

12.º FÓRUM  
GESTÃO DE DADOS  
DE INVESTIGAÇÃO

27-28 DE NOVEMBRO 2025  
INSTITUTO POLITÉCNICO DE BRAGANÇA

27 e 28 DE NOVEMBRO 2025  
INSTITUTO POLITÉCNICO DE BRAGANÇA



Organização



Re.Data

Apoio



INSTITUTO POLITÉCNICO DE BRAGANÇA  
BRAGANÇA POLYTECHNIC UNIVERSITY

Financiamento



REPÚBLICA  
PORTUGUESA



Financiado pela  
União Europeia  
NextGenerationEU

# Ask the Data no Dados.IPB

Adília Alves, [aialves@ipb.pt](mailto:aialves@ipb.pt)

Clarisse Pais, [clarisse@ipb.pt](mailto:clarisse@ipb.pt)

# Dados.IPB

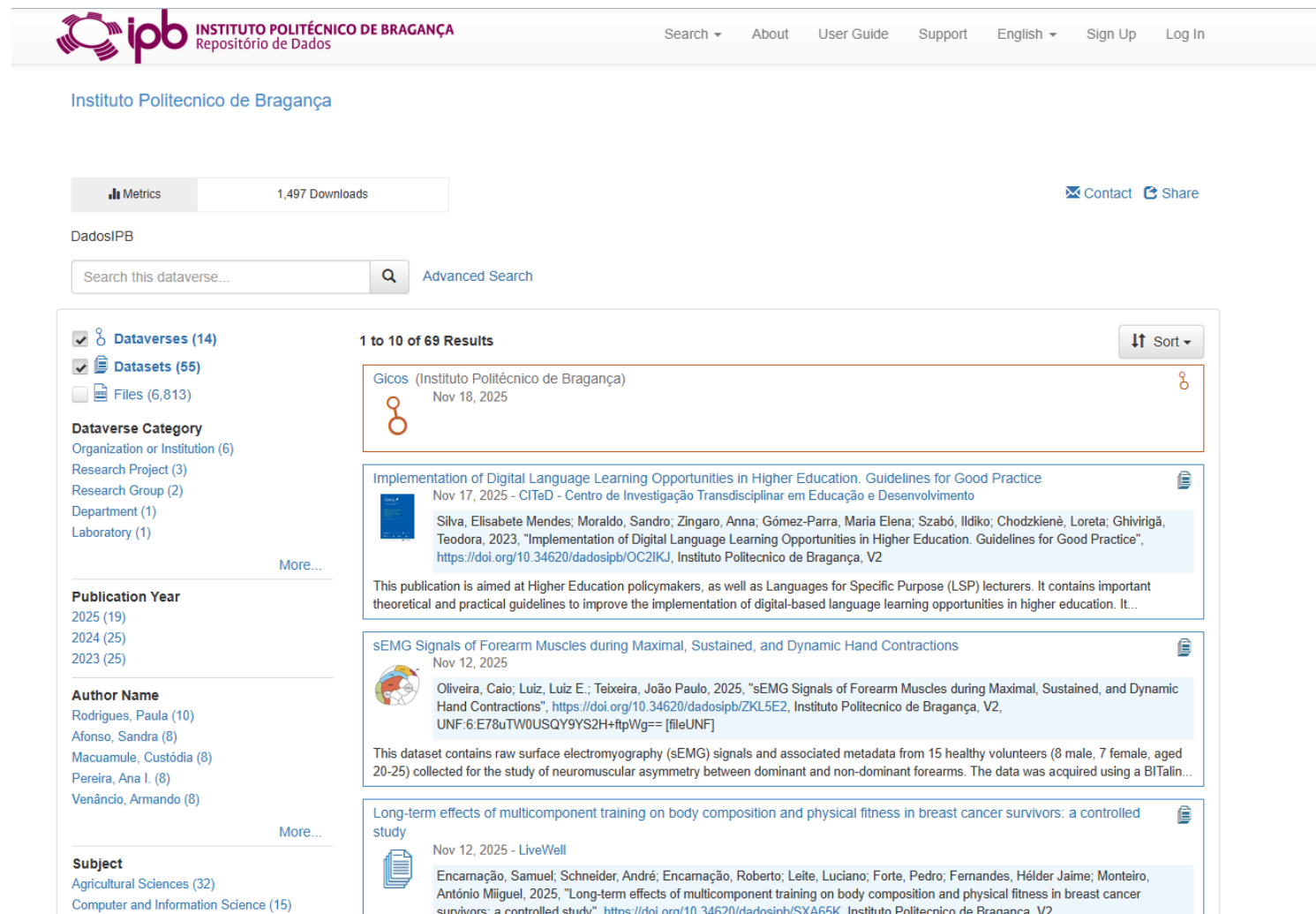
Criação: Outubro de 2018

Produção: 29 de novembro de 2018

Daverses: 14

Datasets: 55

Files: 6813



The screenshot displays the Dados.IPB website interface. At the top, the IPB logo and name are visible, along with navigation links for Search, About, User Guide, Support, English, Sign Up, and Log In. Below the header, the site title 'Instituto Politécnico de Bragança' is shown. A metrics bar indicates '1,497 Downloads'. A search bar with the placeholder 'Search this dataverse...' and an 'Advanced Search' link is present. On the left, a sidebar lists filters: 'Dataverses (14)', 'Datasets (55)', and 'Files (6,813)'. Below these are 'Dataverse Category' (Organization or Institution (6), Research Project (3), Research Group (2), Department (1), Laboratory (1)) and 'Publication Year' (2025 (19), 2024 (25), 2023 (25)). The 'Author Name' section lists: Rodrigues, Paula (10), Afonso, Sandra (8), Macuamule, Custódia (8), Pereira, Ana I. (8), and Venâncio, Armando (8). The 'Subject' section lists: Agricultural Sciences (32) and Computer and Information Science (15). The main content area shows '1 to 10 of 69 Results'. The first result is 'Gicos (Instituto Politécnico