



Deliverable D4.2

Standard Operating Procedures defined

Project Title Grant agreement no	Genomic Data Infrastructure Grant agreement 101081813		
Project Acronym (EC Call)	GDI		
WP No & Title	WP4: European Level Operations		
WP Leaders	Christophe Trefois (5. PNED)		
Deliverable Lead Beneficiary	EMBL-EBI		
Contractual delivery date	31/10/2024	Actual delivery date	28/10/2024
Delayed	No		
Partner(s) contributing to deliverable	EMBL-EBI, CSC, CAD, CNRS, UU		
Authors	EMBL-EBI, UU		
Contributors	Marcos Casado Barbero (EMBL-EBI) Silvia Bahena (EMBL-EBI) Mallory Freeberg (EMBL-EBI) Thomas Keane (EMBL-EBI) Gabi Rinck (EMBL-EBI) Erik Hedman (UU) Markus Englund (UU) Mattias Strömberg (UU) Elisavet Torstensson (UU) Bianca Hendriksze (Health-RI)		



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

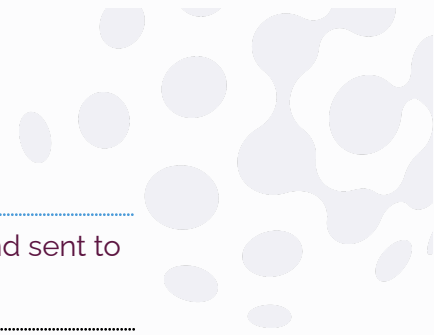


	Jorge Oliveira (IST)
	Pedro Ferreira (IST)
	David Salgado (CAD)
	Milan Ojsteršek (UM)
	Stefanie Kirschenmann (CSC)
	Dylan Spalding (CSC)
	Dilza Campos (VIB)
	Alja Videtič Paska (SI)
	John Hancock (UL)
	Miguel Santos (IST)
	Jorge Silva (UA)
	Anaïs Breton-Bellot (LNDS)
	Svetomir Hitov (MUS)
	Martin Georgiev (MUS)
Acknowledgements	Rob Hooft (Health-RI), Juan Arenas Marquez (ELIXIR)
Reviewers	Jeroen Beliën (Health-RI/Amsterdam UMC)
	Laura Whelan (RCSI)
	Pinar Alper (ELIXIR, LNDS)

Log of changes

Date	Mvm	Who	Description
28/10/2024	v1.0	Yaseen Jakhura (ELIXIR-Hub)	Deliverable submitted to EC
28/10/2024	v0.5	Marcos Casado Barbero (EMBL-EBI)	Update post GitHub SOP release





07/10/2024	v0.4	Yaseen Jakhura (ELIXIR-Hub)	Quality checked and sent to GDI-MB review
04/10/2024	v0.3	Marcos Casado Barbero (EMBL-EBI)	Addressed external reviews
12/09/2024	v0.2	Marcos Casado Barbero (EMBL-EBI)	First draft of document sent for external review
04/09/2024	v0.1	Marcos Casado Barbero (EMBL-EBI)	First draft of document sent for T4.3 internal review



Contents

1. Executive Summary	7
2. Contribution towards project outcomes	8
3. Methods	10
Deliverable Scope	10
Methodology	10
Survey and Landscape Analysis	10
Tools and Workflow	10
SOP Development and Alignment	11
GitHub Repository Management	12
4. Description of work accomplished	13
4.1 Overview of Activities	13
4.2 Detailed Description of Work Accomplished	13
4.2.1 Pre-Task Group Formation (Q3-Q4 2023)	13
4.2.2 Formation of the Task and Finish Group (January 2024)	14
4.2.3 Development of the GDI SOP Framework	14
4.2.4 Creation of Documentation and Platform Preparation	17
4.2.5 Roles and Responsibilities	18
4.2.6 SOP Life-Cycle	18
4.2.7 Initial SOP Development and Review	20
4.2.8 ZenHub Workspace Setup and Task Tracking	21
4.2.9 Creating DOIs	22
4.2.10 Alignment with GDI Project Goals	22
4.2.11 Review cycle of the SOPs	23
4.3 Interaction with Other Work Packages (WPs)	24
4.4 Challenges and Solutions	25
5. Results	26
6. Discussion	28
7. Conclusions & Impact	29
8. Next steps	31
9. References	32





List of figures

Figure 1. Summary of outputs (left) and activities that led to them (right) of T4.3.	13
Figure 2. Classification of SOP documents based on their maturity level.	16
Figure 3. Representation of SOP Templates and Instances across types.	17
Figure 4. Summary of the SOP life-cycle, presented at the Pillar II technical workshop in Brno (2024).	19
Figure 5. Examples of new SOP requests, listed in the GitHub repository as GitHub Issues.	21
Figure 6. Examples of new SOP requests, listed in the GDI SOP Zenhub board as GitHub tickets.	22
Figure 7. Example of what an automatic SOP Review reminder looks like.	24

Document Accessibility

Please note that some documents referenced in this deliverable are hosted within the GDI Google Drive space and may not be accessible to external readers. For access, please complete the registration form¹.

¹ <https://forms.gle/eqGpQtFdBmnUvhqP6>



Glossary

Abbreviation	Description
B1MG	Beyond 1 Million Genomes
BBMRI	Biobanking and Biomolecular Resources Research Infrastructure
DOI	Digital Object Identifier
EGA	European Genome-phenome Archive
FEGA	Federated EGA
GDI	Genomic Data Infrastructure
GH	GitHub
ISM	Information Service and Management
ISO	International Organization for Standardization
MB	Management Board
OC	Operations Committee
ORR	Organisational Roles and Responsibilities
SDPC	Security and Data Protection Committee
SLA	Service Level Agreement
SOP	Standard Operating Procedure
T4.3	GDI Task 4.3
TRL	Technology Readiness Level
WP	Work Package





1. Executive Summary

The deliverable D4.2 focuses on **creating and managing Standard Operating Procedures (SOPs) for the Genomic Data Infrastructure (GDI) network** to ensure standardised practices across diverse nodes. Its main objectives include developing a framework for SOP management, creating both European-level and Node-specific SOPs, and establishing a system for SOP request, development, review, approval and authorization.

Key achievements include:

- Formation of a **task group** involving representatives from nine GDI nodes to guide SOP development. Across 2024, these nodes covered both onboarding and deployment stages².
- Creation of a **flexible SOP framework**, supporting both common and node-specific needs.
- Establishment of a **GitHub-based platform** for SOP management, with automated workflows.
- **Completion of three SOPs**, with several others in development.

While progress has been made, some delays occurred due to the absence of key committees required for formal SOP approval. Looking forward, efforts will focus on expanding the SOP library, transitioning to a GDI-managed platform, aligning SOPs with standards and developing a monitoring framework for nodes, including self-assessments.

²docs.google.com/presentation/d/1Sre1njYQoadjOJ4GIXlhGxGbaD2GH54WzfhUruvI_Y/edit#slide=id.g2b7014b4bc8_0_370



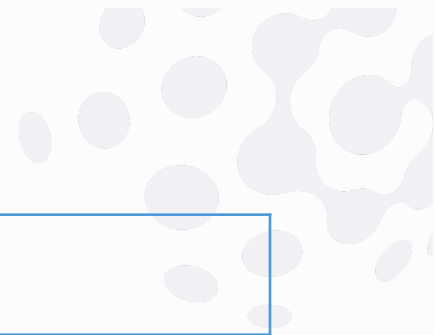


2. Contribution towards project outcomes

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

	Contributed
<p>Outcome 1</p> <p>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.</p>	No
<p>Outcome 2</p> <p>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.</p>	No
<p>Outcome 3</p> <p>Clear description of the roles and responsibilities related to personal data and privacy protection for humans and computers, applicable during the project lifetime and after its finalisation.</p>	Yes
<p>Outcome 4</p> <p>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 (e.g., IT and biotech companies), healthcare systems and public authorities at large.</p>	No
<p>Outcome 5</p>	No





<p>Sustained coordination mechanism for the GDI and the GoE multi-country project launched in the context of the 1+MG initiative.</p>	
<p>Outcome 6</p> <p>Communication strategy – to be designed and implemented at the European and national levels.</p>	No
<p>Outcome 7</p> <p>Capacity-building measures necessary to ensure the establishment, sustainable operation, and successful uptake of the infrastructure.</p>	No
<p>Outcome 8</p> <p>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.</p>	No





3. Methods

Deliverable Scope

This deliverable aims to create and manage Standard Operating Procedures (SOPs) within the Genomic Data Infrastructure (GDI) network³. It addresses the highly heterogeneous set of nodes within the GDI network by standardising and aligning SOPs. Its scope includes:

- **Documenting current operational practices** related to security, data protection, and node operations.
- **Developing a repository** of European-level SOP **instances** and Node-specific SOP **templates**.
- **Establishing a system for SOP management**, including request, development, review, approval, and dissemination.

