European Genomic Data Infrastructure

# Deliverable D5.5

*Report outlining QoS metrics, reporting processes, and node TRLs defined*

| | | | |
|---|---|---|---|
| **Project Title**<br>Grant agreement no | **Genomic Data Infrastructure**<br>Grant agreement 101081813 | | |
| **Project Acronym** (EC Call) | GDI | | |
| **WP No & Title** | WP5: Technical Coordination and Outreach | | |
| **WP Leaders** | Tommi Nyrönen (6. CSC) | | |
| **Deliverable Lead Beneficiary** | 6. CSC | | |
| **Contractual delivery date** | 31/10/2023 | **Actual delivery date** | 04/12/2023 |
| **Delayed** | Yes | | |
| **Partner(s)**<br>contributing to deliverable | CSC, EMBL, CRG | | |
| **Authors** | Dylan Spalding (CSC) | | |
| **Contributors** | N/A | | |
| **Acknowledgements** | N/A | | |
| **Reviewers** | Thomas Keane (EMBL-EBI)<br>Teresa D'Altri (EGA-CRG)<br>Mallory Freeberg (EMBL-EBI) | | |

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

European Genomic Data Infrastructure

## *Log of changes*

| Date | Mvm | Who | Description |
|---|---|---|---|
| **31/10/2023** | 0v1 | Dylan Spalding (CSC) | Initial document sent for review |
| **14/11/2023** | 0v2 | Dylan Spalding (CSC) | Reviewer comments addressed |
| **24/11/2023** | 0v3 | Nikki Coutts (ELIXIR Hub) | Copy circulated to the GDI-MB for review |
| **04/12/2023** | 0v4 | Dylan Spalding (CSC) | Comments addressed |
| **04/12/2023** | 1v0 | Nikki Coutts (ELIXIR Hub) | Final version submitted to the EC portal |

## *Table of contents*

## Contents




GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

European Genomic Data Infrastructure

# 1. Executive Summary

Version 1.0.0 of the list of metrics to monitor and record the service levels and operation of the European Genomic Data Infrastructure (GDI) are listed in this document. Input was given into the metrics by a range of different parties, including those who currently operate data services, as well as other metrics such as those from ELIXIR. Additional input was taken from other work within the GDI, such as deliverables and workshops, with the aim to make the metrics cover the full range of GDI services and operations.

These metrics were developed as a basis for the Standard Operating Procedures (SOPs) being defined, linked to the Helpdesk Roadmap, and mapped to the quarterly Operational Readiness level monitoring of the nodes. This allows nodes to demonstrate progress towards operation by showing the generation, monitoring, and where necessary logging of the specific metrics listed here. The metrics were classified into operational groups, and sub-classified into stages of the GDI data journey.

The metrics are designed to monitor the operational status of the GDI, as well as feedback the impact and performance of the GDI to stakeholders. Operationally these metrics will be fed to the Operations Committee as defined in milestone MS10, however the processes for reporting these metrics will evolve as the legal entity of the GDI is defined as well as the SOPs.

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

## 2. Contribution towards project outcomes

With this deliverable, the project has reached or the deliverable has contributed to the following project outcomes:

| | Contributed |
|---|---|
| **Outcome 1**<br><br>Secure federated infrastructure and data governance needed to enable sustainable and secure cross border linkage of genomic data sets in compliance with the relevant and agreed legal, ethical, quality and interoperability requirements and standards based on the progress achieved by the 1+MG initiative. | No |
| **Outcome 2**<br><br>Platform performing distributed analysis of genetic/genomic data and any linked clinical/phenotypic information; it should be based on the principle of federated access to data sources, include a federated/multi party authorisation and authentication system, and enable application of appropriate secure multi-party and/or high-end computing, AI and simulation techniques and resources. | Yes |
| **Outcome 3**<br><br>Clear description of the roles and responsibilities related to personal data and privacy protection, for humans and computers, applicable during project lifetime and after its finalisation. | No |
| **Outcome 4**<br><br>Business model including an uptake strategy explaining the motivation, patient incentives and conditions for all stakeholders at the different levels (national, European, global) to support the GDI towards its sustainability, including data controllers, patients, citizens, data users, service providers | No |

