FCCN Serviços digitais fct | fct Fundação para a Ciência e a Tecnologia | PRR Plano de Recuperação e Resiliência | REPÚBLICA PORTUGUESA | Financiado pela União Europeia NextGenerationEU

GLIM
BioData.pt

Detailed Work Plan

No âmbito do PNCADAI: Medida RE-C05-i08 - Ciência Mais Digital do PRR - Programa de Recuperação e Resiliência

I. Identificação do Centro de Gestão de Dados de Investigação /Research Data Management Centre Identification

Acrónimo/Acronym	GLIM-BioData
Título Completo em Português/ Full Title in Portuguese	Portal para a Gestão de Dados da Vida e Saúde
Título Completo em Inglês Full Title in English	Gateway for Living Data Management

II. Identificação da Entidade Responsável/Identification of the Responsible Entity

Nome Completo/Full name	BioData.pt - Infraestrutura Portuguesa de Dados Biológicos
Morada/Address	Rua da Quinta Grande, 6, 2780-156 Oeiras
NIF	516 416 120
Categoria/Category	Infraestrutura Científica e de Investigação e pertence ao Roteiro Nacional de Infraestruturas Científicas

III. Identificação dos pontos de contacto/Identification of contact persons

Primário/Primary	Inês Chaves
Secundário/Secondary	Luciana Peixoto

IV. Resumo da Proposta/Proposal Summary

The Gateway for Living Data Management centre (GLIM-BioData) will be the keystone initiative needed to cement BioData.pt as a primary provider of national expertise within the domain of management and valorisation of Portugal's life and health scientific data. Building on BioData.pt extensive experience in scientific data management, GLIM-BioData will focus on fostering the sharing and reuse of life and health data through establishing curation and management guidelines and practices, refining them through two high-relevance uses cases in biodiversity and genomic data, and providing state-of-the-art training in bioinformatics, data management, and data stewardship. A key aspect of GLIM-BioData is the establishment of tailored processes for data management in both sensitive and non-sensitive domains.

In the first use case, GLIM-BioData will support the public annotation and sharing of non-sensitive biodiversity datasets in accordance with the FAIR principles of data management and stewardship, to monitor the dissemination of virulent bacterial clones in aquatic environments and develop predictive models for assessing human health risks. Through our collaborations in the EMO-BON and FAIR-EASE initiatives, GLIM-BioData will focus on enhancing data reusability with a focus on anthropogenic pressures, providing crucial tools for predicting and mitigating future outbreaks of hazardous marine bacteria, and thereby integrating environmental and human health outcomes.

The second use case focuses on the curation and management of sensitive data, building on our collaboration in the FEGA. GLIM-BioData will establish a helpdesk for FEGA Portugal to support and capacitate users in national institutions in the data deposition process, as well as in their reuse by providing datasets for sensitive life and health data. It will also develop guidelines and templates for data stewardship of sensitive data that address legal and ethical issues, including compliance with the General Data Protection Regulation (GDPR).

Training is a core part of GLIM-BioData and a key driver of its impact. Building on BioData.pt's extensive experience, it will offer both in-depth training in data stewardship — to train the next generation of life and health data stewards — and training in research data management (RDM) to empower researchers and students. All GLIM-BioData trainers will be certified by ELIXIR, the European life sciences infrastructure.

GLIM-BioData's long-term sustainability plan includes monitoring activity impact, strengthening collaboration within research communities, and widening our participation in European projects.