This task is interconnected with all Work Packages (WPs) within Pillar II (WP3-6) to aid with the deployment and maintenance of national nodes' services. Among them, WP3 (Deployment of 1+MG National Nodes) in particular, by ensuring that the SOPs support the national nodes as they progress through the maturity assessment and, more importantly, that once operational, they are interoperable with each other. Furthermore, the other two GDI pillars are intertwined: Pillar I has contributed with advice revolving the governance of GDI SOPs and their development cycle; while the entirety of Pillar III will benefit from our efforts in developing a framework for managing SOPs, especially with regard to catering for GDI use-cases (WP7) through SOPs.

Methodology

Survey and Landscape Analysis

The initial phase involved forming a **"Task and Finish Group"** comprising representatives from nine GDI nodes (Belgium, Bulgaria, Finland, France, Luxembourg, Netherlands, Portugal, Slovenia, Sweden). The group's objective is to keep track and work towards Task 4.3 (T4.3) goals. Initially, we gathered information on nodes' current approaches to SOP management.

Tools and Workflow

The SOP life cycle involves multiple tools and workflows to ensure efficient management and communication across all involved actors:

- **Document Storage:** GDI's Google Drive suite⁴ is utilised to host development files and internal documentation. Furthermore, GDI's GitHub (GH) workspace⁵ is used as the public end of files and documentation.

³ <https://drive.google.com/file/d/1Fv--38mgOcz3l2wEy5eGhecZ7y-KxmAL/view>

⁴ <https://drive.google.com/drive/u/o/folders/17mC-CnAvtaEwZJnoWxGWBnem8OewuCb7>

⁵ <https://github.com/GenomicDataInfrastructure/standard-operating-procedures>



- **Communication:** Emails and mailing lists are the primary communication channels, supported by Slack⁶ for internal discussions.
- **Project Management:** A Zenhub board⁷ is used for tracking tasks and issues and is integrated with the GitHub repository.
- **Collaboration and Meetings:** Fortnightly Zoom meetings facilitate coordination among task members. Minutes⁸ are maintained to document efforts in the task consistently.
- **Surveying:** Google Forms was used to generate a survey for GDI SOP prioritisation⁹.

This comprehensive methodology ensures that the process of SOP development is accessible, transparent, easy to adapt, and capable of supporting the diverse needs of the GDI network.

SOP Development and Alignment

To address the variability in SOP management across nodes, the methodology focused on:

- **Representation and Collaboration:** Representatives from various nodes contributed to both minutes and collaborative sessions, ensuring diverse perspectives were considered.
- **SOP Categorisation:** SOPs were categorised into European-level and Node-specific SOPs. European-level SOPs instances are designed to be directly implementable by all nodes, while Node-specific SOP templates allow customisation based on individual node needs.
- **Template Standardisation:** A generic SOP template¹⁰ was created, outlining relevant sections of an SOP (e.g., scope, motivation, procedure...). This template serves as the foundation for all future GDI SOPs.
- **Documentation:** We created a suite of documentation to aid with the development of the SOPs, including a Charter¹¹, Information Service and Management (ISM)¹², Organisational Roles and Responsibilities (ORR)¹³, GitHub (GH) management¹⁴, SOP accessioning¹⁵, SOP Review Checklist¹⁶, SOP styling guide¹⁷, SOP index¹⁸, and main repository README¹⁹.

⁶ <https://gdi-elixir.slack.com/archives/Co6RJQJNHFC>

⁷ <https://app.zenhub.com/workspaces/t43-gdi-sops-667c1c5532726a00b93d51e4/board>

⁸ https://docs.google.com/document/d/1sWsbZle7ydmu8SZLnA00OnWoorO_9crBSgttnXTpygQ

⁹ <https://docs.google.com/forms/d/1y5KaJhzus9FdwSWfwq8BuXmXp5c0Sw1dU4rrtLZye3U>

¹⁰ github.com/GenomicDataInfrastructure/standard-operating-procedures/blob/main/docs/GDI-SOP_sop-template.md

¹¹ github.com/GenomicDataInfrastructure/standard-operating-procedures/blob/main/docs/GDI-SOP_charter.md

¹² github.com/GenomicDataInfrastructure/standard-operating-procedures/blob/main/docs/GDI-SOP_information-service-management.md

¹³ github.com/GenomicDataInfrastructure/standard-operating-procedures/blob/main/docs/GDI-SOP_organisational-roles-and-responsibilities.md

¹⁴ github.com/GenomicDataInfrastructure/standard-operating-procedures/blob/main/docs/GDI-SOP_github-management.md

¹⁵ github.com/GenomicDataInfrastructure/standard-operating-procedures/blob/main/docs/GDI-SOP_sop-accessioning.md

¹⁶ github.com/GenomicDataInfrastructure/standard-operating-procedures/blob/main/docs/GDI-SOP_review-checklist.md

¹⁷ github.com/GenomicDataInfrastructure/standard-operating-procedures/blob/main/docs/GDI-SOP_style-guide.md

¹⁸ github.com/GenomicDataInfrastructure/standard-operating-procedures/blob/main/sops/README.md

¹⁹ github.com/GenomicDataInfrastructure/standard-operating-procedures/blob/main/README.md





GitHub Repository Management

The main GH repository, **GenomicDataInfrastructure/standard-operating-procedures**²⁰, is structured to facilitate transparency, collaboration, and efficient management:

- **Structure:** The repository includes directories for documentation (docs), European-level SOPs instances (sops/european-level), node-specific SOP templates (sops/node-specific), GitHub workflows and templates (.github), and maintenance scripts (scripts).
- **Collaboration Tools:** GitHub features such as issues, pull requests, discussions, and releases are utilised to manage SOP development and updates. Automated workflows ensure compliance with repository standards and format.
- **Zenodo Integration:** The repository is linked with GDI's Zenodo account (ELIXIR Europe)²¹ to automatically mint DOIs for released SOP bundles, enhancing accessibility and citation.

In order to track the development and maintenance of GDI SOPs, we use Git²² through GitHub. It enables a branching approach to node implementations of GDI SOPs, where nodes can import and fork the main GH repository. This allows nodes to customise and manage their own clones of the repository while keeping a clear tracking of the remote.

²⁰ github.com/GenomicDataInfrastructure/standard-operating-procedures

²¹ zenodo.org/communities/gdi/records?q=&l=list&p=1&s=10&sort=newest

²² <https://git-scm.com/>





4. Description of work accomplished

4.1 Overview of Activities

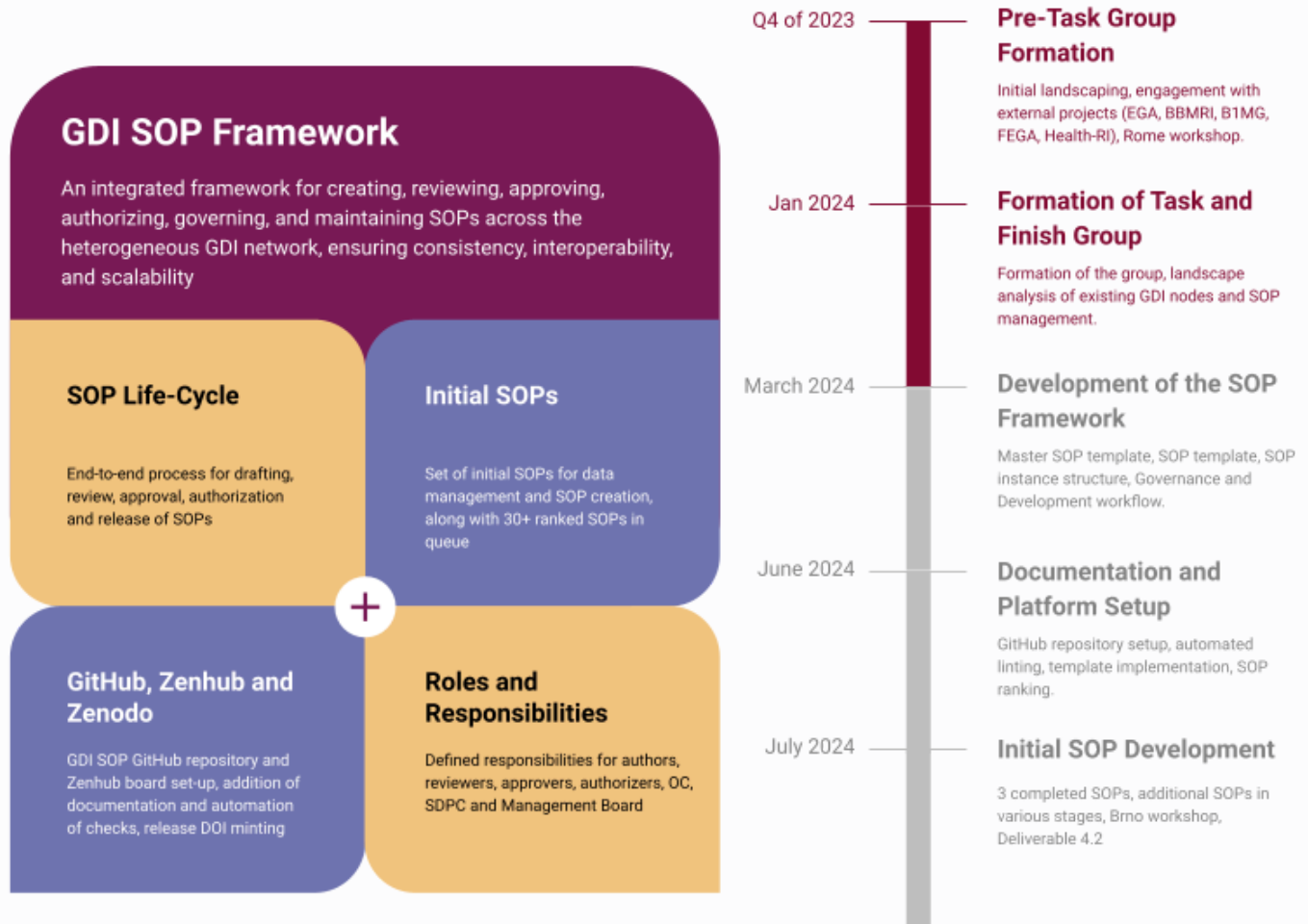


Figure 1. Summary of outputs (left) and activities that led to them (right) of T4.3.