de Bragança)' dated Nov 18, 2025. The second result is 'Implementation of Digital Language Learning Opportunities in Higher Education. Guidelines for Good Practice' dated Nov 17, 2025, by Silva, Elisabete Mendes; Moraldo, Sandro; Zingaro, Anna; Gómez-Parra, Maria Elena; Szabó, Ildiko; Chodzkienė, Loreta; Ghivirigă, Teodora, 2023. The third result is 'sEMG Signals of Forearm Muscles during Maximal, Sustained, and Dynamic Hand Contractions' dated Nov 12, 2025, by Oliveira, Caio; Luiz, Luiz E.; Teixeira, João Paulo, 2025. The fourth result is 'Long-term effects of multicomponent training on body composition and physical fitness in breast cancer survivors: a controlled study' dated Nov 12, 2025, by Encarnação, Samuel; Schneider, André; Encarnação, Roberto; Leite, Luciano; Forte, Pedro; Fernandes, Hélder Jaime; Monteiro, António Miguel, 2025.

# Dados.IPB

Os dados armazenados nesta plataforma abrangem diversas áreas de conhecimento e estão guardados em distintos formatos (ex: .xlsx, .docx, .pdf, .jpg, .tab, .csv).

- Recentemente é possível:
  - pré-visualizar ficheiros armazenados no Dados.IPB sem fazer o download.
  - fazer perguntas aos ficheiros .tab e às imagens
- Foram utilizadas ferramentas externas (<https://guides.dataverse.org/en/latest/admin/external-tools.html>) e desenvolvidas novas ferramentas.

# Dados.IPB - Ask the Data

A extensão *Ask the Data*

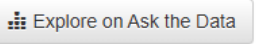
(<https://github.com/IQSS/askdataverse/tree/main/askthedata>) é uma

ferramenta externa ao Dataverse que permite fazer perguntas em

linguagem natural aos dados contidos

nas tabelas do Dados.IPB (dados

tabulares, com extensão .tab).

 Explore on Ask the Data

Tell me what you want to know

What is this data about?

