



HAL
open science

Data standards for plant phenotyping: MIAPPE and its implementations

Cyril Pommier, Guillaume Cornut, Thomas Letellier, Célia Michotey, Pascal Neveu, Manuel Ruiz, Pierre Larmande, Paul J. Kersey, Hanna Cwiek Kupczynska, Pawel Krajewski, et al.

► 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-02789754v1>

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

elixir, EMPHASIS, Bioversity International, Crop Ontology

INRA | Cyril Pommer / MIAPPE Plant Phenotyping Data standard | 14 Jan 2018

PLANT PHENOTYPING DATA STANDARD NEED

INRA | Cyril Pommer / MIAPPE Plant Phenotyping Data standard | 2

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

INRA | Cyril Pommer / MIAPPE Plant Phenotyping Data standard | 3

Plant Phenotyping Experiment Data Interoperability

- High Data Interoperability need
 - Intra dataset interoperability
 - Phenotyping networks consolidation
 - Project integration
 - Inter dataset interoperability
 - large scale breeding
 - genetic analysis
 - climat change studies
 - ...
- Interoperability and sharing principle

FAIR

INRA | Cyril Pommer / MIAPPE Plant Phenotyping Data standard

Plant Phenotyping Experiment Data FAIR Data Principles

F Findable	A Accessible	I Interoperable	R Reusable
Ids Metadata Indexed	Open Protocols Perennial Metadata	Linked Data Vocabularies	License Well described Provenance Standards

<https://www.ontoforce.com>

Wilkinson, M. D. et al. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* 3, 160018 (2016).

INRA | Cyril Pommer / MIAPPE Plant Phenotyping Data standard

MIAPPE STANDARD

INRA | Cyril Pommer / MIAPPE Plant Phenotyping Data standard | 6

International standards

National Networks

Global Networks

European Networks

International data standards

- BrAPI Web services
- RDA
- miappe Minimal information
- Crop Ontology
- Controlled vocabularies Trait dictionaries
- MCPD

Logos: ifit, BFF, FRANCE GRILLES, WheatIS, WHEAT INITIATIVE, CGIAR, Elixir Data, EMBL-EBI, EPPN, EMPHASIS

Footer: INRA Cyril Pommer / MIAPPE Plant Phenotyping Data standard 7

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

Hanna Cwik-Kupczyńska, Thomas Altmann, Daniel Arend, Elizabeth Arnaud, Dijun Chen, Guillaume Cornus, Fabio Fiorani, Wojciech Frohberg, Astrid Junker, Christian Klukas, Matthias Lange, Cezary Mazurek, Anahita Nafissi, Pascal Neveu, Jan van Deverren, Cyril Pommer, Hendrik Poorter, Philippe Rocca-Serra, Susanna-Assunta Sansone, Uwe Scholz, Marco van Schriek, Umit Seren, Björn Usadel, Stephan Weise, Paul Kersey and Pawel Krajewski III

Towards recommendations for metadata and data handling in plant phenotyping

Pawel Krajewski^{1,2}, Dijun Chen², Hanna Cwik¹, Aali D.J. van Diik³, Fabio Fiorani⁴, Paul Kersey⁵, Christian Klukas⁶, Matthias Lange⁷, Augustyn Markiewicz⁸, Jan Peter Nap⁹, Jan van Oeveren¹, Cyril Pommer¹, Uwe Scholz², Marco van Schriek¹⁰, Björn Usadel¹¹ and Stephan Weise¹²

This Article

1. Eup. Mol. C (2015), 66 (18): 5417-5427
doi: 10.1093/eup/671
First published online: June 4, 2015

Footer: INRA Cyril Pommer / MIAPPE Plant Phenotyping Data standard 7

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
 - Biodiversity International - CGIAR
 - Elizabeth Arnaud, Paul Kersey, Pawel Krajewski, Matthias Lange, Cyril Pommer, Björn Usadel
- 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.