GLIM-BioData is backed by a team of expert researchers from 17 different institutions across the country and covering a wide variety of topics. This thematic and geographic diversity is highly synergistic and a crucial aspect to ensuring the success of GLIM-BioData's mission in advancing Portugal's capacity to manage life and health data, foster collaboration and enhance research in alignment with European standards. This will be leveraged by the participation in Consortium-GDI.

a. State of the Art in Research Data Management From BioData.pt

BioData.pt's mission is to support the management, dissemination, and enhancement of biological data generated by research, clinical, and industrial activities in Portugal through

resources in computing, bioinformatics, and data management. Additionally, BioData.pt is responsible for managing the Portuguese node of the European ELIXIR network (European Life Sciences Infrastructure for Biological Information), a key component of the ESFRI (European Strategy Forum on Research Infrastructures) network.
BioData.pt is presently coordinating the implementation of the Portuguese node of the European "Genomic Data Infrastructure" network, which it will manage when deployed. The infrastructure will host the genomes of Portuguese individuals as part of the European 1+ Million Genomes initiative. In this context, BioData.pt, together with SPMS (Serviços Partilhados do Ministério da Saúde), is also the entity responsible for the data management and computational architecture proposal for the national strategy to implement genomic medicine in Portugal. Additionally, BioData.pt is responsible for implementing the National node of the European Genome Archive (EGA) network, which enables the management of sensitive data obtained by scientific projects in Portugal, giving European and international visibility to research conducted in Portugal and participating in European efforts in the field of Big Data analysis.
As a member of the National Roadmap for RI (RNIE), BioData.pt maintains a close partnership with FCT and its digital services unit FCCN, with whom a cooperation protocol has been signed.
BioData.pt is currently a member of the RCTS Network, maintained by FCCN. As a distributed digital infrastructure, the support of FCCN has provided access and support to RCTS services such as software licences, hosted web platforms, connectivity solutions and access to RNCA National Network for Advanced Computing. BioData.pt | ELIXIR Portugal web services are already running on RNCA infrastructure, and reaching a worldwide audience, especially with respect to the popular Beacon and FEGA Portugal federated human data services.

b. Brief Description of the BioData.pt's Open Science Strategy

BioData.pt's mission is centred on Open Science as a fundamental factor in developing the national scientific fabric. BioData.pt provides services for the management, collection, analysis, and sharing of biological data, promoting alignment with open data policies and the enhancement of scientific research. These benefits are significant considering the implementation of Open Science policies by the European Commission (EC), including the obligation to make all data generated FAIR (Findable, Accessible, Interoperable, and Re-usable), but also of immense potential given the benefits that data re-use brings for scientific advancement. BioData.pt has long supported its members in sharing their scientific data with notable examples of freely available tools and datasets that due to their high impact are supported through ELIXIR (Fig 1).

Name	Provider
Bioinformatics Docker Images Project	i3S - Institute for Research and Innovation in Health
CorkOakDB	BioData.pt
DMPortal	BioData.pt
EvoPPI	i3S - Institute for Research and Innovation in Health
PHYLOViZ	INESC-ID
Plant sRNA Portal	
YEASTRACT	INESC-ID

Figure 1: List of computational tools and databases freely available and supported by BioData.pt

In addition, BioData.pt further supports another 20+ through a grassroots movement whereby Biodata communities share their data and tools.
Our Open Science strategy is also strongly anchored in training services, which ensure our members are skilled in the best practices of Open Science (see section II). BioData.pt has been committed to promoting and developing tools for data management and sharing . To empower national researchers and research technicians in best practices for Open Science and FAIR data sharing, BioData.pt developed a training programme - "Ready for BioData Management?", which trained over 600 individuals. Moreover, the "Training Data Stewards for Life Sciences" trained a first generation of experts, totalling 25 individuals from at least 12 institutions creating a network of support in the implementation of data management, open science and FAIR data principles in Portuguese R&I organisations. All BioData.pt|ELIXIR Portugal material is available at TeSS.

To further promote Open Science, several BioData.pt members participate in working groups from the "Fórum GDI" (e.g., Training and Skills for FAIR Data and Management, Data Repositories: Technology, Organisation, and Certification).