Answer please

The data appears to be a dataset related to a scientific or experimental study, likely in the field of health, fitness, or physiology. Here are some observations based on the column names and the first few rows:

- Column Names:** The first row contains a long list of column names that include various measurements and attributes such as: - Age - Body Weight (BW) - Height - Body Mass Index (BMI) - Body Fat Percentage (%BF) - Lean Mass (LM) - Basal Metabolic Rate (BasalMET) - Various other physiological measurements (e.g., Vfat, ULS, LLS, etc.)
- Data Structure:** The data appears to be structured in a way that each participant (or observation) has multiple measurements taken at different times or conditions (indicated by the suffixes like `\_1` and `\_2`).
- Participant Information:** The first column seems to contain identifiers for participants (e.g., "Part.1", "Part.2", etc.), and there is a "GROUP" column that may indicate the experimental group to which each participant belongs.
- Experimental Context:** The presence of the term "Experimental" in the "GROUP" column suggests that this data may be part of a controlled study where participants are subjected to specific conditions or treatments.
- Missing Values:** The presence of "nan" values indicates that there are missing data points in the dataset.

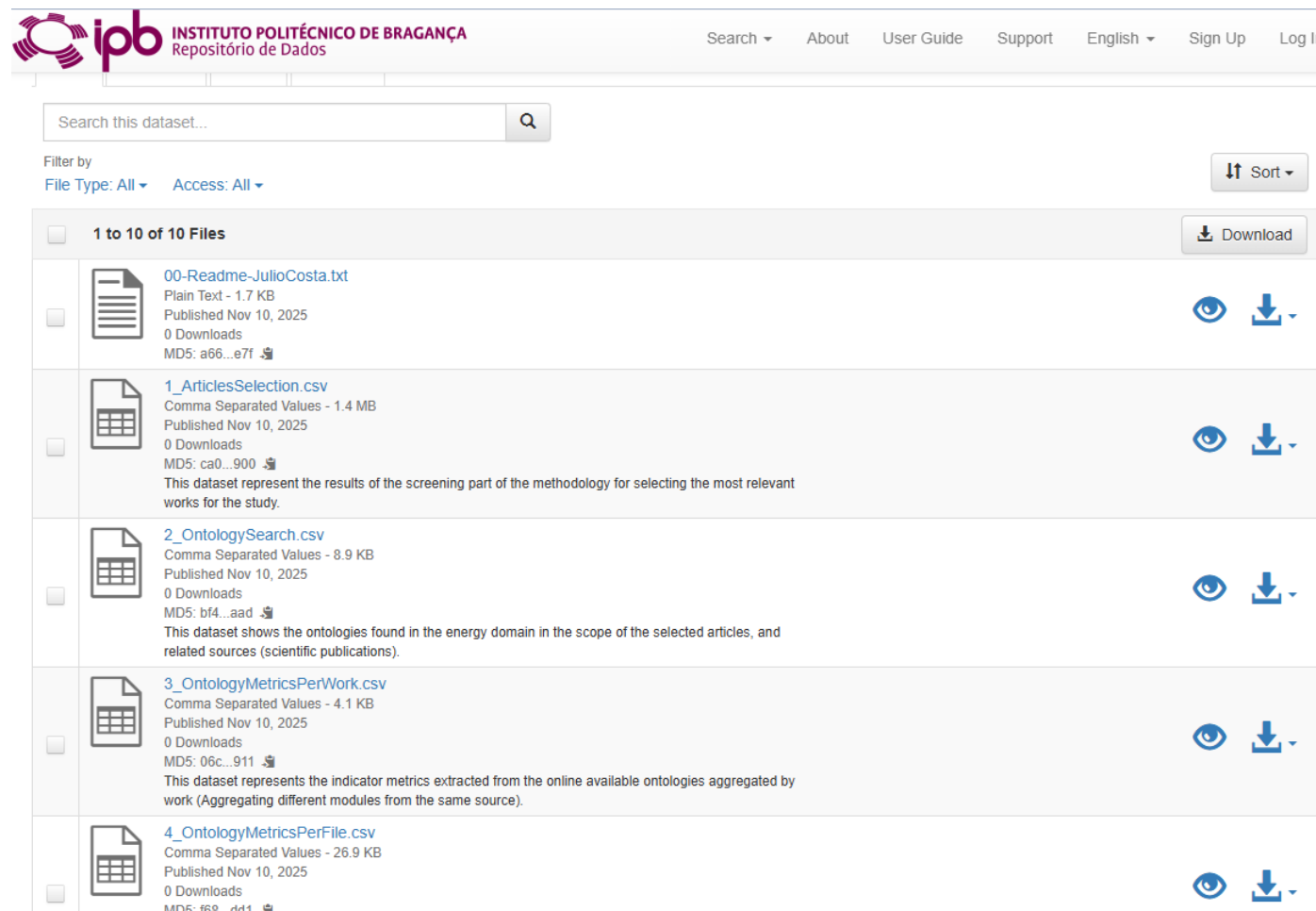
Overall, this dataset likely contains detailed physiological measurements from participants in an experimental study, possibly aimed at understanding the effects of certain interventions on health or fitness metrics.