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

| | |
|---|---|
| (e.g., IT and biotech companies), healthcare systems and public authorities at large. | |
| Outcome 5<br><br>Sustained coordination mechanism for the GDI and for the GoE multi-country project launched in the context of the 1+MG initiative. | No |
| Outcome 6<br><br>Communication strategy – to be designed and implemented at the European and national levels. | No |
| Outcome 7<br><br>Capacity building measures necessary to ensure the establishment, sustainable operation, and successful uptake of the infrastructure. | No |
| Outcome 8<br><br>Financial support to the relevant stakeholders to enable extension, upgrade, creation and/or physical connection of further data sources beyond the project consortium or to implement the communication strategy and for capacity-building. | No |

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

European Genomic Data Infrastructure

# 3. Methods

To operate an infrastructure that provides a service(s) to customers, it is required to monitor the performance and security of these services. This helps to ensure that the current service level is adequate for the customers and stakeholders, allows planning for any necessary changes in capacity or performance, and report the performance of the infrastructure to relevant stakeholders. Existing operational service providers of genomic data access were asked for input into the types of metrics and reporting processes that were necessary to ensure that a federated network of nodes would perform in such a way to meet the needs of their users and stakeholders. Specific operators were Sensitive Data Services[1], and the central European Genome-phenome Archive (EGA)[2], as well as other nodes of Federated EGA[3] who are also part of the GDI.

Reference was made to the 'statistics' part of the EGA website, to determine what metrics they recorded, especially the ones related to the impact of the resource.

Additionally a workshop was held during the Pillar II meeting in Rome[4] between 09.10.23 and 11.10.23 on stress and compliance testing. By defining the specific stress and compliance tests required within the federated infrastructure, the necessary metrics that need to be monitored can be defined to ensure these tests cover all the operations of a deployed node, as well as the whole federated infrastructure.

Another important aspect of defining the metrics and reporting processes is the relationship with the standard operating procedures (SOPs) which are currently being defined in WP4, and are due to be released by November 2024. We collaborated with the leads of this task (Task 4.3) and agreed to utilise the existing SOPs[5] used by Federated EGA as a basis for the metrics that would need to be monitored to ensure the correct application of an SOP.

To help enable to monitoring of the progress of nodes through the various phases of GDI (onboarding, deployment, operational), the metrics were also mapped to the specific steps in the GDI Operational Readiness Monitoring[6] tool, which is completed by the nodes every quarter. These steps are mapped onto the Technical Readiness Levels (TRLs) and the GDI phases, and where possible, the Maturity Model.

Each metric was assigned a group related to the type of monitoring the metric was used for, as well as a sub-group for the stage of the data journey[7] that the specific metric relates to (where applicable)

[1] https://research.csc.fi/sensitive-data-services-for-research
[2] https://ega-archive.org/
[3] https://ega-archive.org/about/projects-and-funders/federated-ega/
[4] https://docs.google.com/document/d/1BVQ20ap2LOzmonWWIdANWdVf-EBD9gw4G4wuavzEJwE/edit#heading=h.qjdgxs
[5] https://drive.google.com/drive/folders/14yFvXOxRyGl-ENoqIB5TdoqIUdL-qmfk
[6] https://docs.google.com/spreadsheets/u/0/d/1m9QckgkYXy-6JLSwzcMXNhERpV61nW_OO2Vq1KS55l0/edit
[7] https://zenodo.org/records/8279697

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

(Tables 1&2) . These were checked against the metrics[8] used by ELIXIR for the Core Data Resources[9] (CDRs) to ensure that the same types of metrics used for ELIXIR CDRs were also used in GDI.

**Table 1**: List of groups for the metrics used to monitor GDI performance.

| Group | Description |
|---|---|
| Helpdesk | Metrics required to ensure the efficient operation of the node and virtual helpdesks |
| Security | Metrics to monitor the compliance of the infrastructure to data protection principles and regulatory requirements |
| Impact | Metrics used to monitor the impact of the GDI, both with users, stakeholders, and to the community as a whole |
| Service Monitoring | Metrics from the compliance and stress tests, as well as KPIs and associated SLAs |

**Table 2**: Stages of the 1+MG Data Journey

| Term | Description |
|---|---|
| Data Preparation | Pre-processing of the data to ensure it meets the agreed standards, including the required metadata |
| Data Inclusion | Transfer of the data, physical and/or legal, into the GDI |
| Data Storage and Management | Storage of the data, versioning, including compliance to all necessary regulations |
| Data Discovery | Discovery of the data, either via the the User Portal or via Beacons |
| Data Access | The mechanism(s) through which data access is granted |
| Data Use | The authorised processing of the data for approved purposes to achieve the desired result |
| Data Archiving | Archiving, where necessary (e.g. research finding verification). |

[8] https://f1000research.com/articles/5-2422
[9] https://elixir-europe.org/platforms/data/core-data-resources

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

# 4. Description of work accomplished

## 4.1 Metric List

To be fit for purpose, the metric chosen must support the Helpdesk Roadmap[10] which is based on the FitSM[11] standard for IT service management, and hence the processes within FitSM. This relies on suitable metrics to ensure the service is operating within the acceptable levels, and must support the monitoring and reporting processes defined in FiTSM, particularly the ones listed in Table 3.

