

Stanisław GALIJ¹, Grzegorz PAWLAK¹, Sławomir GRZYB²

¹Poznań University of Technology

²West Pomeranian University of Technology in Szczecin

WPROWADZENIE KLUCZOWYCH WSKAŹNIKÓW SUWERENNOŚCI DANYCH (KDSI) W CHMURZE PUBLICZNEJ

Streszczenie. Niniejszy artykuł wprowadza Kluczowe Wskaźniki Suwerenności Danych (KDSI), framework oparty na Key Performance Indicator (KPI) służący do ilościowego pomiaru suwerenności danych w środowisku chmury publicznej. W przeciwieństwie do istniejących modeli koncentrujących się na ocenie suwerenności cyfrowej państw lub dostawców, proponowane podejście ocenia suwerenność na poziomie wdrożenia chmurowego, umożliwiając zautomatyzowany pomiar w systemach nadzorujących.

INTRODUCING THE KEY DATA SOVEREIGNTY INDICATORS (KDSI) FOR PUBLIC CLOUD ENVIRONMENTS

Summary. This paper introduces the Key Data Sovereignty Indicators (KDSI), a Key Performance Indicator (KPI) based framework for quantifying data sovereignty in public cloud environments. Unlike existing models focused on national or provider-level readiness, the proposed approach evaluates sovereignty at the cloud deployment level, enabling continuous, automated measurement through observability systems.

1. Introduction

How can data sovereignty be measured, particularly in public cloud environments where sovereignty is inherently lower by design due to factors such as data localization constraints, third-party access, and the applicability of foreign jurisdictions and legislation?

Several academic studies ([4], [5], [7], [8]), practical frameworks, and composite indexes have already attempted to address the broader concept of digital sovereignty and to measure and compare sovereignty levels across digital environments.

One of the most recent initiatives is the European Union's Cloud Sovereignty Framework (CSF) [2], introduced by the European Commission in 2025 to support the evaluation of cloud service tenders and procurement processes.

The CSF comprises eight sovereignty objectives: 1) Strategic Sovereignty, 2) Legal and Jurisdictional Sovereignty, 3) Data and AI Sovereignty, 4) Operational Sovereignty, 5) Supply Chain Sovereignty, 6) Technology Sovereignty, 7) Security and Compliance Sovereignty, and 8) Environmental Sustainability. Each objective is assessed using a five-level Sovereignty Effectiveness Assurance Level (SEAL) scale.

The five SEAL levels range from L1 (Initial – ad hoc, undocumented controls) through L3 (Defined – standardized and documented controls) to L5 (Optimizing – continuous improvement of the sovereignty posture). This progression mirrors the Capability Maturity Model Integration (CMMI) scale adapted for sovereignty assessment purposes. The assessment results are subsequently aggregated to calculate the Overall Sovereignty Score, which is derived as a weighted aggregate of the individual sovereignty objectives. The weighting factors are assigned as follows: Strategic Sovereignty (SOV1) – 15%, Legal and Jurisdictional Sovereignty (SOV2) – 10%, Data and AI Sovereignty (SOV3) – 10%, Operational Sovereignty (SOV4) – 15%, Supply Chain Sovereignty (SOV5) – 20%, Technology Sovereignty (SOV6) – 15%, Security and Compliance Sovereignty (SOV7) – 10%, and Environmental Sustainability (SOV8) – 5%.

Academic approaches to sovereignty measurement have also emerged. One example is the Digital Sovereignty Index, introduced in 2022 [5], which is based on three existing indexes: the e-Government Development Index (EGDI) [5], the IMD Digital Competitiveness Index (DCI) [4], and the Barnett Sovereignty Index [1].

The e-Government Development Index (EGDI), produced by the United Nations Department of Economic and Social Affairs since 2001 and published biennially, evaluates the level of e-government development across all 193 United Nations Member States.

The IMD Digital Competitiveness Index (DCI), introduced in 2021 by the IMD World Competitiveness Center, ranks economies according to their digital competitiveness and innovation readiness. However, it does not directly measure the level of digital sovereignty.

The Barnett Sovereignty Index, introduced in 2017 see [1] and currently considered somewhat outdated, covers 122 countries and evaluates national sovereignty using equally weighted metrics related to legal recognition, governmental control.

The Digital Sovereignty Index appears to be more an indicator of countries' physical and political sovereignty, as well as their readiness for digital transformation, than a direct measure of digital sovereignty itself.

