

# BioData.pt

## Plant Phenotyping-Genotyping Data Management Workshop



# The plant phenotyping data standard MIAPPE v1.1

Daniel Faria

# What is MIAPPE?

- **Minimum Information About a Plant Phenotyping Experiment**
- Metadata standard for plant phenotyping / phenomics
- Details the metadata fields needed to describe a plant phenotyping dataset so that it is readily interpretable and self-contained

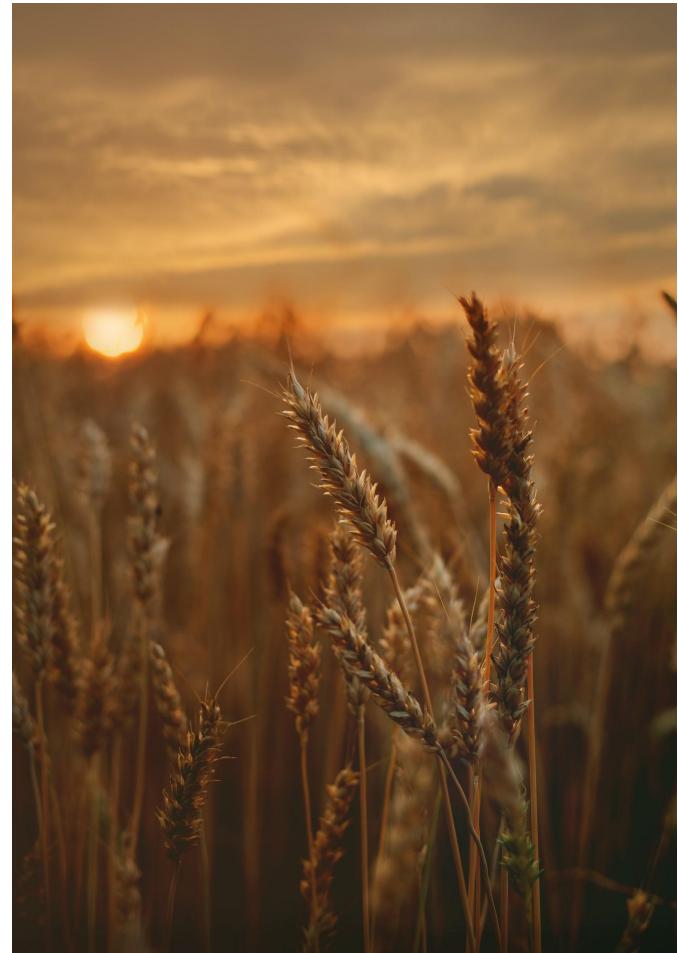


Photo by [Tetyana Kovyrina](#) from [Pexels](#)

# What is MIAPPE?

- Covers both woody and herbaceous plants, either cultivated or wild, in any growth environment
- Includes mandatory and optional fields
- Details how each field should be filled, including ontology recommendations