4.2 Detailed Description of Work Accomplished

4.2.1 Pre-Task Group Formation (Q3-Q4 2023)

The groundwork for Task 4.3 began well before the official formation of the Task and Finish Group, during the third and fourth quarters of 2023. Recognizing the complexity and diversity of SOP management across GDI nodes, an **initial landscaping** activity was conducted. This activity aimed to identify the primary operational roles and topics that should be covered by a common set of SOPs.



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

Additionally, we sought to understand the existing and required resources necessary to support these SOPs and to establish a viable mechanism for their management.

To gather comprehensive insights, we engaged with established projects and teams, including the European Genome-phenome Archive (EGA)²³, Biobanking and Biomolecular Resources Research Infrastructure (BBMRI)²⁴, Beyond 1 Million Genomes (B1MG)²⁵, Health-RI²⁶, and Federated EGA (FEGA)²⁷ operation models. These engagements, along with close collaboration with Work Packages 3-6, provided critical input on **best practices and challenges in SOP management**.

The insights gained were further refined at the GDI Pillar II technical workshop²⁸ held in October 2024 in Rome, Italy, titled "Data Management Workshop on SOPs and Data Management Plan," during the GDI Pillar II technical workshop. This workshop was instrumental in **brainstorming potential SOPs and classifying them into four key topics**²⁹, which later served as the foundation for our SOP development framework. These four major categories are: (1) Data protection and security; (2) Data & metadata management; (3) infrastructure & software ; and (4) Helpdesk & operations.

4.2.2 Formation of the Task and Finish Group (January 2024)

In January 2024, the **Task and Finish Group** was officially established, with its first meeting held on January 15th. The group was formed to bring together representatives from various GDI nodes, tasked with addressing the diverse needs of the network concerning SOP development and management. The group's formation marked a significant milestone, as it formalised the collaborative efforts needed to standardise SOP practices across the GDI network.

Following its formation, the Task Group prioritised the systematic landscaping of existing GDI nodes to assess their **current approaches to SOP management**. This involved gathering detailed information on the tools and methods used by different nodes, revealing a wide range of practices, including the use of Git repositories, internal websites, and cloud-based solutions such as Google Drive and Microsoft SharePoint. The diversity in these approaches underscored the need for a flexible yet standardised framework for SOPs within the GDI.

4.2.3 Development of the GDI SOP Framework

Building on the insights from the pre-task activities and the landscaping analysis³⁰, the Task Group set out to develop a comprehensive **framework for GDI SOPs**. This framework was inspired in FitSM³¹

²³ <https://ega-archive.org/about/ega/>

²⁴ <https://www.bbmri-eric.eu/about/>

²⁵ <https://b1mg-project.eu/about/>

²⁶ <https://www.health-ri.nl/en>

²⁷ <https://ega-archive.org/about/projects-and-funders/federated-ega/>

²⁸ <https://docs.google.com/document/d/1BVQ20ap2LOzmonWWIdANWdVf-EBD9gw4G4wuavzEJwE>

²⁹ <https://docs.google.com/document/d/1RFI770aWwaoTinb1GvMDn5KbQU1NzNBuv85vcYiKxP8>

³⁰ https://docs.google.com/presentation/d/1UUxfPotlBllrAHDpVMv8NGr_OmZKsvnnwqymLE7vLD8

³¹ <https://www.fitsm.eu/>



by using existing SOP models of Health-RI³² to construct its main documentation. Furthermore, it was designed to accommodate the diversity of the GDI nodes while ensuring consistency and interoperability across the network. Its details can be found in the *docs/* directory within the GDI SOP GH repository³³, with especial interest in the Charter, ORR, ISM, GitHub Management, and SOP Accessioning documents.

Depending on the maturity level of an SOP document, we can split them into three main categories (see [Figure 2](#)):

- **Main SOP template.** This document outlines the minimum required structure and elements expected for any GDI SOP. It serves as the foundational template upon which all SOPs are built.
- **SOP template.** This is a more detailed version of the master SOP template, tailored to a specific SOP. It contains GDI-wide information and allows for customisation where needed. SOP templates are only used for Node-specific SOPs, as European-level SOPs are always released as SOP instances (see [Figure 3](#)).
- **SOP instance.** SOP Instance: This is the finalised, production-ready version of a specific SOP. It contains all the necessary details and is considered the fully implemented version of the SOP.

³² https://drive.google.com/drive/folders/1mW7mee9LFlcdJoBb_2IFGEecdju-j4rr

³³ <https://github.com/GenomicDataInfrastructure/standard-operating-procedures/tree/main/docs>





Figure 2. Classification of SOP documents based on their maturity level.

This approach also categorised SOPs into two main types: **European-level SOPs instances**, which are directly implementable across all nodes, and **Node-specific SOPs templates**, which allow for customisation based on individual node needs. Once a GDI node clones and adapts the node-specific SOP templates, they make them their own, transforming them to their node-specific implementations. This categorisation ensures that **while a common standard is maintained, there is also flexibility to accommodate the unique requirements of different nodes**. See [Figure 3](#) for a simplified visual representation of the differences between templates and instances, or refer to the main README's diagram³⁴ for a detailed version.

³⁴

github.com/GenomicDataInfrastructure/standard-operating-procedures/blob/main/README.md#summary-diagram



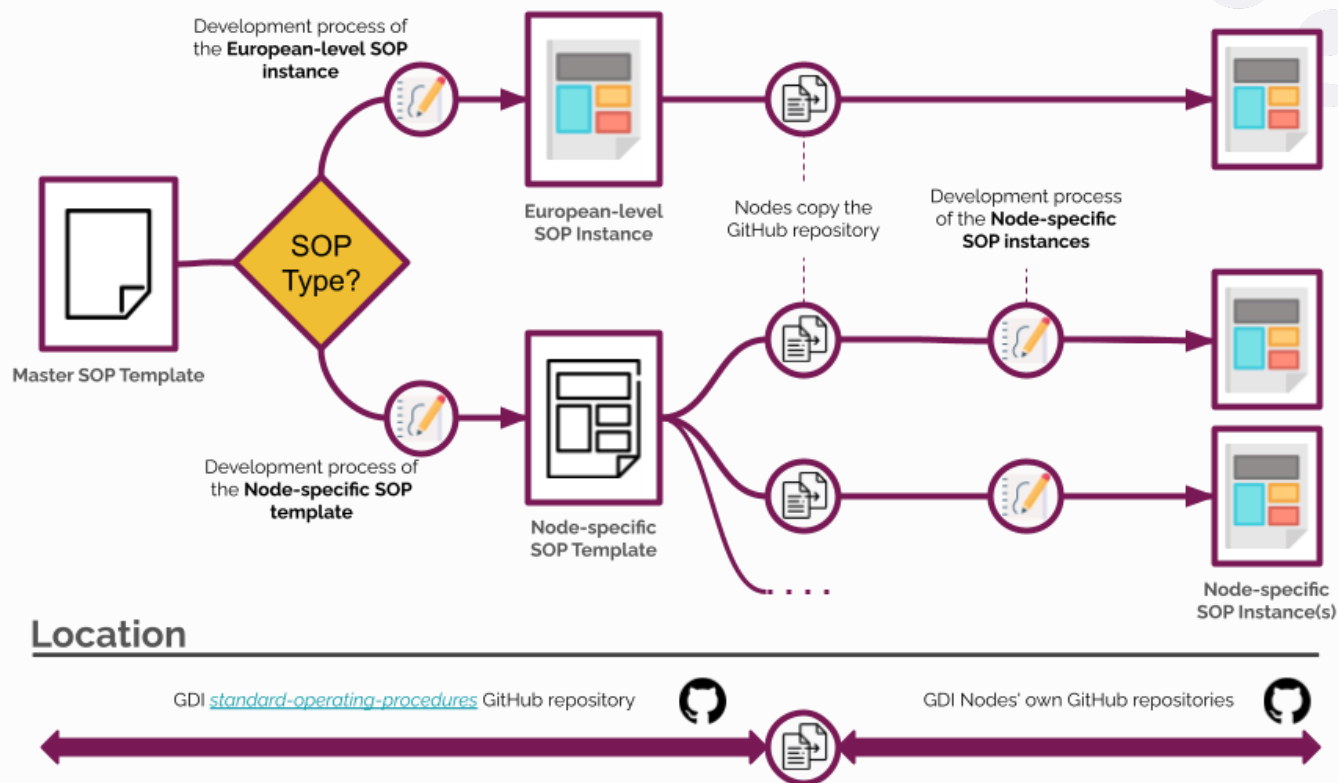


Figure 3. Representation of SOP Templates and Instances across types.

