

Progress Report

2nd Quarter of 2025

GLIM-BioData **Gateway for Living Data Management**

Within the scope of **PNCADAI - Programa Nacional de Ciência Aberta e Dados Abertos de Investigação**

Part of **Medida RE-C05-i08 - Ciência Mais Digital**
do **PRR - Programa de Recuperação e Resiliência**

May to August 2025

Date: 10/09/2025

Prepared by: Luciana Peixoto

Revised: Inês Chaves



Index

1. Progress Monitoring	3
2. Overall Execution	8
3. Additional Relevant Considerations	9
4. Next Steps	12

1. Progress Monitoring

The implementation of the Research Data Management Centre under the GLIM-BioData project is designed to consolidate BioData.pt as the national reference for the management and valorisation of life and health sciences data. By fostering data sharing and reuse, the Centre promotes the adoption of best practices and guidelines that enhance the quality and impact of research outputs.

Two strategic use cases are at the core of this implementation. The first addresses biodiversity, focusing on the annotation and public dissemination of datasets to monitor the spread of pathogenic bacteria in aquatic environments and their potential health risks. The second concerns sensitive genomic data, with the creation of a FEGA Portugal helpdesk, support for data curation, and the development of legal and ethical frameworks that guarantee compliance with GDPR.

Capacity building is also a central goal. Training initiatives in bioinformatics, data management, and data stewardship, delivered by ELIXIR-certified trainers, aim to shape a new generation of highly skilled professionals capable of addressing the challenges of data-intensive life sciences. In parallel, tailored processes are being developed for managing both sensitive and non-sensitive data, ensuring alignment with FAIR principles and the broader vision of Open Science.

Finally, the long-term sustainability of the Centre is pursued through continuous impact monitoring, based on KPIs, the RI-PATHS methodology, and CoARA principles. Strengthened collaboration at the national level and active engagement with European infrastructures such as the GDI Consortium, ELIXIR, FEGA, and EOSC ensure the interoperability and integration required for the Centre to play a lasting role in the European RDM ecosystem.

1.1. Deliverables

Below is an overview of all deliverables and milestones, as well as the implementation concerning the contractual timetable (Table 1).

Table 1 - Application Timetable vs Execution – Deliverables and Milestones

Activity	Deliverable /Milestone	Estimated schedule	% Execution as of the report date	
Detailed Work Plan	D1.1 - GLIM-BioData Centre Work Plan	9 May 2025	Expected	100
			Actual	100
Implementation and adoption of an institutional policy for research data management and	D1.5 - Definition of the institutional policy for RDM and sharing	Up to 9 months after the contract start date (October 2025)	Expected	80
			Actual	70

sharing by the centre's institution(s)	D1.6 - Implementation and adoption of the institutional policy for RDM and sharing by the GLIM-BioData participant institution(s)	By the end of the contract	Expected	0
			Actual	0
	M1.1 - Roadshow about the GLIM-BioData Centre and the integration in the Consortium-GDI	By the end of the contract	Expected	80
			Actual	40
	M1.2 - Roadshow about human-sensitive data and alignment with FEGA	By the end of the contract	Expected	80
			Actual	40
	M3.1 - Report on outreach activities with FEGA stakeholders	By the end of the contract	Expected	0
			Actual	0
	M3.3 - Guidelines to address Legal and Ethics Issues in RI	By the end of the contract	Expected	50
			Actual	10
Availability of data sets with reuse potential	D2.1 - Curation and datasets availability plan for Non-sensitive Life Data	Up to 9 months after the contract start date (October 2025)	Expected	50
			Actual	50
	D3.1 - Curation and datasets availability plan for Sensitive Life and Health Data	Up to 9 months after the contract start date (September 2025)	Expected	75
			Actual	40
	D2.2 - Publication of datasets for Non-sensitive Life Data	By the end of the contract	Expected	0
			Actual	0
	D3.2 - Publication of datasets for Sensitive Life and Health Data	December 2025	Expected	0
			Actual	0
	M2.2 - DMPortal integration in POLEN	July 2025	Expected	100
			Actual	50
	M2.1 - Open-source electronic lab notebook (eLabFTW) implementation	November 2025	Expected	70
			Actual	100
	M3.2 - Standard Operating Procedures for FEGA Portugal Helpdesk	December 2025	Expected	50
			Actual	50
Training sessions on good research data management	D4.1 - Training sessions on best practices in RDM for	3 and 4 July 2025 in CiiMAR	Expected	100

