

BioData.pt

Rede Portuguesa de Informação Biológica



REPÚBLICA
PORTUGUESA



CIÊNCIA, TECNOLOGIA
E ENSINO SUPERIOR

Big (Bio) Data

- Biological research is becoming increasingly data intensive - its now big data science
 - genomics, proteomics, lipidomics, etc.
 - high throughput image generation and analyses
 - robotisation of the laboratory
- There is a lot of data that is used **once** by the team that produces it and then lost
- **These data need to be made available to the community integrated into a larger whole**

These data need to be made available to the community integrated into a larger whole

- education
- curation
- interoperability
- tools
- hardware

Goals of infrastructure

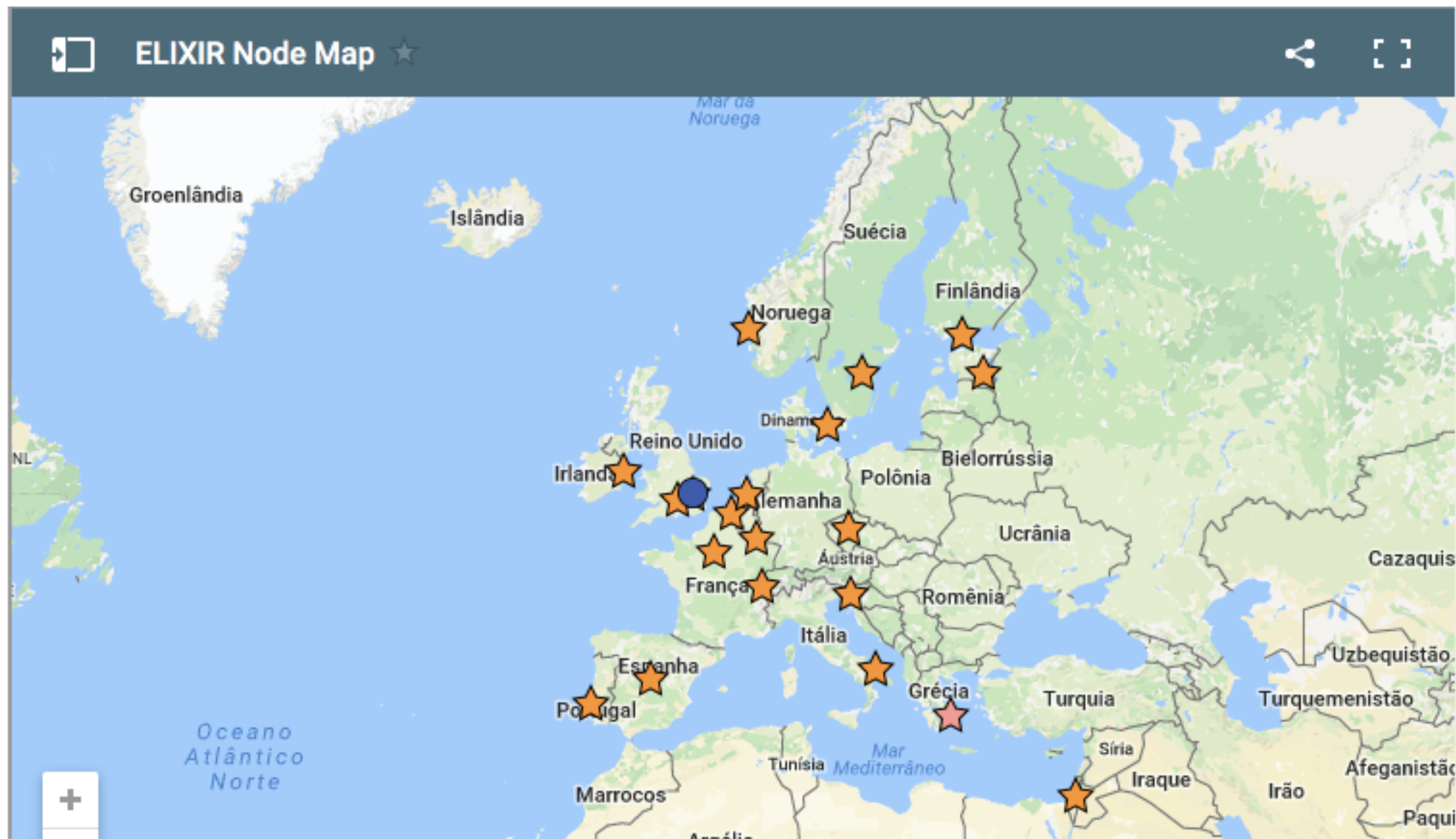
The purpose of [BioData.pt](https://www.biodata.pt) and its international counterpart ELIXIR is to construct and operate a sustainable infrastructure for biological information in Europe to support life science research and its translation to medicine and the environment, the bio-industries and society.

