

1st BioData.pt Association Technical Meeting Report



Date: 29th September 2021

Place: Casa de Exercícios de Santo Inácio, Colares

"Building research infrastructures is about supporting other people and relishing in their success, something not obvious in a scientific world increasingly competitive and individualistic".

José Pereira Leal

Context

This was the first F2F BioData.pt meeting after the COVID-19 pandemics, and also the first Technical Meeting of the BioData.pt Association. Our main goal was to welcome new members and establish for all associates their expectations regarding their contributions to the BioData.pt mission and what their organisations expect to receive. These are highlighted in Table I.

Table I. Highlights of presentations by the Associates

Institution	Representative	Groups	We would like to contribute	We would like to receive
Univ Coimbra	Irina Moreira	Computational Biology@UC	Lead ComputationalBiology@UC and ComputationalBiology@Portugal to the international scene Collaboration - Common goals and projects ELIXIR industry platform: connect with UC	Support - For training events, project writing, etc... Aggregation - The creation of an infrastructure and front-end to aggregate technological platforms (national level)
Univ Aveiro	José Luís Oliveira	LAQV-REQUIMTE iBiMed IEETA	<ul style="list-style-type: none"> – Proteomics partnerships / collaborations – Drug design – Sequencing /genotyping requirements / projects – Biological, biomedical and health problems/challenges – Data, of any kind, for processing, analysis. 	<ul style="list-style-type: none"> – Collaborations Dissemination of tools (bioinformatics, text mining, semantic knowledge) by BioData.pt - like SilicoLife)

BioData.pt - Portuguese Infrastructure of Biological Data

INESC-ID	Mário Gaspar da Silva	INESC-ID	Computing and data services, and YEASTRACT+; Computing power; FAIR data formats specification; Training.	Biological data Computing power Data services Partnerships in health and biology
IGC	Ricardo Leite	Quantitative Biology GTPB Bioinformatic Unit Genomic Facility Evolutionary Biology group	RNA-Seq analysis tools Training (GTPB) Leading the PT Microbiome community and populate a future centralized repository for Microbiome data Continuous and active developing of the presented tools	Collaborations in health and life sciences Advanced training in DMP, RDM toolkit A dedicated Portuguese data portal for Microbiome studies More computational tools based in AI/ML (data analysis, bioinformatic pipelines) Guarantee the persistence of tools
CCMAR	João Machado	Plant Systematics and Bioinformatics Research Group (PSB)	Bioinformatics expertise Computational platform (ceta.ualg.pt) - Plan to develop api to give programmatic access to dataset to seagrass data Development and implementation of OS policies	Support for Open Science Policy development Support for training initiatives Support for EU grant writing
CIIMAR	Filipe Castro	Evolutionary Genomics and	Access to all Technology Platforms, including:	Collaborative efforts in the omic and related fields;

BioData.pt - Portuguese Infrastructure of Biological Data

		<p>Blue Technology Animal Physiology and Functional Genomics Animal Nutrition and Health Global Changes and Ecosystems Services</p>	<p>Bioinformatics: genomics and transcriptomics software genome assembly sequence databases</p> <p>Molecular Biology and Omics: Proteomic Analysis Genomic Identification of organisms and gene based analyses.</p>	<p>Team engagement for in common projects and courses</p>
ITQB-NOVA	Nelson Saibo	<p>Genomics of Plant Stress Unit Systems and Synthetic Biology Lab Protein Modelling Lab Multiscale Modelling Lab Molecular Simulation Lab Control of Gene</p>	<p>Improvement of tools already developed within BioData.pt (e.g. INCREASING implementation study)</p> <p>Brainstorming on management, storage (e.g. to design solutions for distributed data storage), and analysis of different types of data</p> <p>Expertise on the analysis of different types of data (e.g. molecular dynamics data, sRNA data)</p> <p>Expertise on dynamic models of microbes and microbial factory design</p>	<p>Training on state of the art data management, data analysis, as well as on data storage / sharing</p> <p>Solutions for storage of different types of data (e.g. molecular dynamics data, molecular topology information for simulations)</p> <p>Storage capacity and computing facilities</p> <p>Solutions to integrate data (e.g. molecular dynamics) with different databases (e.g. cheminformatic databases)</p> <p>Services on data analysis (e.g. proteomic data)</p>