Aligned with EOSC, BioData.pt actively facilitates the interoperability of datasets to foster a connected and open research environment. Through the provision of robust infrastructure and targeted training, BioData.pt will drive the widespread reuse of high-quality, interoperable data, including 'omics biodiversity data products and workflows developed as part of EOSC initiative projects. Moreover, GLIM-BioData will profit from the implementation of the open data strategy of the Consortium-GDI.

c. Alignment with the Objectives of the PN CADAI

PN CADAI aim 1 - Development of infrastructures and tools for data planning and curation:

BioData.pt hosts and manages several services and repositories created for, by, or with our associates. These services are being supported to achieve the ELIXIR Europe classification of "Core Services". Examples are DMPortal, DSWizard, and CorkOakDB. See: https://biodata.pt/services/node-services

PN CADAI aim 2 - Technical support, consultancy, and training for the scientific community in best practices for data management:

Through GLIM-BioData, we will continue to improve and diversify the array of services currently provided by BioData.pt and ELIXIR Portugal, such as Support, Training, Node Services, and Community Services.

PN CADAI aim 3 - Alignment with the Open Science principles and FAIR data endorsed by Portugal:

BioData.pt is strongly committed to the Principles of Open Science and FAIR Data. Not only do our Research Data Management (RDM) training materials rely heavily on guidelines and good practices towards those principles, but we also promote tools such as ELIXIR's RDMkit and work closely with stakeholders on these matters, e.g. Research Data Alliance Portugal.

PN CADAI aim 4 - Convergence with EOSC:

BioData.pt is the national representative at EOSC-ENTRUST, a Horizon Europe Research Project focused on the joint development of a common blueprint for federated data access and analysis.

PN CADAI aim 5 - Coordination with other initiatives, namely open data in Public Administration and businesses:

We are involved in diverse projects with FCT and FCCN, while also collaborating with industry stakeholders such as P-Bio, The Navigator Company, HeartGenetics, and SilicoLife.

d. Description of the Research Data to be Made Available and Its Potential for Reuse

GLIM-BioData will support existing and future projects that are required to make their data openly available: thereby enabling and promoting data reuse.

Data management support will be provided by the data stewards to the researchers to speed up the process of data publication in open repositories ensuring data quality-assured formats and outputs comply with current research standards. Particular attention will be given to the development of tools to enable the management of sensitive data, such as those from the field of human health.

Non-sensitive Life Science data, namely biodiversity research data from projects like EMBRC EMO BON and EOSC FAIR-EASE, can be reused to provide insights into the diversity, distribution, and ecological roles of species, as well as their environmental impacts on ecosystems. Together with the FAIR-EASE project, GLIM-BioData will enhance the operability of services for managing data within a single system, enabling the integration of marine genomic biodiversity and microbial data with ecosystem services research in the context of global climate change. Synergies gained through interoperability with other environmental science data can be used for ecological research, conservation efforts, climate change studies, and policy-making. Through this interlinked reuse of cross-disciplinary data, researchers can gain a comprehensive understanding of ecosystem interactions and enhance predictive models.

In the context of health data, particularly in the use of genetic and genomic data, the GDI, FEGA, and Beacon projects play a central role. These data, classified as Sensitive Life Data, are being generated in research projects by BioData.pt Associates as well as through the national implementation of the European 1+MG initiative. BioData.pt provides a 'Help Desk' service to all partner institutions and supports the establishment of data transfer agreements between institutions and across borders, including ensuring compliance with the General Data Protection Regulation (GDPR) to provide the necessary framework for the secondary use of data. When these data are used with modern research methods, such as machine learning, they have the potential to revolutionise healthcare, develop precision medicine solutions, and unravel the mechanisms behind many diseases such as cancer and rare genetic disorders.

V. Detailed description of the Work Plan

With the building of the Research Data Management (RDM) Centre GLIM-BioData, we propose extending and enhancing the BioData.pt infrastructure to support the digital transition and to address the major challenges of the PNCADAI, namely in the management of biological data generated in research in Exact and Natural Sciences, and in Medical and Health Sciences. This work plan is thus divided into 5 Work Packages (WPs) as described below. Deliverables (D) and Milestones (M) are presented in each WP and are scheduled in section b.

WP1-Coordination and Management
Leader - Inês Chaves (BioData.pt); Co-leader Luciana Peixoto (BioData.pt)

This WP aims to coordinate and manage the GLIM-BioData centre and articulate all the WPs. This responsibility falls to the BioData.pt team.