Services

- Common Data infrastructure in PT & EU
- Smart specialisation
- Bioinformatics Training for Academia & Industry
- Entrepreneurship/Industry Program
- Consultancy to Academia & Industry

ELIXIR

- ELIXIR coordinates and develops life science resources across Europe so that researchers can more easily find, analyse and share data, exchange expertise, and implement best practices.



BioData.pt

emerging infrastructure for biological information of the
portuguese infrastructure roadmap



Universidade do Minho



P-BIO

Portugal's Biotechnology Industry Organization
Associação Portuguesa de BioIndústria



INSTITUTO
DE TECNOLOGIA
QUÍMICA E BIOLÓGICA
(UNL)
Knowledge Creation



inesc id
lisboa



Champalimaud
Foundation



iBET



TÉCNICO LISBOA



CENTRE OF MARINE SCIENCES



Education



Who we are

The GTPB runs face-to-face Bioinformatics Training Courses regularly at the Instituto Gulbenkian de Ciência since 1999. Up to now, **more than 5000 course participants** have acquired practical skills that they can use with a high degree independence. The Programme consists in a series of short, intensive hands-on courses delivered and fully documented in English. The design of the courses is based on sets of carefully chosen exercises, flanked by short lectures and participative interaction sessions.

[General access conditions](#)