BioData.pt - Portuguese Infrastructure of Biological Data

		Expression Lab		Platform for collaborations with Academic groups and industry
Univ Minho	Miguel Rocha	CEB / Biosystems	<p>Improve our ecosystem of tools for metabolic modeling and engineering - integrated front-end for reannotation/ modeling services</p> <p>Development of tools for machine/ deep learning applied to biological and biomedical data - for now, python-based tools for scripting</p> <p>Strong participation in the future training plan of BioData</p> <p>Active involvement in the setup of the Systems Biology community - opportunities for synergies in BioData, both related to microbial biotech and systems medicine</p> <p>More active involvement in the plant sciences community</p> <p>Extend involvement of UMinho in BioData to other units with a more</p>	<p>Synergies with MACC</p> <p>Bioinformatics training forum - discuss the setup of a national program for Bioinformatics training</p> <p>Improve the web site of BioData to better show available services and tools</p> <p>Industry forum (co-organized with P-BIO ?)</p> <p>Organization of a national bioinformatics conference annually + collaboration with other events (as BOD)</p>

BioData.pt - Portuguese Infrastructure of Biological Data

			<p>“user-oriented” profile</p> <p>Focus group of Machine Learning in PT</p>	
CEBAL	Liliana Marum	AgroGenomic Bioinformatic and Animal Genomics Agro-food valorization Process engineering Bioactive compounds	Networking collaboration Biological data in Plant Science and Microbiology -Genomic; -Transcriptomics; -Metabolomics; -Phenotypic data. Difficulties detected: -Data availability - Research works involve different	Networking collaboration Training in bioinformatics skills; data management and biostatistics Genome assembly and annotation Small and long-non-coding RNA analyses Integration of different omics Occasional of computational resources (Cloud computing)
iBET	Inês Isidro	Animal Cell Technology Labs Food & Health Division Labs	Case studies (data) Access to wet lab (biological platforms and analytics) Collaborations with biopharma and academic partners	Collaborators/services bioinformatics (transcriptomics) and MS data pipelines
IST-SI	Fernando Mira da Silva	Computational Infrastructures	Infrastructure Ongoing and future developments:	Cooperation with other groups in the field of HPC and infrastructures

BioData.pt - Portuguese Infrastructure of Biological Data

		Team	<ul style="list-style-type: none"> -Arm CPU support Single-root input/output virtualization (SR-IOV) -Dedicated hardware support -Automated workflow support 	<p>Contributions and proposals for improved infrastructure services</p> <p>Keep HW at state of the art level</p> <p>Support for technical & specialized human resources</p>
IST-iBB	Miguel Cacho Teixeira	iBB YeastRACT+ / Micro Biotech team	<p>Tools:</p> <p>YEASTRACT+ new features:</p> <ul style="list-style-type: none"> -Community Database -Comparative Genomics Tools for regulatory prediction -Automated construction of global regulatory models <p>Models:</p> <ul style="list-style-type: none"> -Genome-scale metabolic models for pathogens (iRV781, for C. albicans; others coming out soon) -Mixed metabolic-regulatory models for yeasts (S. cerevisiae coming out soon) <p>Marine Metagenome Data Data analysis tools Culture Collections</p>	<p>Facilitated access to european partners, particularly within the Microbial Biotech / Marine Metagenomics Communities</p> <p>Facilitated access to national partners, particularly within the Microbial Biotech / Marine Metagenomics Communities</p> <p>Free access to the BioData.pt computing infrastructure</p> <p>Access to BioData.pt Crash Courses (in Bioinformatics)</p> <p>Funding?</p>

Ciências ULisboa	Cátia Pesquita	BioISI cE3c MARE-ULisboa LASIGE	Data across all BioData communities to share and reuse: plants, marine resources, microbiology, health... Data integration and interoperability Semantic data integration tools and services (integrating different data types/format using biomedical ontologies) Machine learning applications for health and life sciences data Text mining tools in health Predictive machine learning in health	Data producers & consumers Support and training in data stewardship and management, FAIR, data repositories, etc Bioinformatics & Computer Science: interesting problems in data interoperability and data and text analytics opportunities to expand the reach of developed tools Opportunities for funding and knowledge transfer
---------------------	----------------	--	--	---

Challenges

The urge to facilitate access to the tools and services developed and/or provided by BioData.pt was addressed. BioData.pt has a broad list of resources capable of responding to a set of challenges faced by some areas of Portuguese science, and it is our priority making them easily found by the research community in need. Measures will be taken to push BioData.pt services and resources to the entry page of the website. It was also recognised that the national research community needs a more diversified offer of bioinformatics and data management programmes, and BioData.pt can give a precious contribution to that. A reflection on how BioData.pt responds to these needs will continue within this research infrastructure.

Changing faces

This moment was also the time for Miguel Rocha to take over the Presidency of the Administration Council of BioData.pt. Miguel succeeds José Pereira Leal, to whom we are very grateful for so many years leading this Research Infrastructure. José has been an inspiration for all of

BioData.pt - Portuguese Infrastructure of Biological Data

us and we sincerely hope that he continues with us in a journey that we expect to be pleasant and long-lasting. Irina Moreira is now Vice-President together with Cátia Pesquita, and the other two board members are Cymon Cox and José Pereira Leal.



Technical meeting working session