Task 1.1 Preparing and monitoring the work plan for the GLIM-BioData
Preparation of a comprehensive and exhaustive work plan for the GLIM-BioData (**D1.1**) until **9 May 2025**, monitoring its implementation, and preparation of four-monthly progress reports (**D1.2 - 13 May 2025**, **D1.3 - 10 September 2025**, **D1.4 - End of project**).

Task 1.2 Define, implement, and adopt a policy for management and sharing research data
Development of a policy for the management and sharing of biological research data, specifically Life and Health Data, by all GLIM-BioData members. This data management policy will be aligned with the National Strategy for Open Data under development by the entity INCoDe 2030 (**D1.5 - October 2025**), as well as promoting the adoption of the policy (**M1.1 - End of project**) and monitoring its implementation (**D1.6 - End of project**). This policy may be a requirement for funding / support / certification of BioData.pt's services.

Task 1.3 Articulate and coordinate GLIM-BioData's actions with the Consortium-GDI
Participate in the Consortium-GDI meetings (**D1.7 - End of project**), namely in the 1st Re.Data General Assembly Meeting (26 February 2025) with GDI Centers and in the following, and coordinate the GLIM-BioData actions with the consortium's principles (Consortium Invitation to participate in the Kick Off Meeting, as well as in the round table "Research Data Management Centres for Living Data: Driving Open Science and Innovation" organised by GLIM-BioData in Bioinformatics Open Days (26 March 2025)), promoting the integration of members into the networks developed by the Consortium. Inês Chaves, as WP1 Leader, will integrate the National Interest Group on Open Science and Research Data Management. Materials and resources of the Re.Data Consortium - such as the Characterisation of Research Data Management Support Professionals survey - will be disseminated with GLIM-BioData members.

Task 1.4 Dissemination and outreach
Disseminate the project's results as described in the **Communication plan** - Visual Identity (Logo, Website, Social Networks, Newsletter); Internal Communication (mailing lists, Teams channels, team meetings to be held bi-monthly via video conference); External Communication and Knowledge Transfer (dissemination of GLIM-BioData information in BioData.pt media channels, publication of results in indexed scientific journals with high impact and through presentations at scientific conferences, deliverables published in press releases, dissemination of training activities); Disclosure of Impact (bi-monthly KPI monitoring, namely through newsletter subscribers, web page views, followers and shares on social media, number of published scientific articles, oral presentations and training activities participations).
Promotion of a **Roadshow** about the GLIM-BioData principles (**M1.1 - End of project**), and its articulation with the Consortium-GDI. Roadshow about human-sensitive data and alignment with FEGA Portugal (**M1.2 - End of project**). The scheduled roadshows includes the associates: UCIBIO Lisboa - January 8, UCIBIO Porto - January 21, CiiMar - April 28, iBET/ ITQB - May t.b.d., GiMM and CEBAL - June t.b.d., FCUL and Nova Medical School - July - t.b.d., CCMAR - September t.b.d. and Universidade de Coimbra - October t.b.d.. Other associates or possible future associates can be visited if justified. During the roadshows, the data management and sharing policy will be presented, along with the consortium's activities, established networks, and efforts to promote the integration of members into these networks — particularly the national network of data stewards (**D5.1 - March**) in articulation with the WP5.

To achieve the objectives of the WP1, Deliverables (D) and Milestones (M) are defined:

D1.1 - GLIM-BioData Centre Work Plan - **9 May 2025**
D1.2 - 1st Progress report - **13 May 2025**
D1.3 - 2nd Progress report - **10 September 2025**
D1.4 - Final report - **End of Project**
D1.5 - Definition of the institutional policy for RDM and sharing - **October 2025**
D1.6 - Implementation and adoption of the institutional policy for RDM and sharing by the GLIM-BioData participant institution(s) - **End of Project**
D1.7 - Participation in the Consortium-GDI coordination activities - **End of Project**

M1.1 - Roadshow about the GLIM-BioData Centre and the integration in the Consortium-GDI- **End of Project**
M1.2 - Roadshow about human-sensitive data and alignment with FEGA - **End of Project**

WP2-Management of Non-sensitive Life Data for Sharing and Reuse
Leader - Bruno Louro (CCMAR), Co-leader Gil Poiares-Oliveira (BioData.pt)

The objective of this WP is to reuse currently available non-sensitive data from biodiversity omics and enhance these data by making their respective transformed data products interoperable with other data streams within Earth sciences. A collaboration with the Centre GDI FAIRway, GBIf and iRe:RESEARCH is being fostered.