**Table 3**. List of FiTSM processes which rely extensively on the metrics defined in this document. The Who column determines who is responsible for the task, with 'N' indicating a node level task, 'E' a European level operations task, and a task for the node and European level operations.

| Process | Process Code | Roadmap Tasks | Who |
|---|---|---|---|
| Service Level Management | SLM | Define service catalogue<br>Generate knowledge base<br>Define Service Level Agreements (SLAs), if necessary<br>Define Operational Level Agreements (OLAs), if necessary<br>Define service targets | B<br>E<br>E<br>B<br>B |
| Service Reporting | SRM | Define reports required | B |
| Service Availability and Continuity Management | SACM | Produce service availability and continuity plans to support SLAs / OLAs as required (with WP6)<br>Monitor availability | B<br>B |
| Capacity Management | CAPM | Define capacity plan<br>Monitor utilisation of resources | N<br>B |
| Customer Relationship Management | CRM | Identify key customers and communication channels<br>Define service review and complaint handling procedures | E<br>E |
| Release and | RDM | Define a release and deployment strategy | E |

[10] https://zenodo.org/records/8017873
[11] https://www.fitsm.eu/

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

| | | | |
|---|---|---|---|
| Deployment Management | | Create monitoring and retrospective procedures for release and deployment actions | E |
| Continual Service Improvement | CSI | Identify opportunities for improvement of services and SMS<br>Ensure consistent evaluation of services and SMS<br>Support onboarding of new nodes to operational status<br>Survey other federated helpdesk operators | E<br>E<br>E<br>E |

Compliance and stress testing are currently included as part of service monitoring. One of the aims of stress testing is to determine the boundaries between safe, at risk, and failure operation statuses. These boundaries will be influenced by the infrastructure choices of each node, and hence cannot be quantified until the stress tests have been performed. Similarly to the stress tests, the metrics can also be used to help define the appropriate service level expected from each node, and the infrastructure as a whole. These support the helpdesk roadmap which requires the expected service level per node to be defined (specifically Service Level management in Table 3), so that appropriate mitigation SOPs can be used if the quality of service drops below a certain level.

The results from the compliance and stress test workshop, plus the input from other production services, were listed into Table 4, and in this spreadsheet[12]. Each row corresponds to a metric, and each metric has an ID, group, associated data journey stage, and operational readiness levels associated with the metric. A majority of the metrics are collected at the national level, so the operation of each national node can be monitored. It is recommended that where a national node is federated that the same metrics are monitored and recorded for each node within the national federated node. This is especially important for security and service level metrics as it allows problems to be quickly identified, isolated, and fixed. In the case of a federated node, the metrics should be summed or aggregated to generate the metrics for the node as a whole. The metrics for each node can then be summed or aggregated to provide the same metric for the whole of the GDI infrastructure.

The impact group of metrics can be used to demonstrate the impact of the infrastructure, and as the User Portal comes online it is expected that additional metrics may be possible to record, such as publications referencing GDI data.

Metrics are given as a generic metric which should be collected by the smallest operational object as possible, as these can be aggregated in multiple ways depending on the use case, and they can be used to help determine the different logging requirements within the GDI. For example, an impact data inclusion metric would be the number of countries who are adding data to the GDI, however by expressing the metric as 'Countries adding data to GDI' the source data can be aggregated as required to support different national and regional use cases. Many metrics may belong to multiple different groups, particularly the service monitoring and security group. However for clarity those metrics which are not solely related to security are classified under their primary group and data

[12] https://docs.google.com/spreadsheets/d/1O1A4NfdEQNXrlNlvZULljHdhxzH1M-YRFKhzfttPWxA/edit#gid=0

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

journey stage. However the relevant security SOPs will reference these metrics as part of the security and data protection monitoring of the infrastructure.