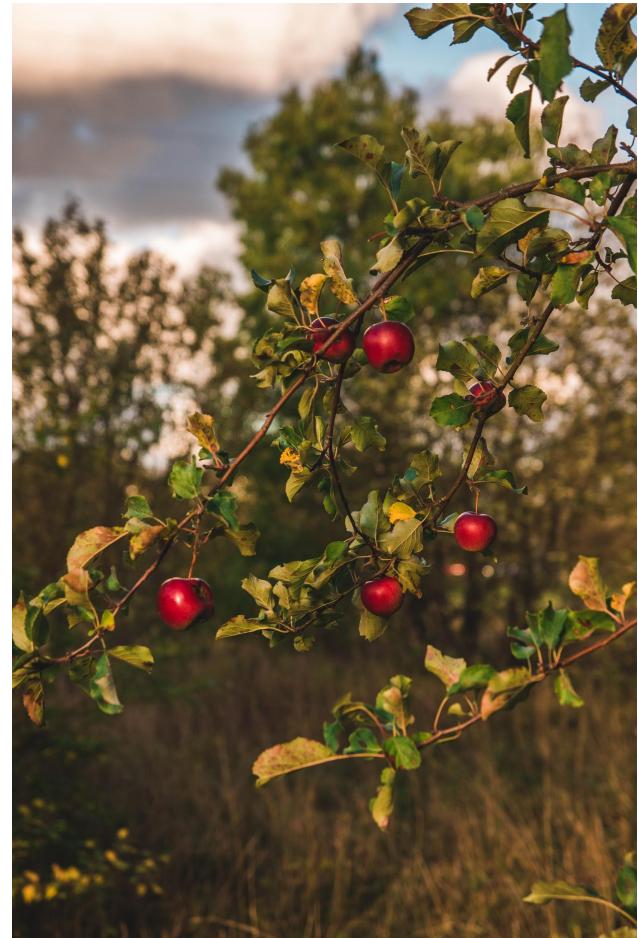
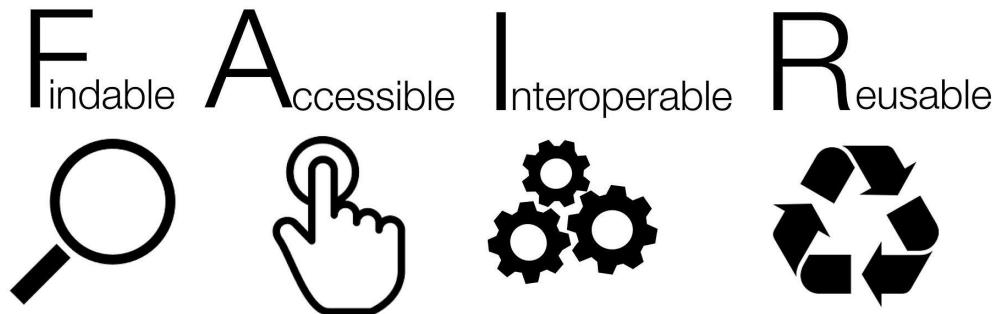


Photo by [Kristina Paukshtite](#) from [Pexels](#)

