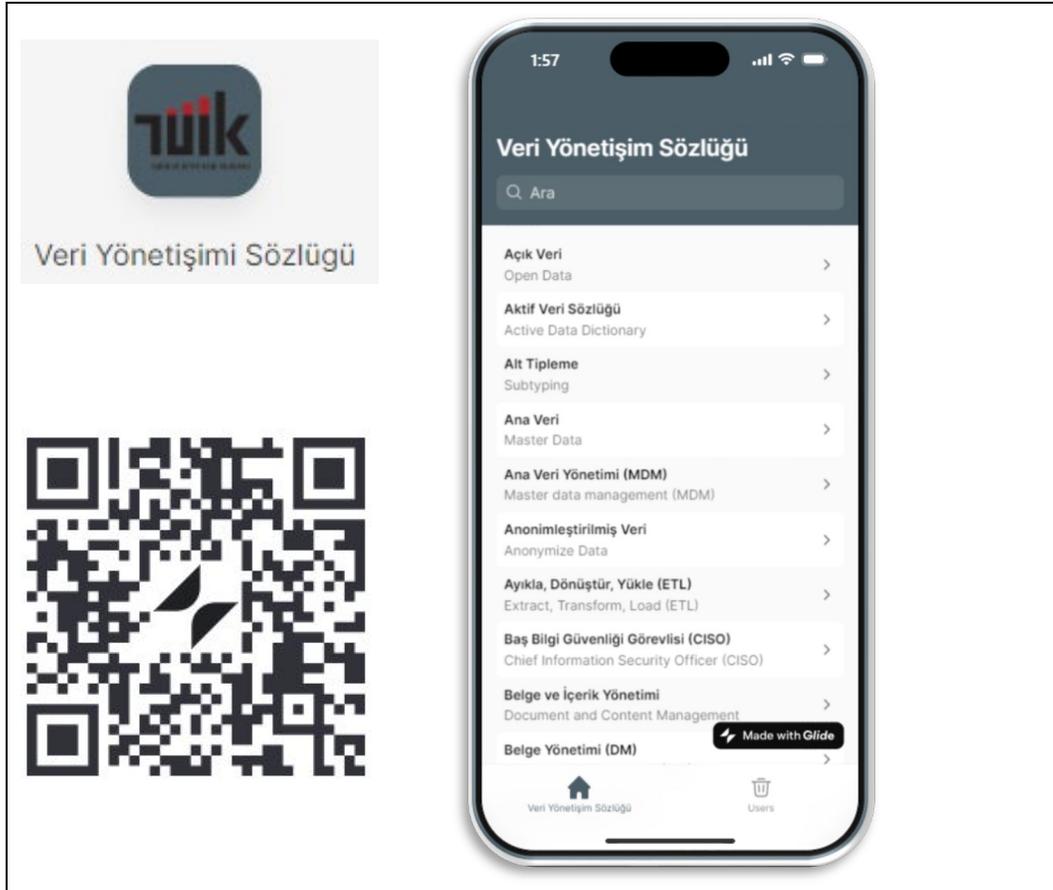
III.   CONCLUSION

18.   Clear and explicit policies and guidelines for national data governance should be established. This means establishing a legal framework covering issues such as data collection, storage, sharing, security and confidentiality. TurkStat's long-standing experience in data management provides a strong foundation for its efforts to develop a robust data governance framework for the TSS. By designing a comprehensive data governance framework for the TSS, TurkStat aims to protect the integrity of data assets, increase their accessibility, make them available to all and ensure their usability. This framework includes main components such as policy formulation, determination of guidelines, compliance with legislation, development of data management practices, promotion of a data-driven culture, improvement of technological infrastructure and encouragement of strategic partnerships.

19.   This comprehensive framework aims to identify and address national legislative needs in data governance, define roles and responsibilities, and ensure the integration of public data. Another objective is to bring together and present data governance components, which are currently dispersed, with a single entry point. The fact that some dead links, outdated information and gaps have already

been encountered while doing this is a proof of how accurate the work has been done. These findings are also kept separately by creating a gap analysis document.

20. With the steps it has taken on data governance, TurkStat is making significant contributions to strengthening national data governance and bringing the TSS in line with modern standards. TurkStat's efforts contribute to making Türkiye's official statistics more reliable, accessible and useful.

21. TurkStat's data governance policy can be expanded to the national level to include not only bodies within TSS but also all stakeholders such as public institutions, private sector and non-governmental organizations. Creating a national awareness on data governance and spreading this awareness to all stakeholders would be an important step towards transforming Türkiye into a data-driven economy. Besides, TurkStat's pioneering work in this area will make Türkiye a role model internationally. By adapting to and exceeding international standards on data governance, Türkiye can become a reference point for data-driven development models.

22. TurkStat's data governance efforts have the potential to be an important reference point for determining Türkiye's data governance framework, establishing a more robust and sustainable data governance infrastructure, and planning and actions to be taken in line with this framework. The importance of these studies and the main benefits to be achieved through a national data governance ecosystem can be summarized under the following headings:

- Establishment of a data governance ecosystem where roles and responsibilities are clearly defined.

- Integration of public data.

- More informed decision systems based on reliable statistics and better data.

- Reduced costs and risks; increased efficiency.

- Information society and new data-driven capabilities.

- Building a data-driven culture.

23. TurkStat continues its data governance efforts with all stakeholders to realize Türkiye's data potential and build a better future.

* * * * *