Another attempt to assess the digital sovereignty of 23 countries is the Digital Dependence Index (DDI) [7]. The study is based on data from 2019 and includes a comparison of selected trends with data from 2010. The DDI consists of three sub-indices: Hardware, Software, and Intellectual Property.

The study also proposes a four-tier Digital Dependence Index (DDI) framework, which classifies countries according to their level of digital reliance based on DDI values:

- Low sensitivity ($0 < \text{DDI} \leq 0.25$): Indicates a very high degree of digital autonomy, with domestic digital technologies occupying a dominant position.
- High sensitivity ($0.25 < \text{DDI} \leq 0.50$): Indicates that domestic providers supply the majority of digital technologies, reflecting a substantial level of resilience and self-sufficiency.
- Low vulnerability ($0.50 < \text{DDI} \leq 0.75$): Suggests that most digital technologies are sourced from global markets, resulting in moderate dependence on external suppliers.

- High vulnerability ($0.75 < \text{DDI} < 1$): Reflects a very low degree of digital autonomy, with foreign digital technologies playing a dominant role in the national digital ecosystem.

The authors conclude that the three countries with the lowest DDI values (indicating the highest levels of digital sovereignty) are the United States, followed by China and the Republic of Korea. Brazil, in contrast, has the highest DDI value among the countries analyzed, see Figure 1.

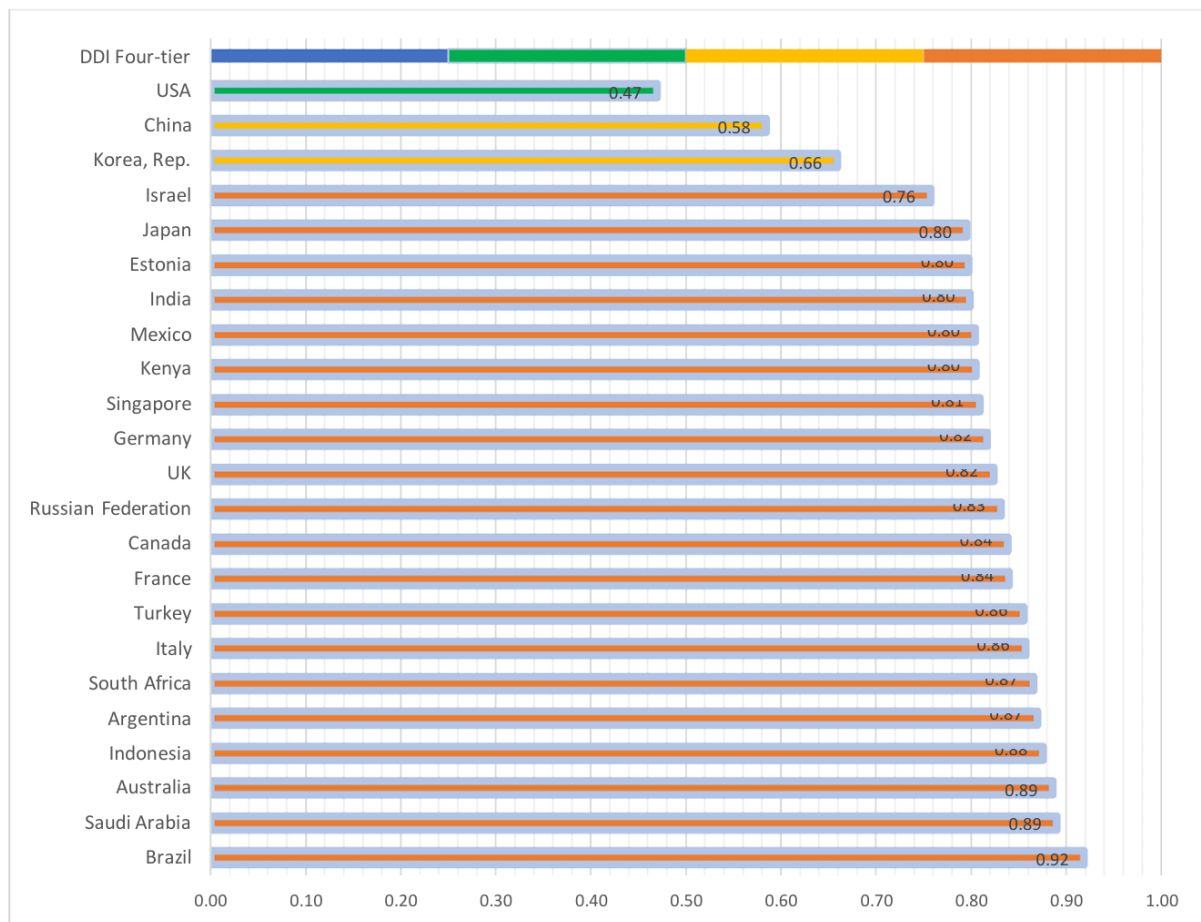


Figure 1. Digital Dependence Index (DDI), source [7].