Task 2.1 Hosting and planning of omics diversity datasets availability
A data management plan will be provided by the data steward to the researchers to speed up the process of data publication in open repositories, ensuring data quality-assured formats and outputs comply with current research standards, with emphasis on those applying to the field of biological diversity (DarwinCore) (**D2.1 - October 2025**). The non-sensitive data produced in the context of research by the GLIM-BioData in biodiversity and microbiota will be annotated using the state-of-the-art metadata standards and deposited for public access in DMPortal. To assure the maintenance and sustainability of DMPortal, the migration to POLEN will start (**M2.1 - July 2025**). This will support for authenticated GLIM-BioData researchers to annotate and share datasets publicly in an intuitive manner and compliance with FAIR principles (**D2.2 - End of Project**). We will provide training on accessing and reusing publicly available datasets to promote data reusability in line with the EOSC interoperability framework. This will be covered in Task 4.3.

Task 2.2 EMO BON and EOSC FAIR-EASE biodiversity data reuse case
Aligned with the One Health approach, in this project, we will use EMO-BON and FAIR-EASE datasets to monitor the spread of virulent bacterial clones in aquatic environments and develop predictive models for human health risks. In GLIM-BioData, machine learning algorithms will identify genes and metabolic functions as biomarkers for clinically important bacteria by continually integrating omics datasets from European marine ecosystems with environmental data. Adopting the FAIR-EASE interoperability methods will enable real-time monitoring, risk assessment, and analysis of human impact, providing tools to predict and mitigate future outbreaks, thus linking environmental and human health outcomes.
RO-crates will facilitate annotating and linking datasets within a package, enhancing their findability, accessibility, interoperability, and reusability (**D2.2 - End of project**).

Task 2.3 Electronic laboratory notebooks

An instance of the open-source electronic laboratory notebook, eLabFTW (**M2.2 - November 2025**), will be installed and available to all members of GLIM-BioData. Users will be able to deposit data and custom JSON metadata directly into BioData.pt repositories or public repositories via their eLabFTW interface.

To achieve the objectives of WP2, Deliverables (D) and Milestones (M) are defined:

D2.1 - Curation and datasets availability plan for Non-sensitive Life Data - **October 2025**
D2.2 - Publication of datasets for Non-sensitive Life Data - **End of Project**

M2.2 - DMPortal integration in POLEN - **July 2025**
M2.1 - Open-source electronic lab notebook (eLabFTW) implementation - **November 2025**

WP3-Management of Sensitive Life and Health Data for Sharing and Reuse
Leader Jorge Oliveira (INESC-ID/IST), Co-leader Miguel Cisneiros (BioData.pt)

The objective of WP3 is the development of policies for management of Sensitive Life and Health Data following RGPD, the publication of datasets after curation and following the policies, curation of the datasets (**D3.1 - September 2025**) and report on outreach activities with FEGA Portugal stakeholders (**M3.1 - End of project**). Collaborations with Centre GDI CeSDHR and AIBILI will be fostered.

Task 3.1 Create FEGA Portugal Helpdesk
Establish a helpdesk for the FEGA Portugal by deploying a Request Tracker system, to keep all the user queries organised and define the Standard Operating Procedures (SOP) for the Operation of the Helpdesk (**M3.2 - December 2025**). This SOP must be aligned with the ones proposed by the Central EGA Operational Committee.
A Data Steward will be allocated to the helpdesk and be trained to fully understand the FEGA submission system. This Helpdesk will support and capacitate users in national institutions in the data deposition process, as well as in their reuse by providing datasets for sensitive life and health data (**D3.2 - End of project**).