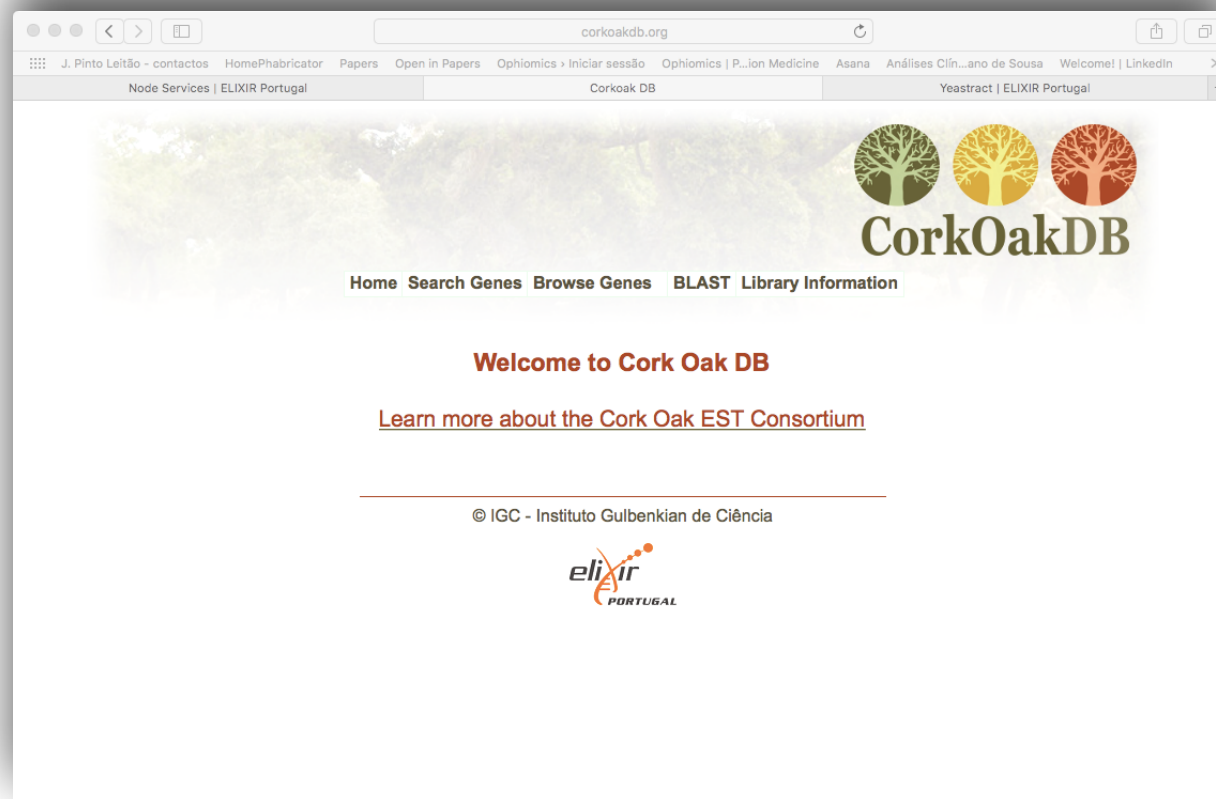
Courses in 2017

Last update, 2017/06/30

From / To (dates)	Application Deadline	Application Status	Title
Mar 6 - Mar 10	Feb 25	Closed	PDA17 Proteomics Data Analysis <i>with Lennart Martens, Harald Barsnes and Astrid Guldbrandsen</i>
Mar 13 - Mar 17	Mar 6	Closed	PGDH17 Population Genetics and Demographic History: model-based approaches <i>with Mark Beaumont, Lounes Chikhi, Willy Rodriguez, Bárbara Parreira and Vitor Sousa</i>
Apr 10 - Apr 13	Apr 3	Closed	ABSTAT17 Advanced Biostatistics <i>with Lisete Sousa and Carina Silva</i>
Apr 17 - Apr 20	Apr 10	Closed Fully Booked!	ADER17 Analysis of Differential Expression

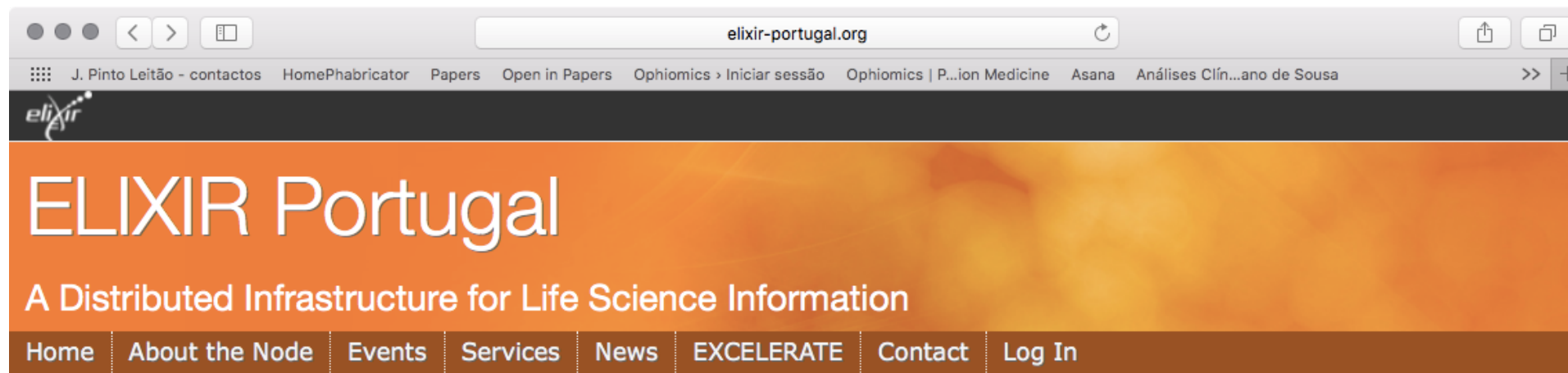
Curation

- Data curation, integration and availability for specific application domains
 - **Woody plants**
 - industrial microbiology
 - **Marine resources**
 - **Neurobiology/behaviour**



Interoperability

- Ontologies, standards, sops, federation, etc..