Footer: INRA Cyril Pommer / MIAPPE Plant Phenotyping Data standard 9

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

Footer: INRA Cyril Pommer / MIAPPE Plant Phenotyping Data standard 10

MIAPPE v1.1 Overview Plant Material

- Biosource : Plant Material identification**
- Multi Crop Passport Descriptor (MCPD) compliant
 - <https://www.biodiversityinternational.org/e-library/publications/detail/faobiodiversity-multi-crop-passport-descriptors-v21-mcpd-v21/>
 - FAO & Biodiversity International
 - Genebanks
- Key : Identification
 - 0. Persistent unique identifier : PUID = URI/DOI
 - 1. Institute code
 - 2. Accession number
 - 5. Genus
 - 6. Species
 - 7. Subspecies : 'subsp.' (for subspecies); 'convar.' (for convariety); '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, ...

Footer: INRA Cyril Pommer / MIAPPE Plant Phenotyping Data standard 11

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

Logos: Crop Ontology (for crop ontology) www.cropontology.org

Screenshot of Crop Ontology interface:

Woody Plant Ontology

Ontology name: Woody Plant Ontology
 Authors: Celia Michalek
 Version: v1
 License: CC BY 4.0
 Labels: [GO](#) [PATO](#) [PDS](#) [PDS label](#)

Method class: Estimation

Method: Observe the plant description (see p11 fao21.pdf)

Method name: AxiT Estimation

Method reference: Physics of d. 20

Traits, methods and scales

Search terms:

NID: Neats presence

Identifier: WOODY:000007
 Name: Neats presence
 Synonyms: Neats presence (WOODY:000496)
 Control of use: This ontology is based for INRA
 Institution: INRA
 Submitted: Celia Michalek
 Date: 13/03/2017
 Crop: Woody Plant

Neats presence

Identifier: WOODY:1000006
 Name: Neats presence
 NID: Presence of neats
 Description: Presence of neats
 Main abbreviation: NID
 Status: Standard for INRA
 Class: Stress: Biotic stress

Visual ontology

Identifier: WOODY:000003
 Name: Visual ontology
 Description: Visual assessment
 Status: Standard for INRA
 Class: Estimation

Presence of neats

Identifier: WOODY:000003
 Name: Presence of neats
 Description: Presence of neats
 Status: Standard for INRA
 Class: Estimation



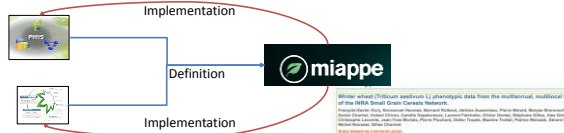
IMPLEMENTATIONS



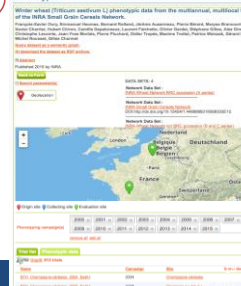
- Databases and repositories
- File archive
- Web services
- Semantic web and RDF



MIAPPE: Databases

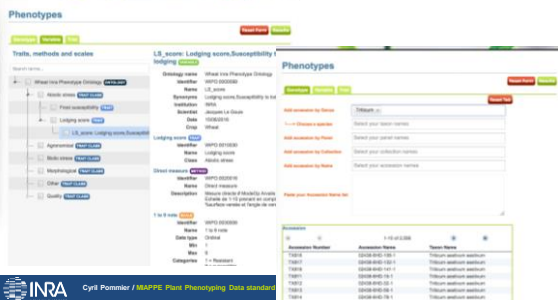


- Production databases
 - PHIS
- Cleaned data publication repositories
 - GnpIS <http://dx.doi.org/10.15454/1.448966215568333E12>
 - eDale
 - PlantPhenoDB
- Implementation
 - Internal
 - Web services API



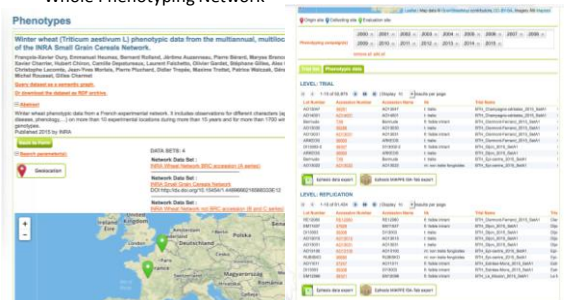
MIAPPE databases & FAIR

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



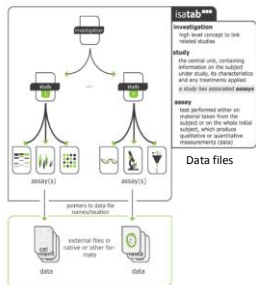
MIAPPE databases & FAIR

- Accessible
- Reusable: License, File standard
- Whole Phenotyping Network



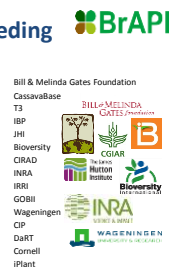
MIAPPE File Archive

- 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 MIAAPPE Web Service : Breeding BrAPI API