# Why do we need MIAPPE?

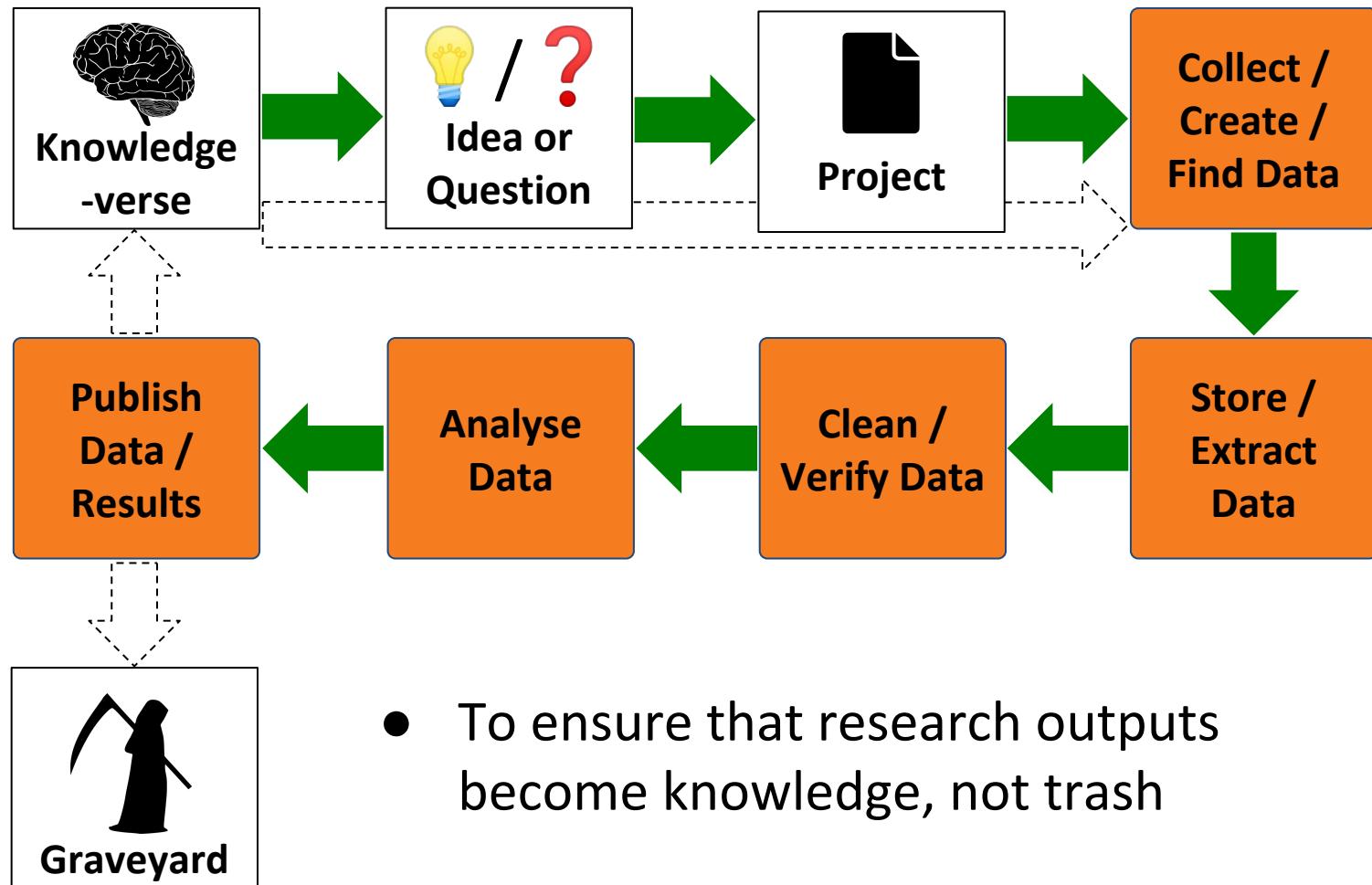


By Sangya Pundir, CC BY-SA 4.0

<https://commons.wikimedia.org/w/index.php?curid=53414062>

- MIAPPE is key for enabling FAIR plant phenotyping data:
  - Structured metadata fields and use of ontologies enable findability and interoperability, which in turn enable reusability

# Why do we need MIAPPE?



- To ensure that research outputs become knowledge, not trash

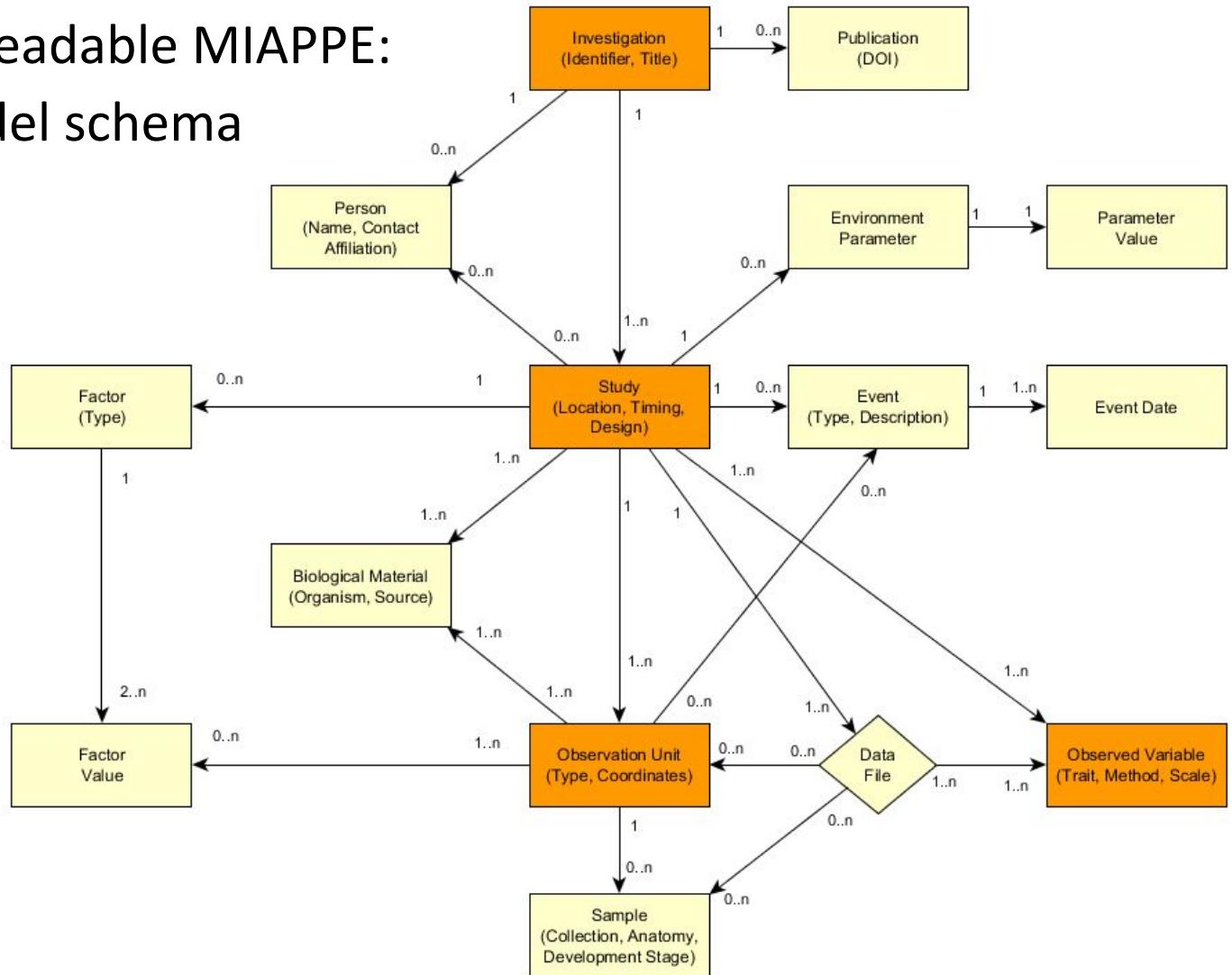
# How is MIAPPE formatted?