The most recent research on digital sovereignty [8] introduces the Digital Sovereignty Index (DSI), which is composed of three sub-indicators: technological sovereignty, state sovereignty, and social sovereignty. The study reveals that, among a group of 50 countries that are either members of the OECD or have large populations, the highest DSI scores were achieved by the United States, followed by Finland and the Republic of Korea. In contrast, Nigeria received the lowest DSI score. The results are presented in Figure 2.

Although the aforementioned indices and frameworks provide valuable insights into countries' digital sovereignty, a significant research gap remains regarding the development of measures and indices capable of assessing data sovereignty within specific organizations or cloud environments. This gap carries direct practical consequences: regulatory frameworks such as the General Data Protection Regulation (GDPR) and the NIS2 Directive impose data sovereignty obligations not at the national

level but on individual data controllers and cloud tenants. Demonstrating sovereign compliance therefore requires quantitative evidence at the level of a specific cloud deployment – a granularity for which no standardized measurement framework currently exists.

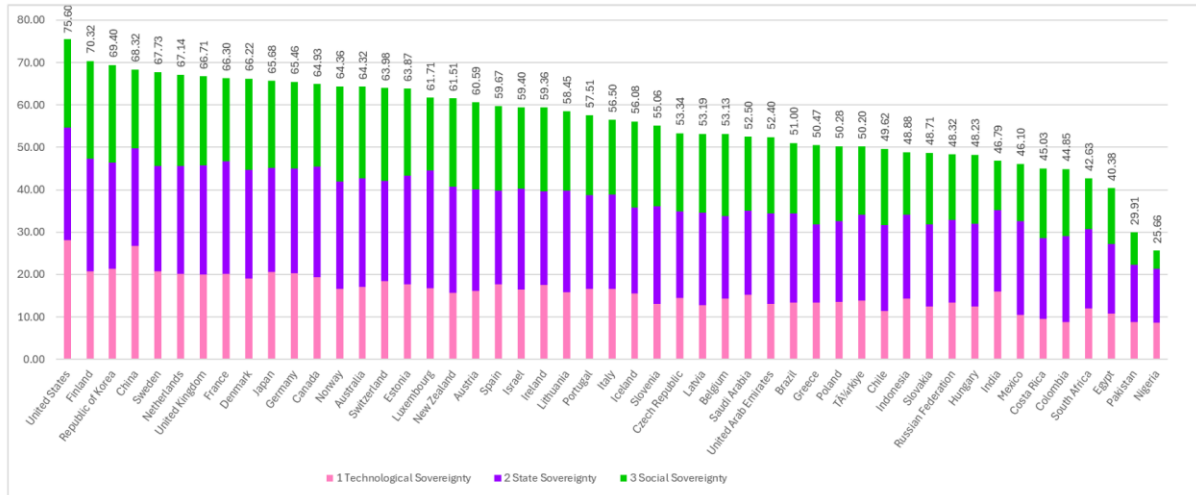


Figure 2. Digital Sovereignty Index and sub indicators, source: [8].

2. Key Data Sovereignty Indicators

The frameworks and indices discussed above primarily assess a country's digital transformation readiness or the digital sovereignty of cloud service provider offerings, rather than the actual characteristics of a specific cloud deployment (i.e., a particular cloud environment instance).

This raises a key question: how one can assess whether a specific cloud landing zone, or a multi-/hybrid-cloud deployment used by an organization, is sovereign? More precisely, what is the level of data sovereignty within a given cloud instance?

As data sovereignty lacks a single, universally accepted definition [3], it is challenging to establish a standardized set of metrics that can unequivocally determine whether data sovereignty requirements are satisfied within a particular cloud environment. Furthermore, the absence of a common definition complicates the development of quantitative methods for measuring the degree of data sovereignty achieved by a given system.