practices according to the plan proposed by the proposing entities	Life Data - "Ready4Intensive" Course - GLIM 1 st Edition		Actual	100
	D4.2 - Training session on best practices in RDM for Life Data - "Ready4Intensive" Course - GLIM 2 nd edition	October 2025 at UCoimbra	Expected	20
			Actual	20
	D4.3 - Training Session for Data Stewards - 1 st session	7 July 2025 Online	Expected	100
			Actual	100
	D4.4 - Training Session for Data Stewards - 2 nd session	17 November 2025 UCIBIO	Expected	20
			Actual	20
	M4.1 - Training on accessing, processing, and curating datasets	22 September 2025 Online	Expected	50
			Actual	20
	M4.2 - Training in eLABFTW	28 April 2025 in CiiMAR in collaboration with Centre GDI FAIRway	Expected	100
			Actual	100
Integration and collaboration in a national network of data stewards	M4.3 - Training in eLABFTW	22 September in CCMAR	Expected	50
			Actual	50
	Completion of 100% of the sessions in the proposed plan	By the end of the contract	Expected	40
			Actual	40
Participation in the Consortium's coordination activities	D5.1 - Integration into the national data stewards network to be operated by the Consortium	By the end of the contract	Expected	100
			Actual	100
	M5.3 -Establishment of a core group of data stewards for sensitive life and health data	November 2025	Expected	80
			Actual	80
Providing a public version of the centre's quarterly progress reports	D1.2 - 1 st Progress report	13 May 2025	Expected	100
			Actual	100

	D1.3 - 2 nd Progress report	10 September 2025	Expected	100
			Actual	100
	D1.4 - Final report	By the end of the contract	Expected	0
			Actual	0
	Information made available every four months	By the end of the contract	Expected	67
			Actual	67
Centre GDI Impact	M5.1 - Assessment of BioData.pt node and communities services	June 2025	Expected	100
			Actual	100
	M5.2 - Participation in projects promoted by European infrastructures (ESFRI) and/or funded by European funds	By the end of the contract	Expected	75
			Actual	90
	M5.4 - Key Performance Indicators definition and monitorization	By the end of the contract	Expected	50
			Atual	50

As part of the activities carried out during this reporting period, four deliverables and two milestones were completed in alignment with the established timeline. These include deliverables D1.3 - 2nd Progress report, D4.1 - Training sessions on best practices in RDM for Life Data - “Ready4Intensive” Courses - GLIM 1st edition 2025, at CiiMAR, D4.3 - Training Sessions for Data Stewards - 1st session (Online), and D5.1 - Integration into the national data stewards network to be operated by the Consortium, and two milestones, M2.1 - Open-source electronic lab notebook (eLabFTW) implementation and M5.1 - Assessment of BioData.pt node and communities services.

1.2. Explanation of Progress and Possible Deviations

During the current reporting period, significant progress was made, with the completion of four key deliverables and two milestones, all in line with the defined schedule. A detailed account of project progress, including any deviations from the initial plan, is provided below, organised by work packages (WP).

WP1 - Coordination and Management

The deliverable D1.3 - 2nd Progress Report will be achieved through the submission of the present document by 10 September, thus ensuring the timely delivery of the project’s mandatory reporting obligations. This deliverable provides a comprehensive overview of the progress made, challenges encountered, and mitigation measures adopted, contributing to transparent monitoring of the project’s implementation.

At the time of this reporting period, the deliverable D1.5 - Definition of the institutional policy for RDM and sharing was expected to be 80% completed. The deliverable will be fully completed in time, by October 2025.

Regarding the deliverable D1.7 - Participation in the Consortium-GDI coordination activities, three General Assemblies were foreseen until the end of the project, of which [BioData.pt](https://biodata.pt) has already participated in the two held so far.

Taking into consideration milestones M1.1 - Roadshow about the GLIM-BioData Centre and the integration in the Consortium-GDI and M1.2 - Roadshow about human-sensitive data and alignment with FEAGA, four Roadshows serving the purposes of the two aforementioned milestones have been successfully conducted to promote the GLIM-BioData Centre. These sessions took place at UCIBIO (Porto and Lisbon), CIIMAR, and the University of Coimbra, which was anticipated, engaging researchers and institutional representatives in discussions about research data management practices and the role of GLIM-BioData in the national and European RDM landscape. Due to institutional constraints, the roadshows to GiMM, NOVA Medical School, FCUL, CEBAL, iBET and ITQB could not be carried out; however, they are in the scheduling phase with the respective institutions after the summer break.

