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
Plant Phenotyping Experiment Data

**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

**Heterogeneous, multiscale variables**

**Highly distributed repositories: Experimental platform, Projects, Institutes**

Interoperability

**High Data 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 Data Principles

- Findable
- Accessible
- Interoperable
- Reusable

**MIAPPE STANDARD**

**International standards**

- National Networks
- Global Networks
- European Networks

**International data standards**

- Web services
- Minimal information

**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
  - Elisabeth Arnaud, Paul Kersey, Pawel Krajewsky, Matthias Lange, Cyril Pommier, Bjorn Usadel

- Current Versions
  - Version 1: Transplant/Elixir + EPPN
  - Version 1.1: Elixir, validation in progress

**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 Plant Material**

- Biosource: Plant Material identification
- Multi Crop Passport Descriptor (MCPD) compliant
- [www.cropontology.org](http://www.cropontology.org)
- Key: Identification
  - 0. Persistent unique identifier: PUID = URI/DOI
  - 1. Institute code
  - 2. Accession number
  - 3. Genus
  - 4. Species
  - 5. Subspecies: ‘subsp.’ (for subspecies); ‘convar.’ (for convarietty); ‘var.’ (for variety); ‘f.’ (for form); ‘Group’ (for cultivar group)
  - 28. Remarks
- ID Field are reused in MIAPPE
- Suitable for non Genbank collections: breeder varietal list, laboratory collection, ...

**Minimum Information About Plant Phenotyping Experiment**

- Standard
  - Data exchange and traceability
  - Repositories
  - Analysis
- Input and output for analysis pipelines developed for phenotyping and genetic

**Towards recommendations for metadata and data handling in plant phenotyping**

- Thomas Atema, Daniel Gendre, Elisabeth Arnaud, Olga Chev, Guillaume Comolli, Fabio Forza, Wolfgang Dohmen, Kenedy Turino, Christian Klute, Matthias Lange, Gertje Meurer, Michele Nelles, Pascal Nemes, Jan van der Meer, Cyril Pommier, Adamde Prenter, Philippe Ricard-Blouin, Nikolaus Roningen Bronze, Jean Schilke, Marie van Schijndel, Olga Chev, Bjorn Usadel, Sophie Vacher, Frédéric Vervoort and Paola Koywood

**MIAPPE v1.1 Overview Variables**

- 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**

- 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
  - Elisabeth Arnaud, Paul Kersey, Pawel Krajewsky, Matthias Lange, Cyril Pommier, Bjorn Usadel

- Current Versions
  - Version 1: Transplant/Elixir + EPPN
  - Version 1.1: Elixir, validation in progress

- See [http://www.miappe.org](http://www.miappe.org) for providing feedback.
IMPLEMENTATIONS

Databases and repositories
File archive
Web services
Semantic web and RDF

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 BrAPI API**

- Resources
  - http://brapi.org/
- Collection of specifications for data retrieval and exchange
- Servers implementations
  - CGIAR international network
  - BISA (Bioinformatics Service Infrastructure)
- Elixir: ExCeLerate
- Emphasis: Phenotyping
- Germplasm API

**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
  - http://github.com/MIAPPE/MIAPPE-ontology
- BrAPI 2 MIAPPE RDF workflow
  - http://ist.blogs.inra.fr/wdi/phenotypes-as-rdf/
- 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/
  - edAI https://edal.ipk-gatersleben.de/
  - PlantPhenoDB at IPGPAS 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
  - Marcin Krajewski
  - Radoslaw Cwiek
  - Kasia Michalek
  - Danuta Kupczynska
  - Anna Michalik
  - Wieslaw Costa
- iBet
  - Marko Vukosav
  - Inês Costa
  - Celia M. Miguel
  - ISG
  - Daniel Feria
- Bioversity International CGIAR
  - Elizabeth Amsalu
  - Marie Angélique Laporte
- European Agroportal
  - Cyril Pommier / Environment
  - Pierre Blondon
  - Célia Kersey
  - Miguel Cornut
  - Cyril Pommier / Phenotyping
  - Pascal Neveu
  - Richard Eksteen