- Human-readable MIAPPE: checklist

MIAPPE				
MIAPPE Check list	Definition	Example	Format	Cardinality
<b>Investigation</b>	Investigations are research programmes with defined aims. They can exist at various scales (for example, they could encompass a grant-funded programme of work, the various components comprising a peer-reviewed publication, or a single experiment).			1 per MIAPPE submission
<b>Investigation unique ID</b>	Identifier comprising the unique name of the institution/database hosting the submission of the investigation data, and the accession number of the investigation in that institution.	EBI:12345678	Unique identifier	0-1
<b>Investigation title</b>	Human-readable string summarising the investigation.	Adaptation of Maize to Temperate Climates: Mid-Density Genome-Wide Association Genetics and Diversity Patterns Reveal Key Genomic Regions, with a Major Contribution of the Vg12 (ZCN8) Locus.	Free text (short)	1
<b>Investigation description</b>	Human-readable text describing the investigation in more detail.	The migration of maize from tropical to temperate climates was accompanied by a dramatic evolution in flowering time. To gain insight into the genetic architecture of this adaptive trait, we conducted a 50K SNP-based genome-wide association and diversity investigation on a panel of tropical and temperate American and European representatives.	Free text	0-1
<b>Submission date</b>	Date of submission of the dataset presently being described to a host repository.	2012-12-17	Date/Time (ISO 8601, optional time zone)	0-1
<b>Public release date</b>	Date of first public release of the dataset presently being described.	2013-02-25	Date/Time (ISO 8601, optional time zone)	0-1
<b>License</b>	License for the reuse of the data associated with this investigation. The Creative Commons licenses cover most use cases and are recommended.	CC BY-SA 4.0, Unreported	Unique identifier	0-1
<b>MIAPPE version</b>	The version of MIAPPE used.	1.1	Version number	1
<b>Associated publication</b>	An identifier for a literature publication where the investigation is described. Use of DOIs is recommended.	doi:10.1371/journal.pone.0071377	DOI	0+

# How is MIAPPE formatted?

- Human-readable MIAPPE:  
data model schema



# How is MIAPPE formatted?

- Computer-interpretable MIAPPE:
  - Ontology (OWL) encoding of the checklist and data model
  - ISA-Tab configuration