The technical implementation for gathering the metric information can occur prior to the operational implementation, and where a metric is linked to a specific TRL the lower (usually technical) TRL is chosen, even if the metric has no meaning at that level. For example, metric 4.3 which measures the time the central DAC takes to grant or deny access to a dataset or virtual cohort can be technically implemented by TRL 6, but will not be operationally implemented until TRL 8. In this case the relevant TRL is set to 6 to ensure that the technical implementation is done and tested during the deployment phase.

**Table 4**: List of metric and associated groups and data journey stage for the GDI.

| ID | Group | Data Journey Stage | Name | TRL |
|---|---|---|---|---|
| 1.1 | Helpdesk | N/A | Ticket human response time | 9 |
| 1.2 | Helpdesk | N/A | Ticket resolution time | 9 |
| 1.3 | Helpdesk | N/A | Total number of tickets received per time period | 9 |
| 1.4 | Helpdesk | N/A | Number of tickets per class, where a class is the classification applied to a ticket once it is first dealt with | 9 |
| 1.5 | Helpdesk | N/A | Response time for tickets transferred from central HD to node HD | 9 |
| 1.6 | Helpdesk | N/A | Number of interactions with the customer before a ticket is resolved | 9 |
| 1.7 | Helpdesk | N/A | Number of unclassified tickets | 9 |
| 2.1 | Impact | Access | Number of users accessing data in GDI | 6 |
| 2.2 | Impact | Access | Volume of data accessed in GDI | 6 |
| 2.3 | Impact | Access | Number of access requests granted | 6 |
| 2.4 | Impact | Access | Total number of access requests | 6 |
| 2.5 | Impact | Access | Total number of unique users requesting access | 6 |

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

| | | | | |
|---|---|---|---|---|
| 2.6 | Impact | Access | Number of access requests per class of data use (e.g. DUO) | 6 |
| 2.7 | Impact | Access | Geographic location of users requesting access | 6 |
| 2.8 | Impact | Discovery | Number of hits to the User Portal | 6 |
| 2.9 | Impact | Discovery | Number of Beacon queries used to define a virtual cohort | 6 |
| 2.10 | Impact | Inclusion | Volume of data GDI use cases or stakeholders adding to GDI | 8 |
| 2.11 | Impact | Inclusion | Countries adding data to GDI | 8 |
| 2.12 | Impact | N/A | Number of publications referencing GDI | 9 |
| 2.13 | Impact | N/A | Number of publications referencing GDI data | 9 |
| 2.14 | Impact | N/A | Number of participants data per node | 6 |
| 2.15 | Impact | N/A | Number of participants per node from a clinical setting | 6 |
| 2.16 | Impact | N/A | Number of incidental findings per node | 9 |
| 2.17 | Impact | N/A | Number of references to a GDI UID per impact factor bucket | 9 |
| 2.18 | Impact | N/A | Number of FTE per node per operation class - HD (1st, 2nd line), DevOps, Management, Data management | 1 |
| 2.19 | Impact | Preparation | Use cases or stakeholders preparing data for GDI | 8 |
| 2.20 | Impact | Storage and management | Types of files within GDI | 6 |
| 2.21 | Impact | Storage and management | Number of files in GDI, per file type | 6 |
| 2.22 | Impact | Storage and management | Volume of data available within GDI | 6 |

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

| | | | | |
|---|---|---|---|---|
| 2.23 | Impact | Use | Number of diagnosis made after accessing genomic data from GDI | 9 |
| 2.24 | Impact | Use | Data type per access request | 6 |
| 2.25 | Impact | Use | Available CPU in GDI | 8 |
| 2.26 | Impact | Use | Available storage in GDI as a whole | 8 |
| 3.1 | Security | N/A | Number of rejected logins to a data endpoint | 6 |
| 3.2 | Security | N/A | Number of rejected logins to a compute endpoint | 6 |
| 3.3 | Security | N/A | Number of DoS attacks per node | 6 |
| 3.4 | Security | N/A | Number of DosS attacker on User Portal | 6 |
| 3.5 | Security | N/A | Number of data breaches per node | 6 |
| 3.6 | Security | N/A | Number of security incidents per node | 6 |
| 3.7 | Security | N/A | Number of participants per access request | 6 |
| 3.8 | Security | Use | Files detected outside authorised use | 9 |
| 4.1 | Service Monitoring | Access | Number of users accessing data per node | 4 |
| 4.2 | Service Monitoring | Access | Volume of data access per node | 4 |
| 4.3 | Service Monitoring | Access | Time take from access request to access decision, central DAC | 6 |
| 4.4 | Service Monitoring | Access | Time taken from access request to access decision, node DAC | 6 |
| 4.5 | Service Monitoring | Access | Number of access requests approved by the central DAC | 6 |
| 4.6 | Service Monitoring | Access | Number of access requests approved by the node dac | 6 |
| 4.7 | Service Monitoring | Access | Number of access requests vetoed | 6 |
| 4.8 | Service Monitoring | Access | Reasons cited for vetoing access by a node | 6 |