Task 3.2 Legal and Ethics Issues
A Data Protection Officer (DPO) service will be available to streamline the whole legal process. This is essential as each dataset deposited in FEGA requires the inclusion of a Data Processing Agreement (DPA) between the institutions.

- Develop and make available a set of templates of DPA and other legal documents.
- Develop a set of guidelines for a correct approach to data stewardship of common legal issues about sensitive data within RI and develop guidelines to address Legal and Ethical Issues in RI (**M3.3 - End of project**).

Task 3.3 Creating Awareness
Many institutions are still unaware of the available solutions for the complete life cycle of sensitive data. This includes its collection, storage, processing, sharing, and re-use. In coordination with WP1, a roadshow **(M1.2 - End of project)**, will be carried out close to the stakeholders, the major institutions dealing with sensitive data, in order to create awareness about the existence of resources like FEGA and Beacon in articulation .

To achieve the objectives of the WP3, Deliverables (D) and Milestones (M) are defined:

D3.1 - Curation and datasets availability plan for Sensitive Life and Health Data - **September 2025**
D3.2 - Publication of datasets for Sensitive Life and Health Data - **December 2025**

M3.1 - Report on outreach activities with FEGA stakeholders - **End of Project**
M3.2 - Standard Operating Procedures for FEGA Portugal Helpdesk - **December 2025**
M3.3 - Guidelines to address Legal and Ethics Issues in RI - **End of Project**

WP4-Training and Capacity Building
Leader Luciana Peixoto (BioData.pt), Co-leader Gil Poiares-Oliveira (BioData.pt)

This WP is dedicated to planning, structuring, developing, preparing, implementing, evaluating, and certifying training in the areas of bioinformatics, data management and data stewards. Synergies with the consortium Re.Data have already been identified and will be fostered.

Task 4.1 "Ready for BioData Management?" Pool of Trainers
The needs of training in the field of bioinformatics, data management and data stewards in all associated institutions will be identified in the BioData.pt Technical meeting. BioData.pt Training Platform will meet to identify possible Trainer. The trainers of GLIM-BioData Centre will be certified by the ELIXIR (following a "Train-the-Trainers" course), and utilise active learning methodologies, particularly project-based learning, which allows for the development of a micro-certification system for the individuals trained by GLIM-BioData Centre.

Task 4.2 "Ready for BioData Management?" Training
Training in management of Life and Health Data will be offered to empowering researchers and technicians in the country. This relies on alignment with the ELIXIR Training Programme.
This intensive course is designed to provide life sciences researchers with a comprehensive introduction to Research Data Management (RDM) across the entire data lifecycle. Through a mix of theoretical content, group discussions, and practical exercises, participants will engage with tools such as Galaxy, DataVerse, and the Data Stewardship Wizard (DSW). The course is aimed at researchers, PhD students, postdocs, and lab managers working with biomedical or life sciences data. No prior experience in RDM is required, although participants should have some familiarity with handling research data. Training will be delivered by experts from BioData.pt using resources from the ELIXIR RDMkit.
Two "Ready4Intensive" training sessions will be promoted (**D4.1 - 3 and 4 July 2025 in CiiMAR and D4.2 October 2025 in UCoimbra**), in Life Data and Health Data. This task also aims to expand the programme by broadening the range of courses and increasing their frequency, as well as formalising the programme within the context of higher education by accrediting the courses by the ECTS system.

Task 4.3 Training in Management of Biodiversity/Life Data
To encourage data reuse, training will be developed on accessing, processing, and curating the EMO BON and FAIR-EASE datasets (**M4.1 - 22 September 2025 Online**). This will support new queries on biodiversity data with improved interoperability through RDM services. With the FAIR-EASE Interdisciplinary Data Discovery and Access Service, data can be used to query existing aggregated biodiversity results and various environmental climatic datasets.

