Data standards for plant phenotyping: MIAPPE and its implementations


To cite this version:
Cyril Pommier, Guillaume Cornut, Thomas Letellier, Célia Michotey, Pascal Neveu, et al.. Data standards for plant phenotyping: MIAPPE and its implementations. 26. Plant and Animal Genome Conference (PAG XXVI), Jan 2018, San Diego, Californie, United States. pp.24 slides. hal-02789754

HAL Id: hal-02789754
https://hal.inrae.fr/hal-02789754
Submitted on 5 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives| 4.0 International License
MIAPPE Plant Phenotyping Data standard

Minimum Information About Plant Phenotyping Experiment and Its Implementations

PLANT PHENOTYPING DATA STANDARD NEED

• Heterogeneous Datasets
  - Single field trial
  - Phenotyping Field networks on multiple years
  - Automated Greenhouse

• Heterogeneous measure types
  - Experimenter measures and notations
    - Low throughput sensors and measurement devices
    - Yield, plant height, disease notations, NIRS
  - Automated measures
    - High throughput Greenhouses or Field
    - Drones, Phenomobiles, sensor networks
  - Images, Multispectral, LIDAR, NIRS

• Interoperability need
  - Intra dataset interoperability
    - Phenotyping networks consolidation
    - Project integration
  - Inter dataset interoperability
    - Large scale breeding
    - Genetic analysis
    - Climate change studies
    -...

• Interoperability and sharing principle
  FAIR

FAIR Data Principles

Minimum Information About Plant Phenotyping Experiment

- Standard
  - Data exchange and traceability
  - Repositories
  - Analysis

- Input and output for analysis pipelines developed for phenotyping and genetic

Measures for interoperability of phenotypic data: minimum information requirements and formatting

Towards recommendations for metadata and data handling in plant phenotyping

Cyril Pommier / MIAPPE Plant Phenotyping Data standard

MIAPPE v1.1 Overview

- Relies on existing standards
  - MCPD, Crop Ontology, Expression (MIAME), Metabolomic (MSI)

- Study Metadata
  - Identification
  - Title, Publications
  - Data files URL

- Timing and Location
  - GPS

- Experimental Design
  - Observation Unit
    - Object being measured
    - Level/types: organ, plant, micro plot, ...
  - Genotype, Plant material identification
  - Treatments, Factors combinations

- Environment
  - Treatments
  - Cultural Practices

- Improvement in progress

MIAPPE v1.1 Overview

- Biosource: Plant Material identification
- Multi Crop Passport Descriptor (MCPD) compliant
- www.cropontology.org
- ID Field are reused in MIAPPE
- Suitable for non Genbank collections: breeder varietal list, laboratory collection, ...

MIAPPE v1.1 Overview

- Observation Variables
  - Phenotype & environment
  - Trait: What is measured, eg plant height
  - Method: How is it measured, eg measuring tape from ground to apical bud
  - Scale: How is it observed, eg cm or notation scale.
  - Variable: trait + method + scale

MIAPPE v1.1 Overview

- MIAPPE Construction
  - Contribution
    - European Infrastructures : Elixir (Bioinformatics), Emphasis/EPPN (Phenotyping)
    - National Institutes: France, Germany, Poland, UK, Portugal, Slovenia, Nederland, Belgium, Italy
  - Steering Committee
    - Elixir: European Bioinformatic infrastructure
    - Emphasis: European plant phenotyping infrastructure
    - Bioversity International - CGIAR

- Current Versions
  - Version 1: Transplant/Elixir + EPPN
  - Version 1.1: Elixir, validation in progress
    - Documentation and model: Biologist friendly
    - Adaptations to new implementations
    - Proposal open for consultation: contribution welcomed
    - See http://www.miappe.org for providing feedback.

Cyril Pommier / MIAPPE Plant Phenotyping Data standard
MIAPPE: Databases

- Production databases
  - PHIS
- Cleaned data publication repositories
  - GnpIS
    - http://dx.doi.org/10.15454/1.4489666216568333E12
  - eDale
  - PlantPhenoDB
- Implementation
  - Internal
  - Web services API

MIAPPE databases & FAIR

- Findable: Indexed metadata
- Interoperable: Vocabularies/Ontologies, Plant material ID

MIAPPE File Archive

- ISA Tab for Phenotyping
  - Investigation/Study/Assay
  - Zip Archive
    - MIAPPE Metadata
    - Raw data
      - CSV
      - Images or binary files
      - Reference to image archive (URI/URL)
    - Elaborated data
      - CSV
      - Provenance

MIAPPE Web Service : Breeding API

- International collaboration
- Vision: To provide a standard Open API for easily, securely, and efficiently exchanging information between systems and applications that support breeding
- MIAPPE BrAPI alignment and compliance
MIAPPE Web Service: Breeding API

- Resources
  - http://paxel.org/
- Collection of specifications for data retrieval and exchange
- Servers implementations
  - CGIAR international network
  - Integrated Breeding Platform
  - Elixir, Emphasis
- Genotype
  - Germplasm
- Clients implementations
  - Flapjack: genotyping data visualization
  - GnpIS
- Ontology Widget
  - https://github.com/gnpis/ontolocity-widget
- R analysis pipelines

MIAPPE Semantic: Plant Phenotyping Experiment Ontology

- Joint initiative: Elixir, Emphasis, CropOntology, RDA
- Goals:
  - Enable computer interpretation of MIAPPE
  - Formally integrate MIAPPE and BrAPI
    - MIAPPE ontology
  - Provide context for publishing datasets in JSON-LD or RDF

MIAPPE Semantic: RDA RDFENO

- MIAPPE OWL Ontology
  - First draft
  - https://github.com/MIAPPE/MIAPPE-ontology
- BrAPI 2 MIAPPE RDF workflow
  - http://www.wheatis.org data standard page
  - http://ist.blogs.inra.fr/wdi/phenotypes/3795510
- Agroportal
  - Agronomy bioportal
  - http://agroportal.lirmm.fr/ontologies/PPEO
- Wheat dataset
  - http://dx.doi.org/10.15454/1.4489666216568333e12
- Future Query: Impact of summer temperature on yield

Adoption

- Plant community involved
  - Elixir (European bioinformatic infrastructure)
  - Emphasis (European Phenotyping infrastructure)
  - Bioversity international CGIAR
- Breeding API is Elixir official Phenotyping standard web service
- MIAPPE and BrAPI high collaboration
  - Adoption and compliance
- Data repositories and management tools
  - GnpIS https://urgi.versailles.inra.fr/gnpis/
  - eDale https://edal.ipk-gatersleben.de/
  - PlantPhenoDB at IPG PAS http://cropnet.pl/plantphenodb/
- In progress: COPO, Elixir plant databases (iBet, WUR, VIB, ...), Brassica Information Portal, ...

Perspectives

- Version 2: Emphasis, Elixir, ...
  - Environment
  - Sensor traceability
  - Phenoharmonis Workshop Montpellier may 2018
- Distributed search, MIAPPE enabled.
  - Data discovery
  - Elixir
  - WheatIS & Emphasis?
  - Open source software
- Dataset Validation
  - Elixir
  - File archive (ISA Tab) and BrAPI based

Acknowledgment & Questions

- IPG PAS
  - Anna Knapczyk-Kopacz
  - Pawel Krajewski
  - EMBL-EBI
  - Paul Kersey
  - iBet
  - Ines Owners
  - Celia M. Milgrom
  - IGB
  - Daniel Faria
  - Bioversity International CGIAR
  - Elisabeth Arnaud
  - Marie Angélique Laporte
  - iBra
  - Richard Eikelen
  - iBet
  - iBra