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

| | | | | |
|---|---|---|---|---|
| 4.9 | Service Monitoring | Access | Number of consent withdrawal per node | 6 |
| 4.10 | Service Monitoring | Access | Number of virtual cohorts defined | 6 |
| 4.11 | Service Monitoring | Access | Latency in revoking access to data | 6 |
| 4.12 | Service Monitoring | Access | Latency in approving access to data | 6 |
| 4.13 | Service Monitoring | Access | Latency in checking a users access rights | 4 |
| 4.14 | Service Monitoring | Access | Latency in releasing a dataset and access being possible | 4 |
| 4.15 | Service Monitoring | Archiving | Datasets or virtual cohorts archived for reuse | 6 |
| 4.16 | Service Monitoring | Archiving | Number of datasets or virtual cohorts with more than 1 version archived for reuse | 6 |
| 4.17 | Service Monitoring | Discovery | Number of registered level discovery queries | 6 |
| 4.18 | Service Monitoring | Discovery | Total number of Beacon queries | 4 |
| 4.19 | Service Monitoring | Discovery | Number of Beacon queries per model object (g_variants, biosample, etc) | 4 |
| 4.20 | Service Monitoring | Discovery | Beacon response time per node | 4 |
| 4.21 | Service Monitoring | Discovery | Beacon Network uptime | 6 |
| 4.22 | Service Monitoring | Discovery | Beacon network latency | 6 |
| 4.23 | Service Monitoring | Discovery | Metadata API uptime per node | 6 |
| 4.24 | Service Monitoring | Discovery | Number of metadata responses per node | 6 |
| 4.25 | Service Monitoring | Discovery | Number of Beacon 'positive' responses per node | 4 |
| 4.26 | Service Monitoring | Discovery | Latency of Beacon response | 4 |
| 4.27 | Service Monitoring | Discovery | Latency of Beacon Network response | 6 |
| 4.28 | Service Monitoring | Inclusion | Files failing submission checks | 6 |
| 4.29 | Service Monitoring | Inclusion | Data ingestion rate per node | 4 |

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

| | | | | |
|---|---|---|---|---|
| 4.30 | Service Monitoring | Inclusion | Number of failed consistency checks | 4 |
| 4.31 | Service Monitoring | Inclusion | Files without necessary metadata | 6 |
| 4.32 | Service Monitoring | N/A | Uptime of a node | 4 |
| 4.33 | Service Monitoring | N/A | Uptime of the User Portal | 6 |
| 4.34 | Service Monitoring | N/A | Uptime of a product per node | 4 |
| 4.35 | Service Monitoring | N/A | Uptime of a service per node | 4 |
| 4.36 | Service Monitoring | N/A | Internode network capacity | 6 |
| 4.37 | Service Monitoring | N/A | Intranode network capacity | 4 |
| 4.38 | Service Monitoring | N/A | Number of times SLA was breached | 8 |
| 4.39 | Service Monitoring | Storage and management | Volume of data available per node | 4 |
| 4.40 | Service Monitoring | Storage and management | Number of files per node, and per file type | 4 |
| 4.41 | Service Monitoring | Storage and management | Types of files within a node | 4 |
| 4.42 | Service Monitoring | Storage and management | Number of files never accessed | 4 |
| 4.43 | Service Monitoring | Storage and management | Volume of data never accessed | 4 |
| 4.44 | Service Monitoring | Storage and management | Volume of data available for distribution as opposed to restricted to the node SPE | 4 |
| 4.45 | Service Monitoring | Storage and management | Volume of data within GDI but not released | 6 |
| 4.46 | Service Monitoring | Storage and management | Volume of storage available in a node | 4 |
| 4.47 | Service Monitoring | Storage and management | Number of datasets or virtual cohorts with more than 1 version | 6 |