4.2.4 Creation of Documentation and Platform Preparation

To support the SOP development process, we prepared **extensive documentation**. See an extensive list of documentation at the SOP Development and Alignment section³⁵ of this deliverable. These documents provide clear guidelines and standards for **creation, management and maintenance** of GDI SOPs, ensuring that all contributors follow a consistent approach.

All this documentation was initially drafted in Google Drive, and later moved to the **main GitHub repository** (*GenomicDataInfrastructure/standard-operating-procedures*)³⁶. This repository was set up to serve as the **central platform for SOP management**, structured into specific directories for European-level SOPs instances, node-specific templates, supporting documentation and housekeeping scripts. To streamline the SOP development process, GitHub issue and Pull Request templates, along with python scripts and GH workflows were implemented. These helped to automate changes to the repository, as well as formatting checks and content creation. One of the main benefits being to ensure that all released GDI SOPs adhere to the minimal required structure and format.

³⁵ docs.google.com/document/d/1FjhsvmN3vmzUDHrJQir8zorDYx-pnqGL9UcMzoJ87_g/edit#heading=h.hzqxjvd6clus

³⁶ <https://github.com/GenomicDataInfrastructure/standard-operating-procedures>



Considerable attention was also given to defining the **SOP review and authorisation process**. Through iterative discussions, the roles and responsibilities of different stakeholders were clarified, ensuring a smooth and efficient workflow from SOP drafting to final approval.

4.2.5 Roles and Responsibilities

One of the first steps was to define the roles involved in the SOP development process. **Key roles** identified include SOP authors, reviewers, approvers, authorisers, trainers and repository maintainers (both at european and node-level)³⁷.

Transversal to these roles, we adapted the framework to fit well with the future **Operations Committee (OC) and Security and Data Protection Committee (SDPC)**, which will be responsible for the development and maintenance of GDI SOPs. Likewise, the workflow includes governing groups, like the **Management Board (MB)**. More information about these groups can be found in the GDI Milestone 10³⁸.

4.2.6 SOP Life-Cycle

The **SOP life-cycle** was carefully structured to guide SOPs from initial request to final approval and release³⁹.

To summarise (see [Figure 4](#)), this process begins with a GDI member submitting a request via a GitHub issue. The OC/SDPC then reviews the request, assigns authors, and oversees the drafting of the SOP using a standardised template. Following a review phase, the SOP is approved by members of the OC/SDPC and finally authorised by the Management Board. After this, the SOP is accessioned in the GitHub repository, where it is included in a formal release. This structured approach ensures transparency, accountability for all roles, and consistency in SOP development.

³⁷github.com/GenomicDataInfrastructure/standard-operating-procedures/blob/main/docs/GDI-SOP_organisational-roles-and-responsibilities.md

³⁸ <https://docs.google.com/document/d/1Vr5ChMWpkO8LDVigeGJv6LuegE4vjhHOJ1aISqWhAG4>

³⁹github.com/GenomicDataInfrastructure/standard-operating-procedures/blob/main/sops/european-level/GDI-SOP0007_SOP-template-creation.md



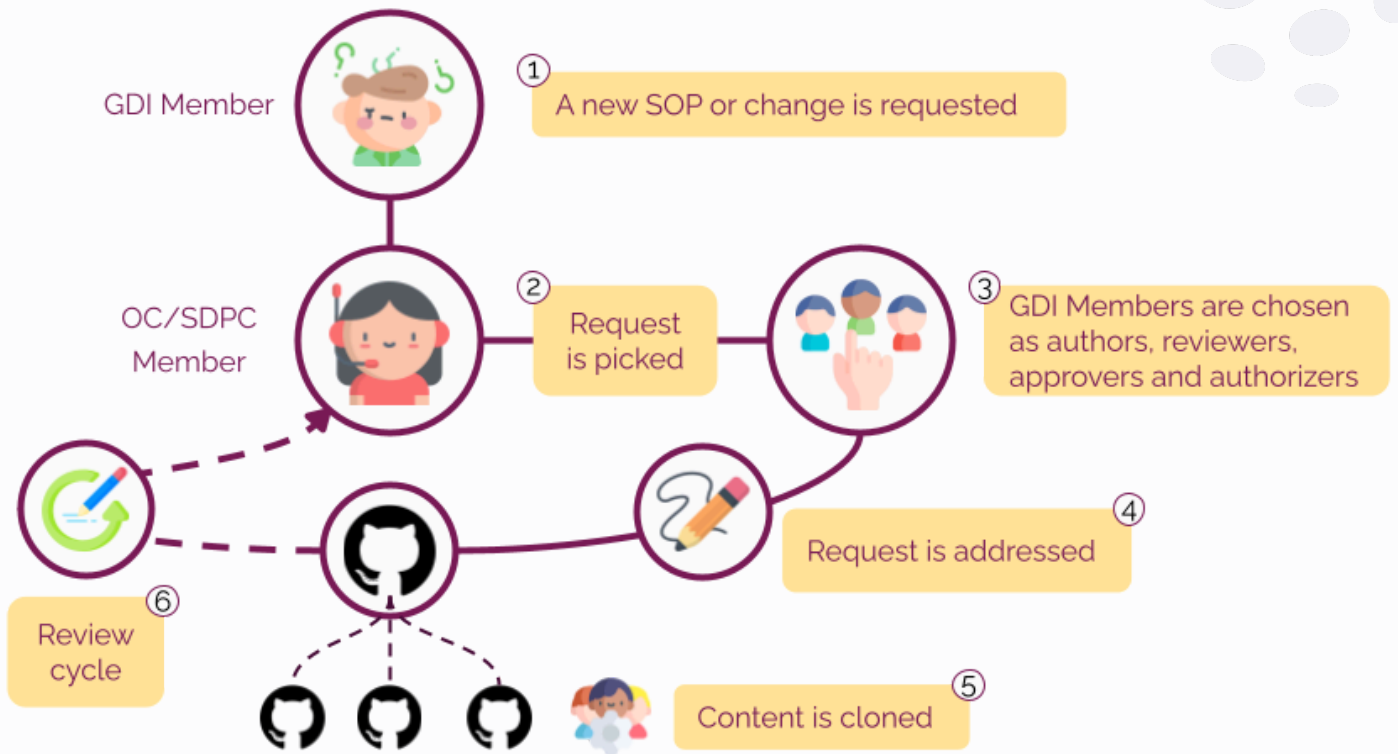


Figure 4. Summary of the SOP life-cycle, presented at the Pillar II technical workshop in Brno (2024).

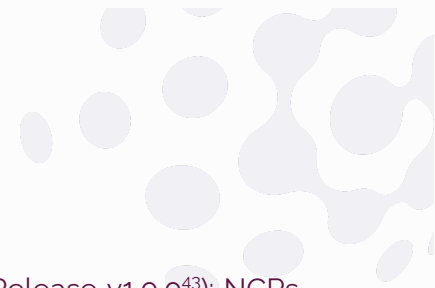
This approach was presented at the Pillar II Technical Workshop in Brno^{40,41}, Czech Republic, in September 2024, where it was well received. Find the **extensive SOP life-cycle** at the main README's diagram⁴².

⁴⁰ <https://docs.google.com/document/d/1QoUkxZDyuBy3m0yA22X-ygUR3nSSFy7ZttCzFGhf8lw>

⁴¹ https://docs.google.com/presentation/d/1R1EIL_LSuPuVH1lu_KTikYDskfOwLdMusgEo-Ob4fkU

⁴² github.com/GenomicDataInfrastructure/standard-operating-procedures/blob/main/README.md#summary-diagram





4.2.7 Initial SOP Development and Review

As of the submission of this deliverable, **three SOPs have been completed** (Release v1.0.0⁴³): NCPs veto EDIC decision⁴⁴, 1+MG DAC Recommendation Approval Process⁴⁵ and SOP Template creation⁴⁶. The full list of released SOPs can be found in the README file of the SOPs directory⁴⁷.

Beyond these initial SOPs, several additional SOPs are currently in various stages of development. These were identified as high-priority through a **ranking exercise conducted via a Google form**⁴⁸ distributed to GDI members (14 responses). This exercise ensured that the most critical SOPs were addressed first, reflecting the immediate needs of the GDI network.

To simulate real-world usage and refine our processes, members of the Task Group submitted **requests for new SOPs via GitHub issues**⁴⁹ (see [Figure 5](#)). From the moment a request is created, it propagates automatically to the task's internal ZenHub board (see [Figure 6](#)), where they are handled systematically by following one of the released SOPs: SOP Template creation⁵⁰. This hands-on approach provided valuable insights into potential bottlenecks, and helped the task group review collectively the SOP life-cycle.