- Resources
 - <http://brapi.org/>
- Collection of specifications for data retrieval and exchange
- Servers implementations
 - CGIAR international network
 - Integrated Breeding Platform
 - Elixir Excelerate
 - Emphasis: PHIS
 - Germinate
 - GnpIS
- Clients implementations
 - Flapjack : genotyping data visualization
 - GnpIS
 - Ontology Widget
 - <https://github.com/gnpis/trait-ontology-widget>
 - R analysis pipelines

1:50 PM From Breeding Data to Decision by BrAPI and Open Rap

Session: Integrated Breeding Platform: Tools, Databases and Applications for Plant Breeding
 Location: Golden Ballroom
 Date: Monday, Jan 15 1:50 PM
 Presentation: 90 minutes

Presenting Author
 Yael Eshkol, FRA, A.
 Hebrew University & Research - Biometrics

miappe MIAAPPE Semantic: Plant Phenotyping Experiment Ontology

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

INRA CYRIL POMMIER / MIAAPPE Plant Phenotyping Data standard 20

miappe MIAAPPE Semantic: RDA RDFENO

- MIAAPPE OWL Ontology
 - First draft
 - <https://github.com/MIAAPPE/MIAAPPE-ontology>
- BrAPI 2 MIAAPPE RDF workflow
 - <http://www.wheatis.org> data standard page
 - <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

Wheat Data Interoperability Guidelines

Phenotypes as RDF

Phenotypes to Semantic web publication

Recommendations

Virtuoso SPARQL Query Editor

Default Data Set Name (Graph: R1)

Query Text

Results Formatted

Results

Securely retrieves the data stored at <http://dx.doi.org/10.15454/1.4489666216568333E12> from RDF files, and returns results in a table.

Results Formatted: HTML CSV JSON XML Turtle N-Triples RDF/XML TriG TriX TriZ TriX TriZ

Options: Show loading of new variables Log settings

Run Query:

INRA CYRIL POMMIER / MIAAPPE Plant Phenotyping Data standard

miappe Adoption

- Plant community involved
 - Elixir (European bioinformatic infrastructure)
 - Emphasis (European Phenotyping infrastructure)
 - Bioversity international CGIAR
- Breeding API is Elixir official Phenotyping standard web service
- MIAAPPE 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 IGPAS <http://cropnet.pl/plantphenodb/>
 - In progress: COPD, Elixir plant databases (iBet, WUR, VIB, ...), Brassica Information Portal, ...

INRA CYRIL POMMIER / MIAAPPE Plant Phenotyping Data standard 22

miappe Perspectives

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



INRA CYRIL POMMIER / MIAAPPE Plant Phenotyping Data standard 23

miappe Acknowledgment & Questions

transPLANT, eXcelerate, EMPHASIS, BrAPI

EMBL-EBI
 • Paul Kersey

IPG PAS
 • Hanna Cwiiek-Kupczynska
 • Pawel Krajewski

INRA, CIRAD, ZURGI, PHENOME

INSTITUTO GULBENKIAN DE CIENCIA
 • Cyril Pommier
 • Anne Françoise Adam
 • Blondin
 • Guillaume Cornut
 • Thomas Letellier
 • Célia Michottey
 • Pascal Neveu
 • Manuel Ruiz
 • Pierre Larmande

IBET
 • Bruno Costa
 • Inês Chaves
 • Célia M. Miguel

IGC
 • Daniel Faria

Bioversity International, Crop Ontology, Bioversity international CGIAR
 • Elizabeth Arnaud
 • Marie Angélique Laporte

VIB
 • Frederik Coppens

WAGENINGEN UNIVERSITY & RESEARCH
 • Richard Finkers

EaRIham Institute

MISTEA

NIB

INRA CYRIL POMMIER / MIAAPPE Plant Phenotyping Data standard 24