# MIAPPE: Investigation

- Top level section loosely corresponding to a research project, which details:
  - Generic information about the project: unique ID, title, description
  - Information about the dataset itself: submission date, public release date, license, MIAPPE version
  - Associated (scientific) publications [for simplicity, all should be listed at this level]

# MIAPPE: Study

- Second level section loosely corresponding to an experiment, which details:
  - Generic information about the experiment: unique ID, title, description
  - Temporal and spatial placement: start date, end date, contact institution, geographical coordinates [by convention a study has a single location]
  - Experimental setting: description and type of growth facility, cultural practices

# MIAPPE: Study

- More information on the experiment setting:
  - Description and type of experimental design
  - Observation unit level hierarchy and description
  - Map of experimental design



Plant icon adapted from [Shannon E Thomas at to \[icon\]](#), CC-BY

# MIAPPE: Person

- Details information about people associated with the investigation (at least 1 person should be listed as contact) or a particular study (option):
  - Contact information: name, email, ID
  - Role
  - Affiliation



Photo by [Naveen Annam](#) from [Pexels](#)

# MIAPPE: Data File

- References and describes a data file that is external to the MIAPPE submission:
  - Link
  - Description
  - Version



By OnePoint Services, CC BY 2.0

<https://www.flickr.com/photos/78830297@N05/14740342074/>

# MIAPPE: Biological Material

- Describes the provenance of the plants used in the study, such as a seed stock or a parent plant that was cloned
- Identification through seed lot number (ID) and/or geographical coordinates (for trees)
- Characterization through NCBI Taxon ID, taxon name, and preprocessing description
- Material source enables a second level of provenance

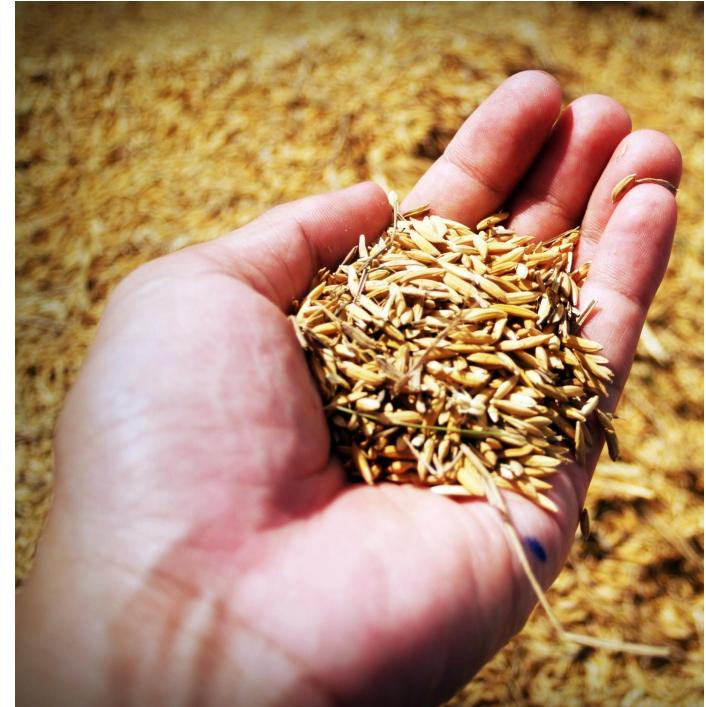


Photo by [icon0.com](#) from [Pexels](#)

# Biological Material vs. Material Source

- The seed stock in your institute used in an experiment ⇒ biological material
- The original seed stock from which your institute acquired the seeds ⇒ material source
- The parent plant cloned for an experiment ⇒ biological material
- The parent(s) or seed stock of the parent plant ⇒ material source
- Each wild tree used in an experiment ⇒ biological material (no material source)

# MIAPPE: Environment

- Details the (fixed) environment parameters of the experiment:
  - Paired environment parameter and value
- Supported by list of environment parameters in MIAPPE appendix