⁴³github.com/GenomicDataInfrastructure/standard-operating-procedures/releases/tag/v1.0.0

⁴⁴github.com/GenomicDataInfrastructure/standard-operating-procedures/blob/main/sops/node-specific/GDI-SOP0002_NCPs-veto-EDIC-decision.md

⁴⁵github.com/GenomicDataInfrastructure/standard-operating-procedures/blob/main/sops/european-level/GDI-SOP0003_SOP-1+MG-DAC-Recommendation-Approval.md

⁴⁶github.com/GenomicDataInfrastructure/standard-operating-procedures/blob/main/sops/european-level/GDI-SOP0007_SOP-template-creation.md

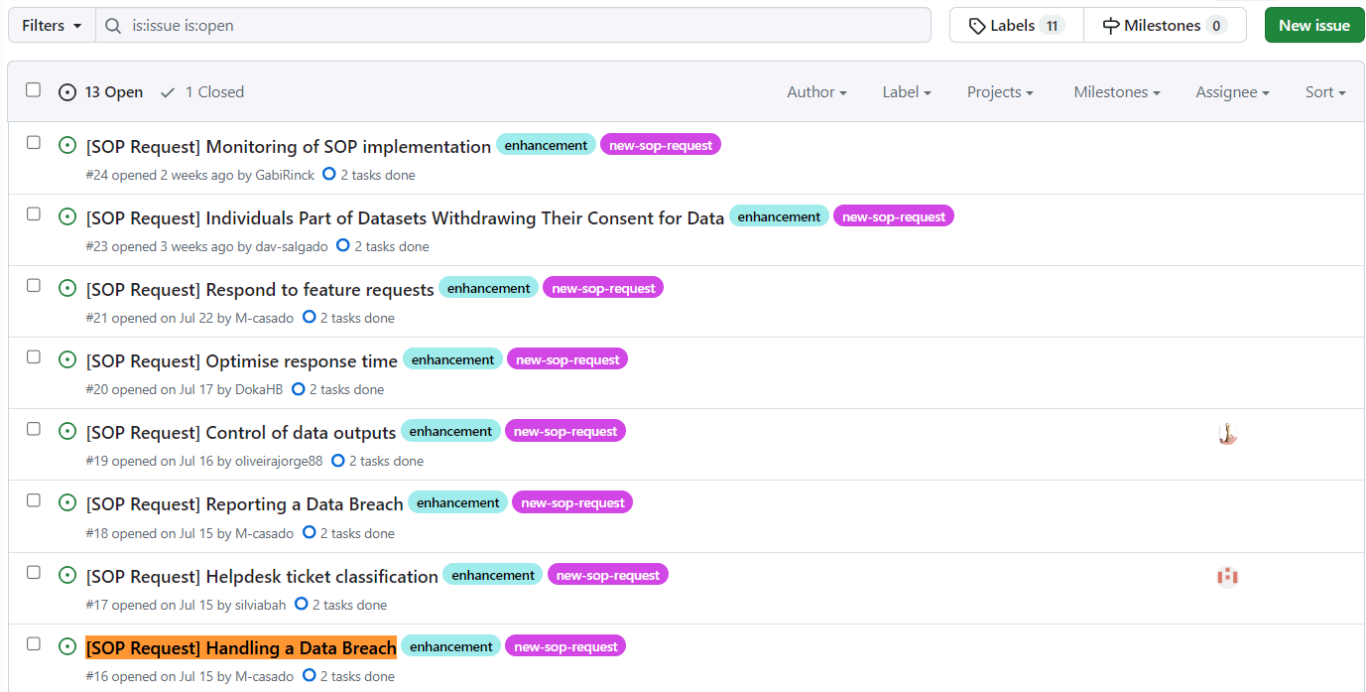
⁴⁷github.com/GenomicDataInfrastructure/standard-operating-procedures/tree/main/sops

⁴⁸<https://docs.google.com/forms/d/1y5KaJhzus9FdwSWfqw8BuXmXp5c0Sw1dU4rrtLZye3U>

⁴⁹github.com/GenomicDataInfrastructure/standard-operating-procedures/issues/new/choose

⁵⁰github.com/GenomicDataInfrastructure/standard-operating-procedures/blob/main/sops/european-level/GDI-SOP0007_SOP-template-creation.md





Filters Labels 11 Milestones 0 [New issue](#)

13 Open 1 Closed Author Label Projects Milestones Assignee Sort

- [SOP Request] Monitoring of SOP implementation enhancement new-sop-request
#24 opened 2 weeks ago by GabiRinck 2 tasks done
- [SOP Request] Individuals Part of Datasets Withdrawing Their Consent for Data enhancement new-sop-request
#23 opened 3 weeks ago by dav-salgado 2 tasks done
- [SOP Request] Respond to feature requests enhancement new-sop-request
#21 opened on Jul 22 by M-casado 2 tasks done
- [SOP Request] Optimise response time enhancement new-sop-request
#20 opened on Jul 17 by DokaHB 2 tasks done
- [SOP Request] Control of data outputs enhancement new-sop-request
#19 opened on Jul 16 by oliveirajorge88 2 tasks done
- [SOP Request] Reporting a Data Breach enhancement new-sop-request
#18 opened on Jul 15 by M-casado 2 tasks done
- [SOP Request] Helpdesk ticket classification enhancement new-sop-request
#17 opened on Jul 15 by silviabah 2 tasks done
- [SOP Request] Handling a Data Breach enhancement new-sop-request
#16 opened on Jul 15 by M-casado 2 tasks done

Figure 5. Examples of new SOP requests, listed in the GitHub repository as GitHub Issues.

4.2.8 ZenHub Workspace Setup and Task Tracking

To manage the workload and tasks associated with T4.3, a **ZenHub workspace**⁵¹ was established, integrated with the GitHub repository. The ZenHub board provides a **visual overview of all ongoing tasks**, including SOP requests, tracking their progress from initial request to final completion and further periodic reviews. This tool has been instrumental in keeping the Task Group organised and ensuring that all deliverables are on track.

ZenHub has also facilitated the coordination of activities across the Task Group, allowing members to efficiently manage their tasks and deadlines. The use of this tool has contributed to the smooth operation of T4.3 and the timely completion of key milestones. It will also be fundamental to the OC and SDPC to track any future SOP-related requests from GDI members (see [Figure 6](#)).

⁵¹ <https://app.zenhub.com/workspaces/t43-gdi-sops-667c1c5532726a00b93d51e4/board>



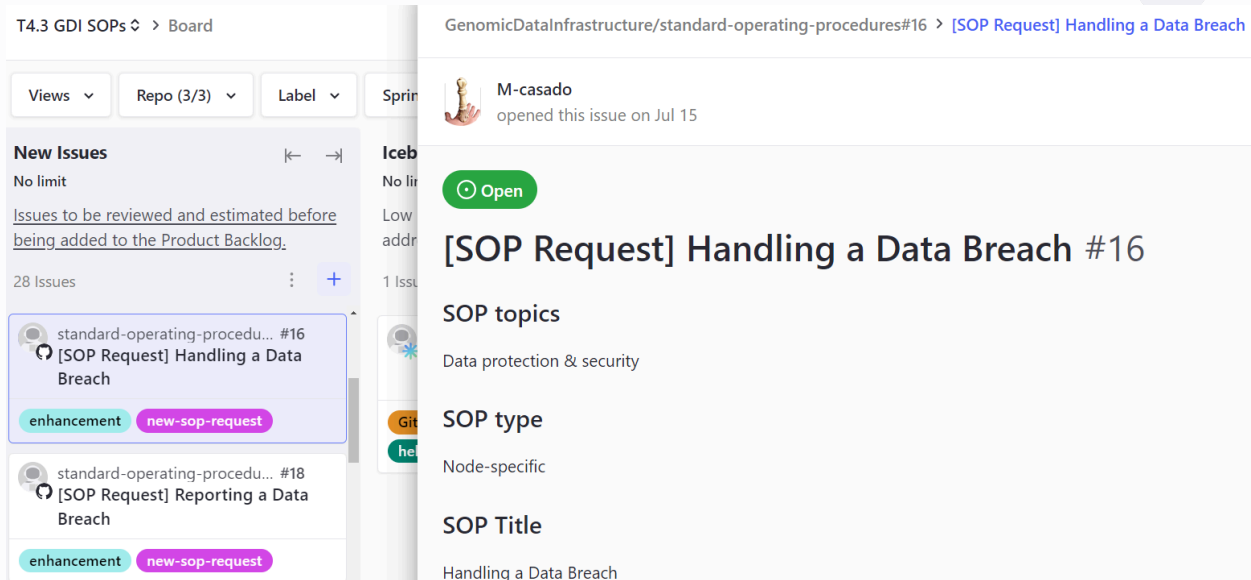


Figure 6. Examples of new SOP requests, listed in the GDI SOP Zenhub board as GitHub tickets.