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

| | | | | |
|---|---|---|---|---|
| 4.48 | Service Monitoring | Storage and management | Files flagged for deletion | 4 |
| 4.49 | Service Monitoring | Storage and management | Files deleted | 4 |
| 4.50 | Service Monitoring | Storage and management | Latency between flagging a file for deletion and completion of the deletion process | 4 |
| 4.51 | Service Monitoring | Storage and management | Files failing consistency checks | 4 |
| 4.52 | Service Monitoring | Storage and management | Files failing encryption | 4 |
| 4.53 | Service Monitoring | Storage and management | Files failing decryption for distribution | 4 |
| 4.54 | Service Monitoring | Use | Data distribution rate per node | 4 |
| 4.55 | Service Monitoring | Use | Number of CPU per node | 4 |
| 4.56 | Service Monitoring | Use | Files failing distribution checks | 4 |
| 4.57 | Service Monitoring | Use | Latency in requesting data and data becoming available | 4 |

## 4.2 Reporting Process

The flow of information, in this case the metrics, would be from the node to the centre, opposite to the ticket flow from the virtual helpdesk to the node helpdesk as described in D4.1. However, as the virtual helpdesk in D4.1 is operated by helpdesk members at the different nodes, this will not in itself work as a central point for gathering, summing or aggregating, and monitoring the metrics generated within the GDI. At a technical level, it is proposed that the metrics determined at node level, for example the volume of data within a node, could be made available via a secure API to a secure part of the User Portal or similar central portal which can monitor the performance of the GDI infrastructure as a whole. Authorised users would be able to log into such a portal, and visualise the metrics, both from particular nodes but also summed across the whole of the GDI. Currently it is proposed that the non-security metrics or those which are not monitored for security implications, are fed to the Operations committee as defined in MS10 - Governance structure, including operational, security, and development committees defined via the secure part of the User Portal or

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

similar. However, as GDI becomes operational, users authorised to access this would indeed to be identified in collaboration with Pillar I as it depends on the final legal framework of GDI, such as an European Digital Infrastructure Consortium (EDIC). Similarly, the security metrics and those metrics which can help identify security incident(s) depending on their value, would be directed to the Security and Data Protection Committee, or sub-group as identified by them, such as a Computer Security Incident Response Team (CSIRT). However, the full process for reporting these metrics, both operational and security, will need to be defined in conjunction with Task 4.3 which defines the SOPs for the GDI.

## 5. Results

The first version of metrics that should be recorded by different actors within the GDI can be utilised to help build the SOPs required for operation of the GDI, as well as the testing requirements and the definition of the required service level across the whole GDI. The metrics are divided into four main groups, which are subdivided into the seven stages of the data journey. This helps identify issues within the data journey itself, while also helping to identify those responsible for the monitoring, recording, and acting upon these metrics.

## 6. Discussion

The list of metrics presented here should be seen as a first version of the list of metrics to record and monitor. It is expected that extra metrics may be needed once nodes start to deploy, and the different compliance and stress tests begin. The definition and operation of the SOPs within the GDI will also affect the metrics used, as these SOPs are yet to be fully defined and will evolve based on stakeholder feedback. As part of the Continual Service Improvement (CSI) FiTSM stage listed in Table 3 the value and applicability of the metrics, as well as the tests and variables they measure, should be reviewed on a regular basis as part of the operation of the GDI. The lists of metrics need to be versioned using Semantic Versioning to ensure that the metrics remain interoperable across the whole of the GDI, and as such is part of the Release and Deployment management (RDM) process within FiTSM.

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.

## 7. Conclusions & Impact

The metrics are a core part of delivering the helpdesk roadmap as well as the SOPs being defined, and ultimately the User Portal, if a central access point for aggregated metrics is agreed upon. The metrics listed here aim to ensure that any variation in service level is identified quickly. For service level reduction the issue can be resolved as fast as possible, while for service level improvement lessons can be learnt as well to ensure that the needs of the GDI stakeholders are met via continual service improvement

## 8. Next steps

To ensure the metrics are collected and monitored, the metrics need to be mapped to the respective product(s) which will collect the information. The location where this information is stored needs to be defined, as well as how the results can be aggregated to give information on the service level of the GDI as a whole. This will require input from both Pillars I & 3, and especially WP6 from Pillar II as these data will form part of the Data Management Plan being implemented by the different nodes.

Technologies need to be determined on how to best record, display, and disseminate the metrics monitored, and under the proposal within the document this would require working with WP4 on the User Portal to enable a central point of access to the aggregated metrics for the GDI as a whole. There will also need to be coordination with Task 4.3 developing the GDI SOPs to ensure that metrics collecting and reporting is represented, as needed.

GDI project receives funding from the European Union's Digital Europe Programme under grant agreement number 101081813.