WP2 - Management of Non-sensitive Life Data for Sharing and Reuse

To accomplish the deliverable D2.1 - Curation and datasets availability plan for Non-sensitive Life Data, EMO-BON data is used to monitor bacterial pathogens using machine learning approaches. To date, metadata has been compiled on sample acquisition, including information such as water temperature, season and an anthropogenic index reflecting human impact. We are also working to apply multiple instance learning (MIL) to raw WGS reads. So far, the necessary metadata and sequence datasets have been prepared. Initial analyses using GO categories and baseline models (Multilayer Perceptron and XGBoost) produced limited results. For the MIL approach, pipelines that link EMO-BON metadata to the ENA Browser have been developed. This enables the retrieval and preprocessing of FASTQ files for model training.

The milestone M2.1 - Open-source electronic lab notebook (eLabFTW) implementation sought to deploy an instance of eLabFTW to be made available to all GLIM-BioData institutions. The platform is available at <https://elab.biodata.pt> and allows everyone with an institutional login to use the platform through Life Science Login. The aim of this instance is not to be used in research, but rather to support training programmes and provide a sandbox for researchers and data stewards to get acquainted with the platform and test its features.

The milestone M2.2 - DMPortal integration in POLEN seeks ways to integrate the BioData.pt Data Management Portal (DMPortal) in Repositório POLEN, managed by FCCN. This will be achieved by meeting with the POLEN team at FCCN to study the best way of assuring this integration. The GLIM-BioData team has responded positively to the survey provided to GDI centres in relation to being added to the Repositório POLEN pilot. Further meetings are expected to take place this month.

WP3 - Management of Sensitive Life and Health Data for Sharing and Reuse

The deliverable D3.1 - Curation and datasets availability plan for Sensitive Life and Health Data is due at the end of September and will consist of a Data Management Plan (DMP) about a dataset of Sensitive Data provided by GiMM. Due to the need for additional authorisations and ethical clearances towards the release and use of sensitive data, a slight delay might occur, but only until the 9th month of the project. The same dataset will be used on the deliverable D3.2, scheduled for December 2025.

Towards the milestone M3.2 - Standard Operating Procedures for FEGA Portugal Helpdesk, our team organised a “SOPathon” - hackathon dedicated to the writing or adaptation of SOPs - that has already produced two new SOPs for the FEGA Portugal Helpdesk, in alignment with the ones proposed by the Central EGA Operational Committee. Another SOPathon will be held, creating another batch of SOPs and concluding this milestone by December 2025.

Regarding the milestone M3.3 - Guidelines to address Legal and Ethics Issues in RI, we are reinforcing our team by proceeding with the contracting of a service of a Data Protection Officer (DPO) that will support the outputs of this milestone. The conclusion of M3.3 remains scheduled for the end of the project.

WP4 - Training and Capacity Building

Regarding the deliverable D4.1, which corresponds to the Training sessions on best practices in RDM for Life Data - “Ready4Intensive” Course - GLIM-1st Edition 2025, it was successfully completed by the occurrence of the training, held on the 3rd and 4th of July, 2025, at CIIMAR (Porto, Portugal) (<https://shorturl.at/Ge20h>; Evidence Report Deliverable D4.1 - Training sessions on best practices in RDM for Life Data - “Ready4Intensive” Course - 1st Edition).

The deliverable D4.3, which comprised the Training Sessions for Data Stewards 2025, was completed through the successful delivery of the first session, which took place online on July 7, 2025 (<https://shorturl.at/wCln4>; Evidence Report Deliverable D4.3 - Training Sessions for Data Stewards - 1st session).

The deliverable D4.4 - Training Session for Data Stewards - 2nd session has been divided into two parts. The first part will take place online on 2 October, while the second is scheduled as an on-site event at UCIBIO in Porto on 17 November. Since the programme has already been finalised and the registration process has already been opened, this deliverable is currently considered 20% completed, with the remaining progress to be achieved upon the execution of both sessions.

Milestone M4.1 - Training on accessing, processing, and curating datasets was originally scheduled for September 22nd at CCMAR. However, due to the trainer's unavailability on that date, the course was rescheduled to November. The updated plan foresees two complementary sessions: an online, more theoretical session (date to be confirmed) and an in-person session on November 19th, to be held at UCIBIO-Porto alongside [BioData.pt](https://biodata.pt) All Hands Meeting.

The milestone M4.3, which comprises a Training session in eLABFTW at CCMAR, in Faro, has already been scheduled, its programme finalised and registrations are now open (<https://events.biodata.pt/event/13/>). For this reason, the milestone is considered 50% completed, with the remaining progress to be achieved upon the successful implementation of the training session.