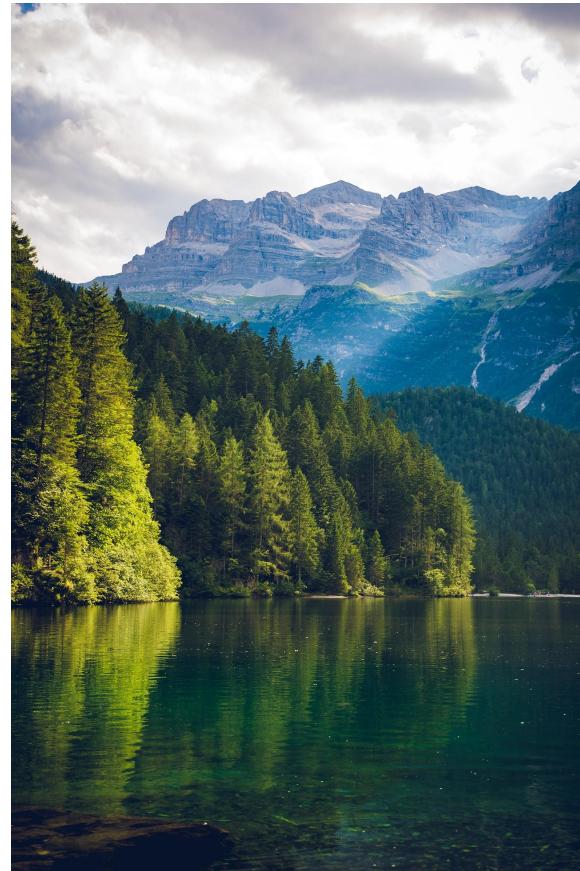


Photo by [Philip Ackermann](#) from [Pexels](#)

# MIAPPE: Experimental Factor

- Details the object(s) of the study, i.e., the aspects the effect of which we are trying to assess:
  - Factor type, description and values (2+)
- Supported by list of experimental factors in MIAPPE appendix



Photo by [Pixabay](#) from [Pexels](#)

# MIAPPE: Event

- Expresses a discreet occurrence of something natural or artificial, at a specific time
- Includes: type, accession number, description, and date
- Can apply to the whole study, or specific observation units



Photo by [Ralph W. Lambrecht](#) from Pexels

# MIAPPE: Observation Unit

- The object of individual observations, which can be the whole study, a plot, a block, a pot, or a single plant
- Includes: ID, type (or level), external ID, spatial distribution, and factor value



Photo by [Max Andrey](#) from [Pexels](#)

# MIAPPE: Sample

- Part of a plant, derived from an Observation Unit
- Includes: ID, plant anatomical structure, plant development stage, description, collection date, and external ID
- May be required for traceability of materials, namely in molecular studies



Photo by [Martin Lopez](#) from [Pexels](#)

# MIAPPE: Observed Variable

- Describes what plant attribute (trait) was recorded during the experiment and how it was recorded (method and scale)
- Includes several fields for variable, trait, method and scale to support both Crop Ontology use and custom use
- Time scale details the time units of recorded observations



Photo by [Alena Koval](#) from [Pexels](#)

# To Sample or not to Sample

- Use of samples is recommended for traceability of materials, especially when complex processing is involved (e.g. molecular studies)
- The use of samples for sub-plant level observations is unnecessary in many cases, as the sampling procedure and plant part can be described in the Observed Variable (e.g. “average fruit weight”, “length of first leaf”)
- It may be necessary to use samples if the plant part is not representative of the whole plant (e.g. when comparing cork thickness of all branches of a tree)

# How can I submit a MIAPPE dataset

- The MIAPPE [ISA-Tab](#) configuration is the only implementation currently ready to use, and is accepted by a number of journals
- There is a [spreadsheet template](#) available, but intended mainly for training purposes
- The Breeding API (BrAPI) has just been updated to be fully MIAPPE compliant, and we are planning on extending it to support MIAPPE dataset submission
- We are working on extending FAIRDOM to enable MIAPPE submission

# Acknowledgments

Evangelia Papoutsoglou, ELIXIR-NL

Cyril Pommier & Anne-Françoise Adam Blondon, ELIXIR-FR

Célia Miguel & Inês Chaves, BioData.pt / ELIXIR-PT



E-mail: [info@biodata.pt](mailto:info@biodata.pt)  
[www.biodata.pt](http://www.biodata.pt)