4.2.9 Creating DOIs

To ensure that each SOP is uniquely identified and easily citable, we established a **linkage between GitHub releases and GDI's Zenodo**⁵². This setup allows for the automatic minting of Digital Object Identifiers (DOIs) for each SOP bundle released, ensuring that all SOPs are properly accessioned and accessible. The workflow consists in an automatic submission to GDI's Zenodo account (*ELIXIR Europe*)⁵³ once a release is created through the main GDI SOP GitHub repository. This submission includes all the content of the repository from the release, as a stable package, along with a DOI. See, for example, doi.org/10.5281/zenodo.14001554⁵⁴ for the SOP GitHub release v1.0.0⁵⁵.

This process not only facilitates the **tracking and citation of SOPs** but also enhances the **transparency and accessibility** of the SOPs developed under T4.3. The use of DOIs ensures that all SOPs are permanently accessible and can be easily, and unambiguously referenced by GDI members and external stakeholders.

4.2.10 Alignment with GDI Project Goals

Task 4.3 has been instrumental in aligning the GDI network's diverse SOP practices with the overarching goals of the GDI project. By developing a standardised framework for SOP creation,

⁵²github.com/GenomicDataInfrastructure/standard-operating-procedures/blob/main/docs/GDI-SOP_sop-accessioning.md#sop-releases-and-doi-minting

⁵³<https://zenodo.org/communities/gdi>

⁵⁴<https://doi.org/10.5281/zenodo.14001554>

⁵⁵github.com/GenomicDataInfrastructure/standard-operating-procedures/releases/tag/v1.0.0



management, and maintenance, this task directly supports GDI's mission of **fostering interoperability and collaboration across European genomic data infrastructures**.

The SOPs developed under T4.3 provide a solid framework to support all participating nodes, regardless of their level of maturity, to implement a common set of operational practices. These are essential to ensure consistency and interoperability across the network, crucial for the successful operations of such a large-scale infrastructure. Furthermore, the involvement of various GDI nodes in the SOP development process has fostered a sense of ownership and engagement, which is crucial for the long-term sustainability and success of the GDI project.

4.2.11 Review cycle of the SOPs

The **SOPs within the GDI network follow a structured review cycle** (see [Figure 4](#)) to ensure their continued relevance, clarity, and compliance. This review process can be triggered via different routes: automatic annual reviews and ad-hoc review requests initiated by GDI members.

1. **Annual Review Cycle.** An automated **GitHub workflow** (`review_reminder.yml`)⁵⁶ runs monthly and triggers a **Python script** (`check_sop_reviews.py`)⁵⁷ to check the last revision date recorded in each SOP's Document History. If a year has passed since the last review, a **GitHub issue is automatically generated** (see [Figure 7](#)). The issue lists required actions for the review process and acts as a clear reminder that the SOP is due for review.
2. **Ad-Hoc Review Requests.** Any GDI member can request an SOP review if they identify a need for updates or improvements. This is done by **submitting a GitHub issue** using the "SOP Content Change Suggestion" template⁵⁸ and adding the SOP-review label.

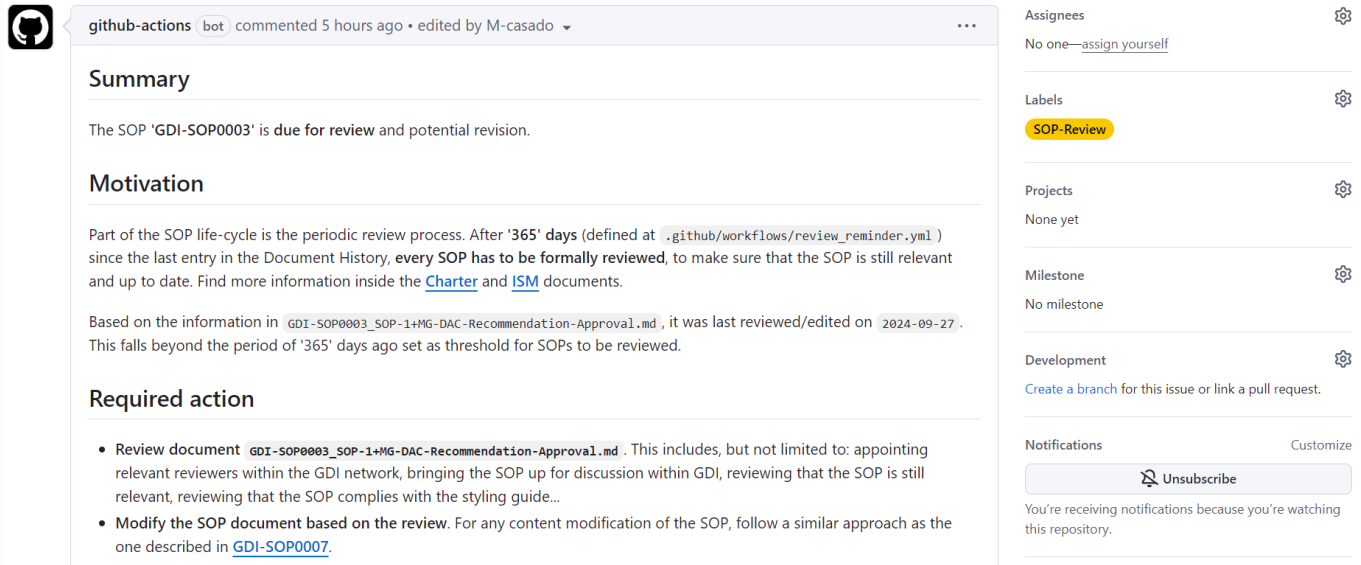
⁵⁶github.com/GenomicDataInfrastructure/standard-operating-procedures/blob/main/.github/workflows/review_reminder.yml

⁵⁷github.com/GenomicDataInfrastructure/standard-operating-procedures/blob/main/scripts/check_sop_reviews.py

⁵⁸github.com/GenomicDataInfrastructure/standard-operating-procedures/issues/new/choose



Open github-actions bot opened this issue 5 hours ago · 0 comments



The screenshot shows a GitHub issue page. At the top, it says 'github-actions bot commented 5 hours ago · edited by M-casado'. The issue title is 'SOP-Review'. The 'Summary' section states: 'The SOP 'GDI-SOP0003' is due for review and potential revision.' The 'Motivation' section explains the periodic review process, mentioning a '365' day threshold and linking to 'Charter' and 'ISM' documents. It also notes that the SOP was last reviewed/edited on '2024-09-27', which falls beyond the 365-day threshold. The 'Required action' section lists two bullet points: 'Review document GDI-SOP0003_SOP-1+MG-DAC-Recommendation-Approval.md' and 'Modify the SOP document based on the review'. On the right side, there are settings for Assignees (No one—assign yourself), Labels (SOP-Review), Projects (None yet), Milestone (No milestone), Development (Create a branch for this issue or link a pull request), and Notifications (Unsubscribe).

Figure 7. Example of what an automatic SOP Review reminder looks like.

Regardless of what triggered the GitHub issue, the OC/SDPC will assess the request and decide whether a formal review is necessary. If deemed justified, **the respective SOP will be reviewed**, with the process overseen by the OC and SDPC.

All review processes, whether automatic or ad-hoc, are tracked on the **ZenHub board**, with issues placed into the "SOP-reviews" pipeline for clarity. The outcomes of the reviews are recorded in the SOP's Document History. If content changes are needed, they follow the established development and approval stages outlined in GDI-SOP0007⁵⁹, ensuring consistency and governance.

4.3 Interaction with Other Work Packages (WPs)

Throughout Task 4.3 close **collaborations with other work packages** have been maintained. These extended to Pillar I, Pillar III and other WPs within Pillar II (WP3, WP4, WP5, and WP6) where, in particular, we can highlight the following interactions:

- Early in the project, **WP3 provided essential insights into the maturity levels and operational needs** of the national nodes, which guided the initial landscaping and SOP development efforts.

⁵⁹github.com/GenomicDataInfrastructure/standard-operating-procedures/blob/main/sops/european-level/GDI-SOP0007_SOP-template-creation.md



- **WP4's focus on European Level Operations** has been closely aligned with T4.3's objectives, particularly in ensuring that the developed SOPs support the broader goal of interoperability between nodes.
- **WP6's work on data management and governance** directly informed the content of several SOPs, ensuring that they align with the latest standards and practices in data protection and security.
- The **two GDI Pillar II Technical workshops**, in Rome and Brno, where members of all work packages pertaining to Pillar II, and some from the other two Pillars, participated in the SOP sessions.
- Overall feedback collected through the **GDI SOPs Priority Survey**⁶⁰ we created, where we obtained responses from WPs 3, 4, 6, 7, 8 and all three Pillars.
- We engaged in **conversations with Pillar I** regarding the overall governance of the GDI SOPs and their development process, where our approach was deemed suitable for GDI.

Additionally, T4.3 has also been guided by cross-WP outputs, like the Milestone 7⁶¹, which centred the first drafted SOPs around Data Governance⁶².

As detailed at [section 4.2.5](#) and Milestone 10⁶³, collaboration across GDI nodes will be maintained tightly in the future through the representation of these in the OC and SDPC.

4.4 Challenges and Solutions

The development of SOPs across a network as diverse as GDI presented several challenges. One of the primary challenges was the significant **variability in SOP content and management across nodes**. To address this, T4.3 adopted a flexible approach, developing a framework that can accommodate GDI node-specific requirements where appropriate, allowing for internal management of both European-level SOPs and node-specific SOPs. Thus, allowing the customisation of the latter while still adhering to a common standard, we ensure both flexibility and consistency.