Plant Experimental Assay Ontology (PEAO)

The Plant domain has been the subject of several attempts to structure and formally define terms and corresponding relations, such as their anatomical features, developmental stages, and the application of particular experimental procedures to a biological problem. However, a focus on experimental assays in order to describe the whole experimental procedure, to the best of our knowledge, has only been attempted in the context of a very general description based on classical views of the scientific method. In this study, we focus on the development and proposal of an ontology dedicated to the description of these experimental procedures, regardless of the scientific questions that prompted the assays. This ontology includes entities from three distinct realms (biological, physical and data), which include both experimental products, their relations and the protocols describing their manipulation. The final outcome is a useful and comprehensive ontology in the plant domain, that will be useful for data integration and querying heterogeneous data.

Platform:

-Data

Website:

[Plant Experimental Assay Ontology Bitbucket](#)

Responsible organisation:

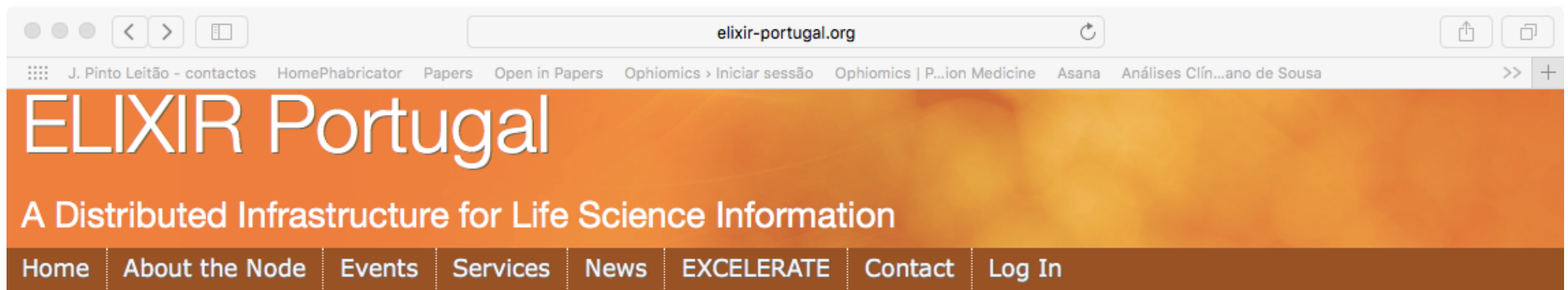
[iBET](#)

Name:

Inês Chaves

Tools

- map, rate, group, visibility, link, sustainability..



Node Services

For the ELIXIR-PT services on bio.tools click [here](#).

Elixir Platform▲	Node Service Name	EDAM Topics	Responsible organisation
-Compute	ELIXIR PT Computing Services	Data Handling	INESC-ID
-Data	CorkOakDB	Plants , Genomics	Instituto Gulbenkian de Ciência
-Data	Plant Experimental Assay Ontology (PEAO)	Ontology and terminology , Plants	iBET
-Tools	PHYLOViZ	Phylogenetics	INESC-ID , IST , IMM
-Tools	Yeasttract	Gene regulatory networks , Transcription factors and	INESC-ID , IBB

Major challenge

- Soft infrastructure - its about people rather than machines
- infrastructure funding focuses on physical infrastructure

Problems and opportunities



A role for the funding agencies (FCT!):

- demand data sustainability, data management plans
- include vouchers in research projects for PT e-infrastructures for data management plan
- final evaluation - need to demonstrate implementation of data plan

Problems and opportunities

- Supporting scientific research is obvious

but

- Supporting/engaging innovation, start ups and established companies harder
 - virtual incubator for (bio)digital start ups
 - (attempt) at engaging companies from the start

Thank you

José Pereira-Leal

jleal@ophiomics.com

Impact

- Scientific
 - enable research in bio-domains
 - train (short/long) bioinformatics professionals
- Socioeconomic
 - enable industrial research
 - supports biodata-intensive company formation/activity