Task 4.4 Training in eLABFTW

Two training sessions will be held in CCMAR and CIIMAR for new users of eLABFTW, an open-source electronic lab notebook, to promote FAIR principles from the start of research data collection (**M4.2 - 28 April in CiiMAR and in collaboration with Centre GDI FAIRway and M4.3 - 22 September in CCMAR**). Aimed at researchers, lab technicians, and data managers, this hands-on session will explore the key features of eLabFTW, including experiment documentation, inventory management, and team collaboration tools, helping participants adopt best practices in data management and reproducibility.

Task 4.5 Training Data Stewards for Life Sciences
Two training sessions of a new edition of the "Training Data Stewards for Life Sciences" courses will be carried out, reinforcing the network of data stewards in Portugal (**D4.3 - 7 July (Online) and D4.4 17 November 2025 (in person)**).
The programme will combine one online and one face-to-face session, including project-based learning activities. It will allow trainees to get extensive practice and consolidate the acquired knowledge, and it will include the elaboration of several assignments, such as Data Management Policies for the institutions of the trainees. Participants will explore topics such as vocabularies, data workflows, FAIR principles, and data sharing using platforms like GitHub, Dataverse, and Beacon. The course will be limited to 32 participants (2 per BioData.pt member institution), and completion of the Ready4Intensive course is required.
The theoretical part of this course will also be professionally recorded and it will be made available publicly along with all the materials of the course.

Task 4.6 Capacity Building
Anchored in training courses and other science dissemination events such as technical meetings, community and platform meetings, and general meetings with members of all associates, and conferences (e.g. the Microbiome Summit 2021 and 2023, organised by BioData.pt), the GLIM-BioData will promote national and international collaboration and create links between academia and industry. To this end, we count on a long-standing partner, P-Bio, which connects with the Portuguese Biotechnology Industries, as well as other partnerships that have emerged, anchored in Knowledge Exchange programmes promoted by ELIXIR.

To achieve the objectives of the WP1, Deliverables (D) and Milestones (M) are defined:

D.4.1 - Training sessions on best practices in RDM for Life Data - "Ready4Intensive" Courses - 1 edition - **3 and 4 July 2025 in CiiMAR**
D.4.2 - Training sessions on best practices in RDM for Life Data - "Ready4Intensive" Courses - 2 nd edition - **October in UCoimbra**
D4.3 - Training Sessions for Data Stewards - 1st session - **7 July 2025 (Online),**
D4.4 - Training Sessions for Data Stewards - 2nd session - **17 November 2025 (in person).**

M4.1 - Training on accessing, processing, and curating datasets - **22 September 2025**
M4.2 - Training in eLABFTW - **28 April 2025 in CiiMAR in collaboration with Centre GDI FAIRway**
M4.3 - Training in eLABFTW - **22 September in CCMAR**

WP5-Impact and Sustainability
Leader Catarina Milho (BioData.pt), Co-leader Luciana Peixoto (BioData.pt)

Develop and implement strategies for the sustainable growth of GLIM-BioData, including financial planning, resource allocation, and enhancement of data management infrastructure, will be promoted in collaboration with the WP1. Opportunities for skills development, career

advancement, and training for GLIM-BioData members will be promoted in collaboration with WP4.

Task 5.1 Facilitation of collaborations within communities and Portuguese Data Steward Network

The GLIM-BioData members will be encouraged to participate in communities of researchers from different scientific fields to foster collaboration among various national research groups, using mailing lists, Slack channels, and shared folders on cloud-based file-sharing infrastructure. Active dialogue with the scientific community will be also encouraged with the creation of dedicated Open Science Working Groups (Non-sensitive Life Data and Sensitive Life and Health Data) within the associated institutions to share experiences and best practices in the domain of Open Science and foster collaboration and integration with the Portuguese Data Steward Network (**D5.1 - March 2025**). Within this network, GLIM-BioData members will attend the monthly meetings, share useful links and resources on Data Stewardship via the network's mailing-list, and contribute for the definition of the RDM Data Professional profile with insights on life-sciences and biological data handling requirements.