Another challenge was the selection of a **suitable platform for SOP management**, given the sensitivity of the information and the diverse technical capabilities of the nodes. While GitHub was selected as the initial platform due to its robust features and widespread familiarity, discussions are ongoing about transitioning to a GDI-managed document system (e.g., GitLab) to better address concerns around data sovereignty and privacy. This potential transition is included in future planning, with T4.3's use-cases detailed⁶⁴. We aim to align our T4.3 use-cases with this solution as soon as it becomes available.

Coordination across multiple nodes and work packages also posed logistical challenges:

⁶⁰ <https://docs.google.com/forms/d/1y5KaJhzus9FdwSWfqw8BuXmXp5c0Sw1dU4rrtlZye3U/edit#responses>

⁶¹ <https://drive.google.com/file/d/1-hj3eM98qZmb5Ew5RjAFz7SyxTk-bySk/view>

⁶² https://docs.google.com/document/d/1a4x1vXReO-_J_ec_fl8iuA3XtE4mxrHg/edit#bookmark=id.qlapvv6guzor

⁶³ <https://docs.google.com/document/d/1Vr5ChMWpkO8LDVigeGJv6LuegE4vjhHOJ1alSqWhAG4>

⁶⁴ <https://gdi-elixir.slack.com/archives/C07B7S1HGG6/p1722345106551379>



- Maintaining **clear communication and timely progress**. To mitigate this, the Task Group secured fortnightly meetings, took clear minutes, set up a Slack channel, implemented a ZenHub workspace, and integrated it with the GitHub repository. This last pair, Github and Zenhub, has proven invaluable in tracking tasks, managing workflows, and ensuring transparency of the process so that all participants are aligned with the project's goals.
- **Governance and responsibilities**. With the sheer size of the GDI project, communicating emerging responsibilities and finding out their actors proved complicated. For example, which would be the groups in charge of approving and authorising SOPs. We approached this challenge by including in the conversation GDI members across vertical (e.g., coordination) and horizontal (e.g., Pillars) layers of GDI.

Lastly, the **accessioning and unique identification of SOPs** presented a technical challenge. The solution was to establish an accessioning guide⁶⁵ and a workflow linking GitHub releases with Zenodo, enabling the automatic creation of DOIs. This ensures that SOPs are not only easily accessible but also citable, providing a robust system for tracking and referencing SOPs across the GDI network.

5. Results

Major Achievements in D4.2:

- **Active Engagement and Collaboration Across Nodes:** Various GDI nodes participated in a dedicated manner to reach a broad consensus on developing SOPs within the network.
- **Development of an SOP Framework:** A framework for categorising SOPs into European-level instances and Node-specific templates was established, ensuring compliance with a unified standard while allowing for node-specific adaptations.
- **SOP Management:** Workflows were established for SOP development, review and authorisation. The entire SOP life-cycle⁶⁶ was meticulously designed to manage the process from initial request to final approval and release of SOPs and templates.
- **Documentation for SOP Development:** Documentation for SOP development was created, offering guidelines for consistent SOP creation and management. This documentation is available in a GitHub repository.
- **GitHub Repository Implementation:** A GitHub repository⁶⁷ was created as a central platform for developing, managing and maintaining SOPs and templates at European and node levels, tracking requests and providing documentation with guidelines.
- **ZenHub Workspace Integration:** A ZenHub workspace⁶⁸ integrated with the GitHub repository was developed to manage tasks related to SOPs efficiently.

⁶⁵github.com/GenomicDataInfrastructure/standard-operating-procedures/blob/main/docs/GDI-SOP_sop-accessioning.md

⁶⁶github.com/GenomicDataInfrastructure/standard-operating-procedures/blob/main/README.md#summary-diagram

⁶⁷github.com/GenomicDataInfrastructure/standard-operating-procedures

⁶⁸<https://app.zenhub.com/workspaces/t43-gdi-sops-667c1c5532726a00b93d51e4/board>



- **Identification of Key Roles:** Roles in the SOP development process were defined, including authors, reviewers, approvers, trainers and repository maintainers. The framework is intended to align with the requirements of committees involved in the review, approval and authorisation processes requirements, with oversight from governance bodies like the Management Board.
- **Initial suite of released SOPs:** Three SOPs were developed and later released^{69,70}, and several additional SOPs are currently in various stages of development.
- **Integration with Zenodo:** GitHub releases were linked with GDI's Zenodo⁷¹ instance to uniquely identify and cite each SOP or template. This setup enables automatic DOI generation upon release for proper cataloguing and accessibility.
- **Alignment with GDI Project Goals:** We successfully aligned our efforts with the overarching goals of the GDI project.

Communication, Outreach and Dissemination Activities:

- **Regular Working Group Meetings:** Meetings of the T4.3 working group were held bi-monthly.
- **Participation in WP4 and Pillar II Meetings:** We participated in WP4 and Pillar II meetings to address T4.3-related issues.
- **Feedback Solicitation:** Through close collaboration and interaction with other WPs of Pillar II, valuable feedback has been received on, for instance, how SOPs should be developed to align with the latest standards and practices in data protection and security and to meet interoperability requirements. From Pillar I, comprehensive feedback has been received on proposed processes for creation, review, approval and authorisation of SOP instances or templates. Feedback has also been received from cross-WP, derived from MS7, on how to write SOPs for Data Governance.
- **Workshops:** We actively participated in two GDI Pillar II Technical Workshops. During the first⁷², in Rome in October 2023, we led the "Data Management Workshop on SOPs and Data Management Plan" with WP6. At the second workshop⁷³, in Brno in September 2024, we presented the "GDI SOPs"⁷⁴.
- **Survey Conducted:** A survey⁷⁵ was conducted to gather valuable insights and feedback on SOPs, helping to prioritise the most relevant SOPs and focus efforts on delivering the most significant value to the consortium.

⁶⁹github.com/GenomicDataInfrastructure/standard-operating-procedures/releases/tag/v1.0.0

⁷⁰<http://doi.org/10.5281/zenodo.14001554>

⁷¹<https://zenodo.org/communities/gdi/about>

⁷² <https://docs.google.com/document/d/1BVQ20ap2LOzmonWWIdANWdVf-EBD9gw4G4wuavzEJwE>

⁷³ <https://docs.google.com/document/d/1Q0UkxZDyuBy3m0yA22X-ygUR3nSSFy7ZttCzFGhf8lw>

⁷⁴ https://docs.google.com/presentation/d/1R1EIL_LSuPuVH1lu_KTikYDskfOwLdMus9Eo-Ob4fkU

⁷⁵ <https://docs.google.com/forms/d/1y5KaJhzus9FdwSWfqw8BuXmXp5c0Sw1dU4rrtLZye3U/edit>





6. Discussion

Summary

1. **Delays in SOP Development, Approval, and Organisational Challenges:** The absence of key committees (OC and SDPC) has caused significant delays in the formal approval of SOPs, leading to bottlenecks in ensuring proper representation and feedback across all nodes.
2. **Challenges in Standardising Diverse Node Practices:** Aligning SOPs across nodes with varying practices and systems has proven difficult, which could lead to resistance to adopting standardised SOPs.
3. **Technical and Operational Constraints:** Uncertainty remains about transitioning from GitHub to a GDI-managed system.
4. **Mapping SOPs to Standards and Frameworks:** The need to align SOPs with recognized standards like International Organization for Standardization (ISO), Technical Readiness Levels (TRLs) or Service Level Agreements (SLAs) has been identified, but no clear guidelines or processes have been established.

Expanded discussion

1. **Delays in SOP Development, Approval, and Organisational Challenges:** A major blocker encountered in Task 4.3 has been the absence of key governance committees—the Operations Committee (OC) and Security and Data Protection Committee (SDPC)—which are essential for the formal approval of SOPs. Without these bodies, SOPs cannot undergo proper review and approval, causing delays in their adoption and implementation. The Task Group has attempted to mitigate these delays by relying on internal feedback and review, but this approach lacks the comprehensive representation needed from all GDI nodes. Furthermore, the Management Board, still relatively new to the SOP development process, has not fully established a structured procedure for final authorisations, which may further stall progress. As a result, while the initial set of SOPs are released, their development and implementation has been slower than anticipated. Furthermore, the process of appointing authors and reviewers for SOPs has proven challenging: SOPs require diverse expertise, and identifying suitable candidates from within the extensive GDI network is not straightforward. This may lead to delays in SOP drafting and review, as well as possible inconsistencies in the quality of SOPs. We were expecting to have SOP instances from nodes among T4.3, but as of now, it has been slowed down by the above bottlenecks.
2. **Challenges in Standardising Diverse Node Practices:** One of the fundamental challenges in developing SOPs has been the need to standardise practices across nodes with highly diverse operational environments and technological capabilities. While the Task Group has made efforts to create a flexible framework that accommodates a version-tracked set of SOPs, this approach may not satisfy the internal needs of all nodes. Some nodes may prefer to maintain their existing SOPs managing systems rather than adopting GDI's. Additionally, although Git-based SOP management was deemed as the best alternative, we recognise the



challenge of integrating it into nodes' existing local solutions. This could cause hesitation and resistance to fully adopting the new framework.