41598

ID;GROUP;Age	1;Age
2;BW_1;BW_2;Height_1;Height_2;BMI_1;BMI_2;%BF_1;%BF_2;LM_1;LM_2;%Wat_1;%Wat_2;BasalMET_1;BasalMET_2;Vfat_1;Vfat_2;ULS_1;ULS_2;ULS_IntArm_1;ULS_I	
Part.1;Experimental	65;65;64.6;61.3;1.62;1.62;25;23;35.7;30.8;39.4;42.5;46.4;44;1250;1244;8;8;5;22;3;16;16;27;-5;4;6;14;4.29;3.61;55;125;.....
Part.2;Experimental	68;68;50.8;50;1.5;1.5;23;22;26.1;21.3;35.6;36;52.5;54;1112;1129;8;6;16;29;10;26;21;28;0;5;7;15;4.02;3.36;59;154;.....
Part.3;Experimental	54;54;78.4;77.1;64.1;64.29;29;39;32.1;45.4;45;45;44;1448;1244;9;9;16;33;9;30;22;28;4;2;3;3;5;03;3;56;27;181;.....

# Dados.IPB - File Previewers

Ferramentas que exibem o conteúdo de ficheiros - incluindo áudio, html, anotações, imagens, PDF, texto, vídeo, dados tabulares, folhas de cálculo, GeoJSON - permitindo que sejam visualizados sem descarregar o ficheiro (<https://github.com/gdcc/dataverse-previewers>).



The screenshot displays the IPB Data Repository website. At the top, there is a navigation bar with the IPB logo, the text "INSTITUTO POLITÉCNICO DE BRAGANÇA Repositório de Dados", and links for Search, About, User Guide, Support, English, Sign Up, and Log In. Below the navigation bar is a search bar labeled "Search this dataset...". Under the search bar, there are filters for "Filter by" (File Type: All, Access: All) and a "Sort" button. A "Download" button is also present. The main content area shows a list of files, with the first five files visible:

File Name	File Type	Size	Published	Downloads	MD5	Preview Icon	Download Icon
00-Readme-JulioCosta.txt	Plain Text	1.7 KB	Nov 10, 2025	0	a66...e7f		
1_ArticlesSelection.csv	Comma Separated Values	1.4 MB	Nov 10, 2025	0	ca0...900		
2_OntologySearch.csv	Comma Separated Values	8.9 KB	Nov 10, 2025	0	b4...aad		
3_OntologyMetricsPerWork.csv	Comma Separated Values	4.1 KB	Nov 10, 2025	0	06c...911		
4_OntologyMetricsPerFile.csv	Comma Separated Values	26.9 KB	Nov 10, 2025	0	f68...d41		

# Dados.IPB - File Previewers

## 00\_readme\_Olive\_Mendes.txt

This file is part of "Olive Cultivar classification based on leaf images".

Version 2.0

### File Citation

Mendes, Joao; Lima, José; Costa, Lino; Rodrigues, Nuno; Pereira, Ana I., 2023, "Olive Cultivar classification based on leaf images", <https://doi.org/10.34620/dadosipb/TVYE8K>, Instituto Politécnico de Bragança, V2; 00\_readme\_Olive\_Mendes.txt [fileName]

[Cite Data File](#)

Learn about [Data Citation Standards](#).

### Dataset Citation

Mendes, Joao; Lima, José; Costa, Lino; Rodrigues, Nuno; Pereira, Ana I., 2023, "Olive Cultivar classification based on leaf images", <https://doi.org/10.34620/dadosipb/TVYE8K>, Instituto Politécnico de Bragança, V2

[Cite Dataset](#)

Learn about [Data Citation Standards](#).