Task 5.2 Identification and Development of Resources

Work towards the identification and development of essential national computational and software resources necessary for cutting-edge research in the life and health sciences, ensuring that these resources comply with European standards and protocols to enhance their visibility and impact (**M5.1 - June 2025**).

Task 5.3 Participation in projects promoted by European infrastructures (ESFRI) and projects funded by European funds

Actively participate in European projects and projects promoted by ESFRI (e.g., the ELIXIR Europe infrastructure) (**M5.2 - End of project**), which have the potential to transform national strategies in the areas of health and life sciences, through knowledge, resources, and collaborations.

Task 5.4 Visibility and Advocacy

Promote the visibility of the activities, developments, and contributions of the GLIM-BioData Centre at both national and international levels. Advocate for the importance of national, data-driven research in bioinformatics in advancing health technologies and life sciences.
Integrate the Portuguese Data Stewards Network (**D5.1 - March 2025**) and within this Network, GLIM-BioData is committed to establishing and maintaining a biological data community sub-group and establishing a core group of data stewards for sensitive life and health data (**M5.3 - November 2025**). Moreover, we will vigorously promote contact and interaction with other working groups, networks, and initiatives at both national and international levels. Additionally, we will assertively align with the Portuguese Data Stewards Network to disseminate vital information, events, training, and resources on data stewardship and RDM.

Task 5.5 KPIs Definition and Impact Monitorisation

Based on the 10 commitments outlined in the Agreement on Reforming Research Assessment, the quality and impact of the GLIM-BioData activities will be monitored in terms of quantitative and qualitative KPIs. This will include recognising a variety of outputs, practices, and activities, ensuring a more comprehensive assessment aligned with the Coalition for Advancing Research Assessment (CoARA) framework, and the RI imPact Assessment paTHways (RI-PATHS) methodology. The quality and impact of the RI's activities will be qualitatively and quantitatively monitored through the use of the impact pathways (P) from the RI-PATHS methodology already implemented at BioData.pt in addition to the commitments set forth by CoARA.
Key Performance Indicators (KPIs) will be used for impact monitoring (**M5.4 - End of project**).

To achieve the objectives of the WP1, Deliverables (D) and Milestones (M) are defined:

D5.1 - Integration and collaboration in the national network of data stewards - **March 2025**

M5.1 - Assessment of BioData.pt node and communities services - **June 2025**
M5.2 - Participation in projects promoted by European infrastructures (ESFRI) and/or funded by European funds - **End of Project**
M5.3 - Establishment of a core group of data stewards for sensitive life and health data - **November 2025**
M5.4 - Key Performance Indicators definition and monitorisation - **End of Project**

a. Work Plan Monitorization

Milestones and Deliverables

	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec	Final
WP1				D1.1 D1.2				D1.3	D1.5			D1.4 D1.6 D1.7 M1.1 M1.2
WP2						M2.1			D2.1	M2.2		D2.2
WP3								D3.1			D3.2 M3.2	M3.1 M3.3
WP4			M4.2			D4.1 D4.3		D4.2 M4.1 M4.3	D4.2	D4.4		
WP5		D5.1			M5.1					M5.3		M5.2 M5.4

b. Overall Budget Proposal

	WP1	WP2	WP3	WP4	WP5	Total
Budget Values without IVA	**38 400€**	**26 500€**	**69 500€**	**43 400€**	**42 400€**	**220 200 €**
Human resources and service provision - includes BioData.pt HR, hiring a Data Steward, hiring a SysAdmin and service provision of a DPO	28 400 €	19 000 €	31 000 €	28 200 €	32 400 €	139 000 €
Training actions - Trainers, fee for Train-the-Trainer, and travel and accommodation	-	-	-	15 200 €	-	15 200 €
Events - Technical meeting, KoM, All hands, Roadshow, travel and accommodation	7 000 €	-	1000 €	-	10 000 €	18 000 €
Technical Resources - server, workstation and data center maintenance	3 000 €	7 500 €	37 500 €	-	-	48 000 €