WP5 - Impact and Sustainability

The assessment of BioData.pt node and communities services, which corresponds to milestone M5.1, was successfully achieved within the planned timeline (Evidence Report Milestone M5.1 - Assessment of BioData.pt node and communities services). This milestone encompassed the evaluation of BioData.pt | ELIXIR Portugal Node and Community Services through the implementation of a two-tier badge system, distinguishing “Node Services” and “Community Services” according to criteria such as maturity, quality, and availability. The process included the creation of an Evaluation Board, the development of a Service Evaluation Form aligned with international ELIXIR guidelines, the launch of an open call for expressions of interest, and the update to the service catalogue on the [BioData.pt](https://biodata.pt) website.

Regarding the milestone M5.2 - Participation in projects promoted by European infrastructures (ESFRI) and/or funded by European funds, it was expected to reach 75% completion by this reporting period and has instead achieved 90%. This progress reflected BioData.pt's active involvement in upcoming European project proposals, namely as a coordinator (PI) in three submissions to the ELIXIR Science tier 2025 call, as a partner in three additional proposals to the same call, and as a participant in four other Horizon Europe calls.

The milestone M5.3, which comprises the establishment of a core group of data stewards for sensitive life and health data, was initiated during the National Data Stewards Network meeting organised by Re.Data, on June 17th 2025 (<https://shorturl.at/mpTGI>). On this occasion, BioData.pt formally presented the creation of a national Core Group of Data Stewards for Life and Health, supported by the GLIM-BioData project. The initiative aims to connect data stewards, bioinformaticians, and professionals across Portugal, fostering collaboration, disseminating best practices, promoting Open Science and FAIR principles, and supporting the visibility of relevant training opportunities, events, and career opportunities. 16 members are already registered, and the next step will be to schedule its first meeting.

The milestone M5.4 - Key Performance Indicators definition and monitorization has reached 50% completion, as initially planned for this reporting period. At this stage, most of the KPIs, as well as the framework, have already been defined, and their monitoring has started.

2. Overall Execution

During the current reporting period (May-August 2025), the GLIM-BioData project maintained a steady pace of implementation, achieving three deliverables and two milestones, all in line with the established timeline. This corresponds to a cumulative execution of nine deliverables and milestones out of 30 (29%), which exceeds the expected progress at this stage.

Progress was distributed across all work packages. In WP1 - Coordination and Management, the deliverable D1.3 (2nd Progress report) is being submitted on time, while D1.5 (Definition of the institutional policy for RDM and sharing) progressed to 70% completion, slightly below the planned 80% but with mitigation measures in place to ensure timely delivery. The deliverable D1.7 (Participation in the Consortium-GDI coordination activities) is on track, with two General Assemblies already attended. In WP2 - Management of Non-sensitive Life Data, both D2.1 (Curation and datasets availability plan for Non-sensitive Life Data) and M2.2 (DMPortal integration in POLEN) advanced to 50% of their completion. Regarding WP3 - Management of Sensitive Life and Health Data, the preparation of D3.1 (Curation and datasets availability plan for Sensitive Life and Health Data) reached 40% completion, while M3.2 (Standard Operating Procedures for FEGA Portugal Helpdesk) advanced to 50%. Milestone M3.3 (Guidelines to address Legal and Ethic Issues in RI) will be due with the contract of the service for DPO. In what concerns WP4 - Training and Capacity Building, two deliverables (D4.1 - Training sessions on best practices in RDM for Life Data - “Ready4Intensive” Course - GLIM 1st Edition, and D4.3 - Training Session for Data Stewards - 1st session) were completed with successful training sessions. Progress was also registered in both D4.4 - Training Session for Data Stewards - 2nd session and M4.1 - Training on accessing, processing, and curating datasets, while M4.3 (Training in eLabFtw) reached 50% of its completion. Finally, in WP5 - Impact and Sustainability, milestone M5.1 (Assessment of [BioData.pt](https://www.biodata.pt) node and communities services) was completed ahead of schedule, while M5.2 (Participation in projects promoted by European infrastructures (ESFRI) and/or funded by European funds) progressed beyond the planned target, reaching 90% completion. Regarding milestones M5.3 (Core Group of Data Stewards for Life Sciences) and M5.4 (KPI definition and monitoring), the first one was initiated and reached 80% completion, while the second one achieved 50% progress.