Access File

Contact Owner

Share

File Metrics

0 Downloads

Preview

Metadata

Versions

Open in New Window

```
Title: Olive Cultivar Classification Based on Leaf Images
Area: Agriculture / Computer Vision / Plant Phenotyping


Description:
This dataset contains more than 4,000 high-resolution RGB images of olive leaves from four cultivars – Cobranço
Sampling was conducted across all seasons to capture phenological variability, including differences due to wat
The collection process followed a randomized methodology: both current-year and older leaves were sampled, from
All samples were collected in Suções (41.49°N, 7.26°W), at an average altitude of 350 meters, within the munic


File structure:
- /images/ - folder containing all olive leaf images (RGB)
- /labels/ - folder containing cultivar labels or metadata (if applicable)
- README.txt - this file

Dataset DOI:
https://doi.org/10.34620/dadosipb/TVYE8K
```






# Dados.IPB - File Previewers

 INSTITUTO POLITÉCNICO DE BRAGANÇA  
Repositório de Dados

Search ▾ About User Guide Support English ▾  Dataverse Admin 37 ▾

Preview Metadata Versions

 Preview ▾  Open in New Window



Find 'Bactrocera Oleae' exemplars

Previewers originally developed by [QDR](https://github.com/QualitativeDataRepository/dataverse-previewers) and maintained at <https://github.com/QualitativeDataRepository/dataverse-previewers>. Feedback and contributions welcome.

O File Preview das imagens permite fazer zoom

# Dados.IPB – Ferramenta externa desenvolvida pelos Serviços Documentação e Bibliotecas IPB

- 1 - Treino de modelo YOLO para identificação de exemplares *Bactrocera Oleae*
- 2 - Exportação de modelo para formato onnx
- 3 - Modelo de deteção executado exclusivamente no browser
- 4 - Importação de funcionalidade como ferramenta externa no dataverse

```
{
  "displayName": "OliveFlyDetection",
  "description": "A tool that finds Bactrocera Oleae exemplars.",
  "toolName": "AskTheData",
  "scope": "file",
  "types": [
    "query"
  ],
  "toolUrl": "https://dados.ipb.pt/askthedata/",
  "contentType": "image/jpeg",
  "httpMethod": "GET",
  "toolParameters": {
    "queryParameters": [
      {
        "fileId": "{fileId}"
      },
      {
        "datasetPid": "{datasetPid}"
      },
      {
        "siteUrl": "{siteUrl}"
      }
    ]
  }
}
```



# Dados.IPB – Ferramenta Externa Desenvolvida nos SDBibliotecas

Pesquisa nas imagens de insetos pertencentes à família *Bactrocera Oleae*

Trabalho futuro para esta ferramenta: Deteções mais englobantes (ex: detetar insetos)




Copyright © 2025


Powered by  **Dataverse**  
Project v. 5.12.1 build 1122-cf9043

About the platform

# Trabalho Futuro – Recuperação de Informação com Base no Conteúdo e nos Metadados


**Adília Isabel Domingues da Cruz Alves** 15:28

Data retrieval systems are beginning to adopt machine learning models to return information according to files content. Maybe it is a future trend to dataverse... In Bragança (Portugal) we are developing a external tool to detect insects in files.

**Philip Durbin** 15:30


Is there documentation for one of these systems that I can look at? I'm still not sure I fully understand.

Once you know the content of a file (thanks to ML), shouldn't you update the file's metadata to say what's in it (e.g. bees 🐝)? And make it searchable? q=mlDeterminedFileContent:bees 15:33

**Adília Isabel Domingues da Cruz Alves** EDITED 15:35


You can always try to 'read' the files without update metadata.  
Photos have much information that can be searched (ex: Has the photo insects? What are the colors of the image? Has the image olive trees?....)

<http://coveo.com/blog/ai-information-retrieval/>

**Philip Durbin** 15:43

Interesting article, thanks.


You seem to be saying that the deep learning system has been trained on the images. That means you can ask questions on the fly like "Does the photo have insects?", etc. 15:44

 You

Are you aware that you can bring your own search service to Dataverse? It's new as of 6.7: 15:44

<https://guides.dataverse.org/en/6.8/developers/search-services.html>

At the community meeting we demo'ed "AI Search", for example. But it could be any kind of search. 15:45

 You

Also, we should probably be talking in a channel (maybe #community?) so others can jump in. 😊 15:45