In information technology, system performance is commonly assessed using Key Performance Indicators (KPIs), which are measurable metrics used to evaluate whether a service or system meets defined operational or business objectives. Among the most widely used KPIs for assessing application performance are request latency and service availability. These KPIs are typically specified through a Service Level Indicator (SLI), which defines the metric being measured, and a Service Level Objective (SLO), which defines the target value for that metric.

For example, for the Request Latency KPI, an SLI may be defined as the 95th-percentile request latency (p95 latency), while a corresponding SLO may specify that the p95 latency should remain below 200 ms.

Similarly, the Service Availability KPI may be defined using an SLI expressed as the ratio of successful requests to total requests. A corresponding SLO could specify that service availability must be greater than or equal to 99.95% over a calendar month. A methodology analogous to the use of Key Performance Indicators (KPIs) for quantifying application performance is proposed. Within this methodology, six Key Data Sovereignty Indicators (KDSI) are defined for the systematic measurement and evaluation of data sovereignty properties:

- **Data Residency Compliance Rate** (score: 0-100%, weight: 20%)

This KDSI measures the proportion of data that remains within a predefined data residency boundary, such as a region, availability zone, or data center. Formally: **DRCR** = $(V_a / V_t) \times 100$, where V_a is the volume of data stored within the defined boundary and V_t is total data volume. Measurement sources include cloud provider location metadata, Cloud Security Posture Management (CSPM) tagging policies, and storage service APIs.

- **Cross-Border Data Retention Rate** (score: 0-100%, weight: 20%)

This metric quantifies the proportion of total data volume that is not transferred outside the defined domestic data centers or sovereignty boundary. Depending on the implementation, the indicator may be computed based on data volume or number of datasets. A higher **CBDRR** score indicates higher sovereignty, as more data remains within the defined boundary during transfer and exchange operations. Formally: **CBDRR** = $(V_n / V_t) \times 100$, where V_n is the volume of data not egressed outside the sovereignty boundary and V_t is total data volume. Equivalently, **CBDRR** = $100 - (V_e / V_t \times 100)$, where V_e is the volume of data egressed outside the sovereignty boundary. Measurement sources: VPC flow logs, network egress telemetry, cloud-native data transfer records.

- **Customer-Managed Encryption Rate** (score: 0-100%, weight: 20%)

This indicator represents the percentage of stored data encrypted using customer-managed encryption keys (CMEK). Formally: **CMER** = $(V_{\square} / V_t) \times 100$, where $V_{\square}$ is the volume of data encrypted with CMEK. Measurement sources: Key Management Service (KMS) audit logs and cloud storage metadata APIs. Reference values: regulated industries (finance, healthcare) typically achieve 85–100%; default cloud storage configurations without CMEK: 0%. A limitation of CMER is that it does not capture key rotation frequency or Hardware Security Module (HSM) usage, which should be tracked as complementary controls.

- **Metadata Containment Rate** (score: 0-100%, weight: 10%)

This metric measures the proportion of metadata that remains contained within the defined sovereignty boundary. The sovereignty boundary may be flexibly defined as a set of data centers, availability zones, or geographic regions. It is important to note that metadata can itself be highly sensitive: query patterns, access timestamps, and schema structures may reveal confidential business logic or organizational behavior. Formally: **MCR** = $((M_t - M_e) / M_t) \times 100$, where M_e is the number of metadata fields exposed outside the boundary and M_t is total metadata fields. Equivalently, **MCR** = $100 - \text{MEI}$, where MEI represents the former exposure-based formulation. Measurement sources: API gateway logs and service mesh telemetry.

- **Data Processing Encryption Rate** (score: 0-100%, weight: 15%)

This KDSI measures the proportion of data that is encrypted during processing operations within the computing environment. It reflects the extent to which data remains cryptographically protected while being actively processed in memory, runtime environments, or application workflows. The primary enabling technology for this indicator is confidential computing, including hardware-based Trusted Execution Environments (TEEs) such as Intel TDX, AMD SEV-SNP, and AWS Nitro Enclaves. Formally: $\mathbf{DPER} = (V_p / V_t) \times 100$, where V_p is the volume of data processed within an encrypted enclave. Measurement sources: hypervisor attestation logs and TEE telemetry APIs.