As of this reporting period, and as mentioned before, nine out of the 30 deliverables and milestones anticipated in the GLIM-BioData project have already been successfully achieved, corresponding to an overall execution rate of 29%. This level of progress is in line with the project’s implementation plan and demonstrates steady advancement towards the established objectives.

Financial Execution

	Total budget until the end of the Project (Values without IVA)	Execution as of the report date
Total Budget for activities	220 200 €	74 585 € (34%)
Human resources and service provision - includes BioData.pt HR, hiring a Data Steward, hiring a SysAdmin and service provision of a DPO	139 000 €	61 000 € (44%)
Training actions - Trainers, fee for Train-the-Trainer, and travel and accommodation	15 200 €	5 500 € (35%)
Events - Technical meeting, All hands, Roadshow, travel and accommodation	18 000 €	5 000 € (28%)
Technical Resources - server, workstation and data center maintenance	48 000 €	3085 € (6.4%)

3. Additional Relevant Considerations

The GLIM-BioData project carried out a wide range of coordination, communication, training, and integration activities during this period.

The GLIM-BioData consortium actively engaged in several national and international initiatives that reinforced its visibility and integration within the broader research data ecosystem. [BioData.pt](https://shorturl.at/GRDcu) participated in the GBIF training course “Publicação de dados de biodiversidade através do GBIF - 2025” (25-27 June, ISA, Lisbon), where Data Steward Miguel Cisneiros represented the project and strengthened expertise in FAIR data practices (<https://shorturl.at/GRDcu>). [BioData.pt](https://shorturl.at/pYqfP) team also participated in major national science events, including the Jornadas de Computação Científica 2025 (7 May, Aveiro) (<https://shorturl.at/pYqfP>) and Encontro Ciência 2025 (9-11 July, Lisbon),

promoting the project's role in advancing Open Science and Research Data Management (<https://shorturl.at/F7DrT>).

At the European level, [BioData.pt](https://shorturl.at/X6Enp) took part in the ELIXIR Heads of Nodes Retreat 2025 (<https://shorturl.at/X6Enp>), represented by ELIXIR Portugal Head of Node Professor Ana Teresa Freitas, and was also represented by Jorge Oliveira at the Genome of Europe Consortium Meeting (<https://shorturl.at/gp3fh>), further consolidating the project's presence in strategic international infrastructures. In parallel, GLIM-BioData actively contributed to the 2025 ELIXIR All Hands Meeting in Thessaloniki (3-7 June), where the BioData.pt team presented their work and exchanged knowledge with European counterparts.

Capacity-building activities were reinforced through participation in the Bioinformatics Workshop hosted by the School of Tourism and Maritime Technology of the Polytechnic Institute of Leiria (<https://shorturl.at/00QRc>), as well as through the ELITMa Module training on Node Data Management Strategy (<https://shorturl.at/IZQ2u>).

These activities highlight the consortium's strong commitment to community building, international collaboration, and the promotion of best practices in research data management across the life and health sciences.

4. Next Steps

In the upcoming reporting period (September-December 2025), the GLIM-BioData consortium will focus on consolidating the progress achieved and advancing in the implementation of the net set of deliverables and milestones. Coordination will continue through regular project meetings, active participation in GDI Consortium activities, and the finalisation of the institutional policy for the Research Data Management and sharing (D1.5), due in October.

In WP2, the consortium will build on the progress achieved with the completion of D2.1 and the 50% advancement of M2.2, by finalising the integration of the [BioData.pt](https://shorturl.at/X6Enp) DMPortal in the POLEN Repository in collaboration with FCCN, ensuring technical and organisational alignment for successful delivery.

For WP3, efforts will be directed towards the completion of deliverable D3.1 and on preparing D3.2, scheduled for December, while continuing the development of SOPs for the FEGA Portugal Helpdesk (M3.2) through a second SOPathon.

Regarding WP4, a second edition of the Ready4Intensive course (D4.2) will be delivered in October at the Universidade de Coimbra, as well as two sessions of the Training for Data Stewards (D4.4), including an online module in early October and an in-person event on 17 November at UCIBIO-Porto, coinciding with the [BioData.pt](https://shorturl.at/X6Enp) All Hands Meeting. The milestone M4.1 is also rescheduled to take place in November, ensuring alignment with the revised plan.

Finally, within WP5 the project will continue monitoring progress against defined KPIs (M5.4), further advance participation in European initiatives under the GID Consortium ELIXIR, FEGA and EOSC, and reinforce national collaborations to ensure the long-term sustainability of the GLIM-BioData RDM Centre.