3. **Technical and Operational Constraints:** The use of GitHub as the primary platform for SOP management, while effective for collaboration and transparency, has limitations. Concerns about data sovereignty and privacy have been raised, particularly with the management of sensitive node-specific SOPs. This has led to discussions about transitioning to a GDI-managed system, which could offer better control over data privacy and security. However, this transition has been delayed due to dependencies on other work packages and ongoing uncertainty about the technical feasibility, timeline, and resources required for implementation. Until these issues are resolved, complete EU-based privacy of the SOP management system cannot be realised. Nevertheless, we expect that nodes' sensitive implementations will require time, which may give precious time for the GDI-managed system to be set.
4. **Mapping SOPs to Standards and Frameworks:** Another identified gap is the lack of mapping SOPs to recognized standards such as ISO, TRLs, or SLAs. Aligning SOPs with these standards is crucial for ensuring their broader applicability, compliance, and acceptance within the GDI network and beyond. However, no clear guidelines or processes for this alignment have been established, which could limit the effectiveness and adoption of the SOPs. Developing a standardised approach to mapping SOPs to these frameworks is needed to enhance the credibility and utility of the SOPs. These mappings would benefit both users, who would know what to expect, and service providers, who would be backed by clear expectations and maturity levels.

Closing Remarks

Despite these challenges, Task 4.3 has made significant strides in establishing a **robust foundation for SOP development within the GDI network**. While there are hurdles to overcome, such as the absence of key committees and the technical complexities of system transitions, the progress made so far has laid a strong groundwork for the future. The creation of a flexible SOP framework, the involvement of diverse nodes, and the ongoing efforts to integrate the system across GDI will ultimately streamline operations and enhance interoperability. Although there is work still to be done, the outcomes of Task 4.3 are already proving valuable and will undoubtedly serve as a crucial resource in the continued development and success of the GDI project.

7. Conclusions & Impact

Summary

1. **Comprehensive SOP Framework Developed:** Successfully established a standardised framework for SOP creation, management, and maintenance within the GDI network.
2. **Documentation and Platform Preparation:** Extensive documentation was created, and a robust GitHub-based platform set up for managing SOPs.



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

3. **Engagement and Collaboration Across Nodes:** Active participation from multiple GDI nodes, leading to SOPs that reflect a broad consensus and cater to diverse needs.
4. **Initial set of released SOPs:** Three SOPs were released and are ready for nodes' implementation, with more in development.
5. **Automation and Workflow Integration:** Implemented workflows for automated SOP formatting, accessioning, and DOI minting, enhancing efficiency and traceability.

Expanded discussion

1. **Comprehensive SOP Framework Developed.** A key achievement of T4.3 has been the creation of a standardised framework that accommodates the diverse needs of the GDI nodes. This SOP framework is designed around two categories of SOPs: European-level instances and Node-specific templates. This allows flexibility while maintaining a common standard. The development of this framework is foundational for ensuring that all nodes, regardless of their maturity level, can implement SOPs that are both consistent and adaptable to their unique contexts.
2. **Documentation and Platform Preparation.** The extensive documentation prepared under T4.3 provides clear guidelines and standards for SOP development and management. The setup of the GitHub repository as the central platform for SOP management, complemented by automation tools and workflows, streamlines the process of SOP creation, review, and maintenance. This preparation not only supports the immediate needs of SOP development but also lays the groundwork for future scalability and sustainability.
3. **Engagement and Collaboration Across Nodes.** The active involvement of various GDI nodes in the SOP development process has been crucial in creating SOPs that are relevant and practical across the network. By incorporating feedback and insights from a diverse set of stakeholders, the SOPs developed under T4.3 are more likely to be adopted and effectively implemented by all nodes. This collaborative approach fosters a sense of ownership among the nodes, which is essential for the long-term success of the SOP development and implementation. Node representation will be maintained through the OC and SDPC in the future.
4. **Initial set of released SOPs.** The progress made on the initial set of SOPs⁷⁶, with three being already released⁷⁷, demonstrates the viability of the frameworks established under T4.3. These ranked SOPs address high-priority areas identified by GDI members, and will be the foundation for further SOP development.
5. **Automation and Workflow Integration.** The integration of automated workflows for SOP formatting, accessioning, and DOI minting represents a significant advancement in the efficiency and traceability of SOP management. These tools not only reduce the manual workload but also ensure that SOPs are consistently formatted⁷⁸, easily accessible⁷⁹ and

⁷⁶github.com/GenomicDataInfrastructure/standard-operating-procedures/tree/main/sops

⁷⁷github.com/GenomicDataInfrastructure/standard-operating-procedures/releases/tag/v1.0.0

⁷⁸github.com/GenomicDataInfrastructure/standard-operating-procedures/blob/main/.github/workflows/lint_sops.yml

⁷⁹github.com/GenomicDataInfrastructure/standard-operating-procedures/blob/main/.github/workflows/compare_index.yml



relevant⁸⁰. The use of Zenodo for DOI minting further enhances the curation and citation of SOPs, making them more accessible and citable within the scientific community.

8. Next steps

Summary

1. **Address representation:** Escalate the ongoing challenge of missing representation from key groups, namely the OC and SDPC.
2. **Expansion of SOP Library:** Continue developing and refining SOPs to build a comprehensive repository.
3. **Releases of SOP's GitHub Repository:** Plan and execute further releases of the GitHub repository, each time a critical mass of SOPs is achieved.
4. **Implementation of Self-Assessments:** Develop and deploy self-assessment tools for operational GDI nodes to evaluate SOP compliance and maturity.
5. **Transition to GDI-Managed Document Storage:** Prepare for the potential transition from GitHub to a GDI-managed document storage solution, such as GitLab.
6. **Mapping to Standards:** approach further to FitSM, map SOPs to relevant ISO standards, and align them with Technical Readiness Levels (TRLs) and Service Level Agreements (SLAs).
7. **Monitoring Framework Development:** Establish a monitoring framework to ensure ongoing compliance with GDI SOPs across nodes.

Expanded Next Steps

Although Milestone 10⁸¹, delivered in April 2023, outlined the governance structure for GDI operations, the Operations Committee (OC) and Security and Data Protection Committee (SDPC) — critical for Task 4.3 — have yet to be formed. These committees are essential for ensuring proper representation of GDI nodes in SOP development and maintenance. Despite our repeated requests to establish these groups in various Pillar II forums, this issue remains unresolved. Escalating this matter is now necessary, and we are committed to keeping it on the agenda in GDI-wide discussions until it is addressed.

In the remaining time of the project, a primary focus will be on **expanding the SOP library**. This involves not only writing new SOPs but also revisiting and refining the existing framework as necessary to ensure it remains relevant and effective.

As the repository continues to grow with more SOPs, we will aim to make further **formal releases of the GitHub repository**. These will serve as public milestones, making a consolidated set of SOPs available to all GDI nodes and external stakeholders. Once SOPs are released, GDI nodes will be able to adapt them, if applicable, to their needs, creating their set of Node-specific SOP instances.

⁸⁰github.com/GenomicDataInfrastructure/standard-operating-procedures/blob/main/scripts/check_sop_reviews.py

⁸¹<https://docs.google.com/document/d/1Vr5ChMWpkO8LDVigeGJv6LuegE4vjhHOJ1aISqWhAG4>



Developing **self-assessment tools** for operational GDI nodes is another key objective. These tools will allow nodes to assess their maturity and compliance with the established SOPs, providing valuable feedback that can guide further improvements and support. This will be in accordance with the future Milestones 13⁸² and 14⁸³ regarding node compliance assessments.

Given the non-EU-based nature of GitHub, we are preparing for a potential transition to a **GDI-managed document storage solution**, such as a self-hosted GitLab. This transition would enhance data sovereignty and security for GDI nodes, particularly those handling sensitive information. With this change, we hope to incentivize the adoption of GDI's SOP approach from the nodes.

An important next step is to **reconcile further the proposed framework with FitSM** (i.e., fill in current gaps), and **map the developed SOPs to relevant standards** (e.g., ISO 9001 for quality management) and Technical Readiness Levels (TRLs). Aligning SOPs with these standards will enhance their credibility and applicability. Additionally, efforts will be made to align the SOPs with Service Level Agreements (SLAs), further integrating them into the operational frameworks of GDI nodes.

To ensure ongoing compliance with the GDI SOPs, we will begin developing a **monitoring framework**. This framework will allow for continuous oversight of SOP implementation across the network, helping to identify areas of non-compliance and enabling support for nodes that may require assistance.

9. References

Some icons of [Figure 2](#), [Figure 3](#) and [Figure 4](#) were made by Freepik, and obtained through Flaticon.com.

⁸² docs.google.com/presentation/d/1RMiG6S6kjSzhW5TQoJdxa2BvRALmDN-piVvVwVK-y4Y/edit#slide=id.p7

⁸³ docs.google.com/presentation/d/1-ridDELcAjSrtw8qZDSIZaNVOCJ-cWqHYJofxbnhln4/edit#slide=id.p7