- **Mutually Authenticated Data Transfer Rate** (score: 0–100%, weight: 15%)

This KDSI quantifies the proportion of data transfers secured using mutual Transport Layer Security (mTLS) during inter-service and inter-system communication. It captures the extent to which data in transit is protected through mutual authentication of communicating endpoints and encrypted transport channels, ensuring bidirectional trust enforcement across infrastructure components. Formally: $\mathbf{MADTR} = (T_m / T_t) \times 100$, where T_m is the number of data transfers secured with mTLS and T_t is total transfers. Measurement sources: service mesh telemetry (e.g. Istio) and API gateway logs.

The scores for all of the above metrics range from 0% to 100%. The weighting scheme is derived from a risk-impact classification aligned with GDPR data protection principles and NIS2 security requirements. Indicators are grouped into three tiers: Critical (Data Residency Compliance Rate, Cross-Border Data Retention Rate, and Customer-Managed Encryption Rate – 20% each), reflecting direct regulatory exposure and data localization obligations; Standard (Data Processing Encryption Rate and Mutually Authenticated Data Transfer Rate – 15% each), reflecting important but implementation-dependent technical controls; and Supporting (Metadata Containment Rate – 10%), reflecting an indirect sovereignty signal. The weights sum to 100% and ensure that a failure in a Critical indicator has a proportionally greater impact on the composite DSI score.

Based on these KDSIs, an overall Cloud Data Sovereignty Index (CDSI) can be calculated using the following equation:

$$\text{CDSI} = \sum_{n=1}^{n=6} \frac{\text{score}(\text{KDSI}_n)}{\text{max. score}(\text{KDSI}_n)} \times \text{weight}(\text{KDSI}_n)\% \quad (1)$$

Moreover, the calculation of KDSI metrics and the CDSI can be automated and integrated into an observability stack to generate real-time measurements, enable visualization through sovereignty dashboards, and trigger alerts when defined data sovereignty thresholds are not met.

This approach is consistent with modern compliance methodologies, in which periodic audits are increasingly replaced by continuous compliance monitoring, see Figure 3.

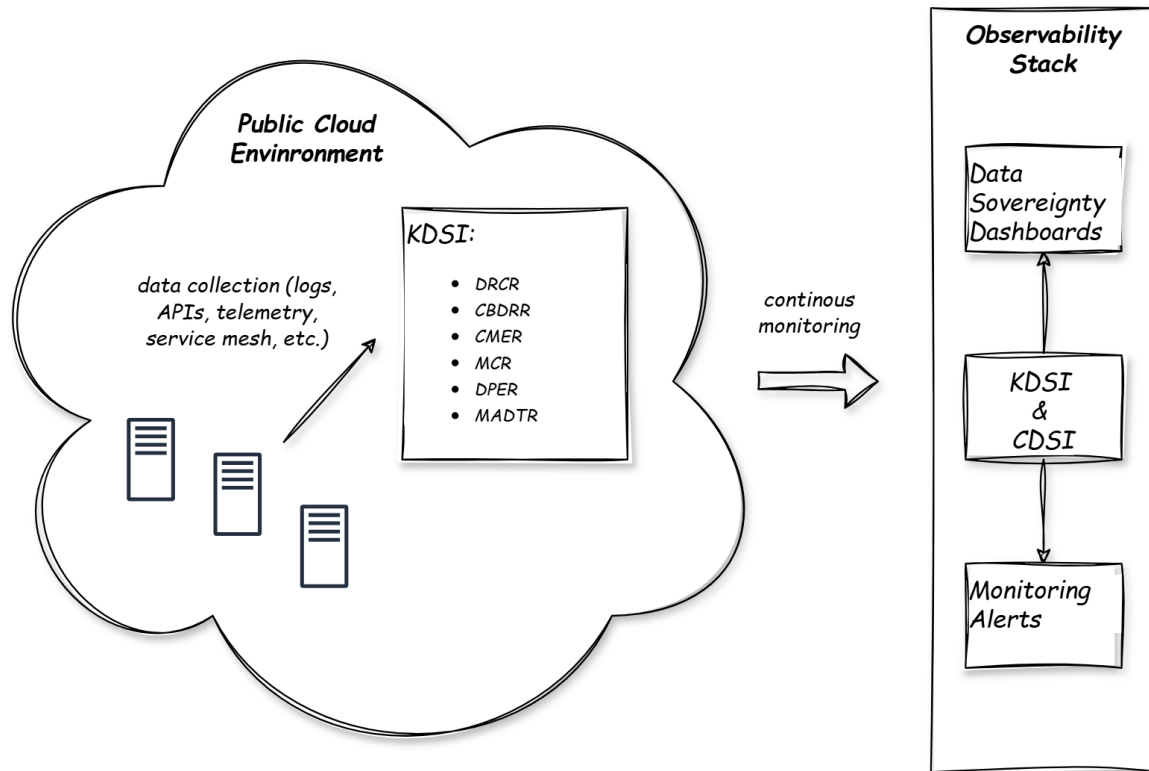


Figure 3. Key Data Sovereignty Indicators collection

3. Conclusions

This paper proposes a novel approach to quantifying data sovereignty in public cloud environments. In contrast to existing models that primarily assess digital sovereignty at national or cloud service provider (CSP) levels, this work introduces a framework designed to evaluate data sovereignty at the cloud deployment level, thereby addressing a gap in current research.

To overcome the lack of a universally accepted definition of data sovereignty, the paper defines a set of quantitative metrics for its assessment at the implementation level. These Key Data Sovereignty Indicators (KDSI) include: Data Residency Compliance Rate, Cross-Border Data Retention Rate, Customer-Managed Encryption Rate, Metadata Containment Rate, Data Processing Encryption Rate, and Mutually Authenticated Data Transfer Rate.

Furthermore, the paper demonstrates that the computation of KDSI and the overall Cloud Data Sovereignty Index (CDSI) can be automated and integrated into modern observability stacks. This enables real-time measurement, visualization through sovereignty dashboards, and alerting when predefined data sovereignty thresholds are violated. Consequently, the proposed approach shifts the paradigm from periodic compliance audits to continuous compliance monitoring.

The framework is explicitly designed with an operational focus, where KDSIs evaluate the observable characteristics of a specific cloud deployment. This provides a practical foundation for system designers and cloud architects to design and maintain sovereignty-compliant systems.

Despite the conceptual complexity and heterogeneous interpretations of data sovereignty, it remains essential to establish quantifiable and continuously measurable indicators to assess and enforce data sovereignty at the cloud deployment level, thereby moving beyond traditional national- or provider-centric evaluation models.

4. Limitations and Future Work

The proposed KDSI framework has several limitations that should be acknowledged. First, the current set of six indicators does not address sovereignty risks arising from cloud provider personnel access (i.e. the risk of insider threat from the cloud service provider), which may be captured in future extensions through an Access Sovereignty Rate. Second, the weighting scheme, while derived from regulatory alignment, has not yet been validated empirically through field studies; future work should include sensitivity analysis of weight assignments using real-world deployment data. Third, measurement of DPER currently depends on TEE availability, which limits applicability to cloud providers and instance types that support confidential computing. Future research directions include: (1) empirical validation of KDSIs across representative cloud deployment configurations; (2) integration of the framework with existing compliance automation tools (e.g. Open Policy Agent); and (3) extension of the CDSI to cover AI workloads processing sovereign data, aligning with the EU AI Act requirements.

LITERATURE

1. Barnett M., A.: Quantifying sovereignty: A new way to examine an essential concept, 2017.
2. Cloud Sovereignty Framework, European Commission, available on-line: https://commission.europa.eu/document/download/09579818-64a6-4dd5-9577-446ab6219113_en , accessed on: 30 April 2026.
3. Hummel P. et al.: Data sovereignty: A review, *Big Data & Society* 8.1, 2021: 2053951720982012.
4. IMD Digital Competitiveness Index, available on-line: <https://www.imd.org/centers/world-digital-comparison/> , accessed on: 27 May 2026
5. Kaloudis M.: Sovereignty in the Digital Age–How Can We Measure Digital Sovereignty and Support the EU’s Action Plan?, *New Global Studies* 16.3, 2022, p: 275-299.
6. UN E-Government Survey 2024, United Nations, Department of Economic and Social Affairs, available on-line: <https://desapublications.un.org/publications/un-e-government-survey-2024> , accessed on: 27 May 2026.
7. Mayer M. and Yen-Chi L.: Digital Autonomy? Measuring the Global Digital Dependence Structure, 2022.
8. Rodríguez Pita M.P., Pérez Martínez J.M, Urueña López A.: From concept to method: A framework for measuring digital sovereignty, ITS 33rd European Conference 2025: "Digital innovation and transformation in uncertain times", Edinburgh, UK, 29th June – 1st July 2025, International Telecommunications Society (ITS), Calgary.