

Implementing an  
**ontology-driven** information system  
to push forward **data sharing**  
with a joint **data standard**  
for **Plant Phenotyping**  
in the Nordic countries

*Or, in other words:*

Our work with getting  
**PHIS/OpenSILEX**  
ready for our plant data

Tatu Hannu Polvinen, University of Helsinki  
Jesper Cairo Westergaard, University of Copenhagen



Tatu

Jesper

# Now versus PHIS: Our facilities' data now. Yours too?

Data in a lot of Excel sheets

On servers, desktops, external harddrives

And a lot of unstructured, ad-hoc naming/annotating of meta-data in the file names and elsewhere

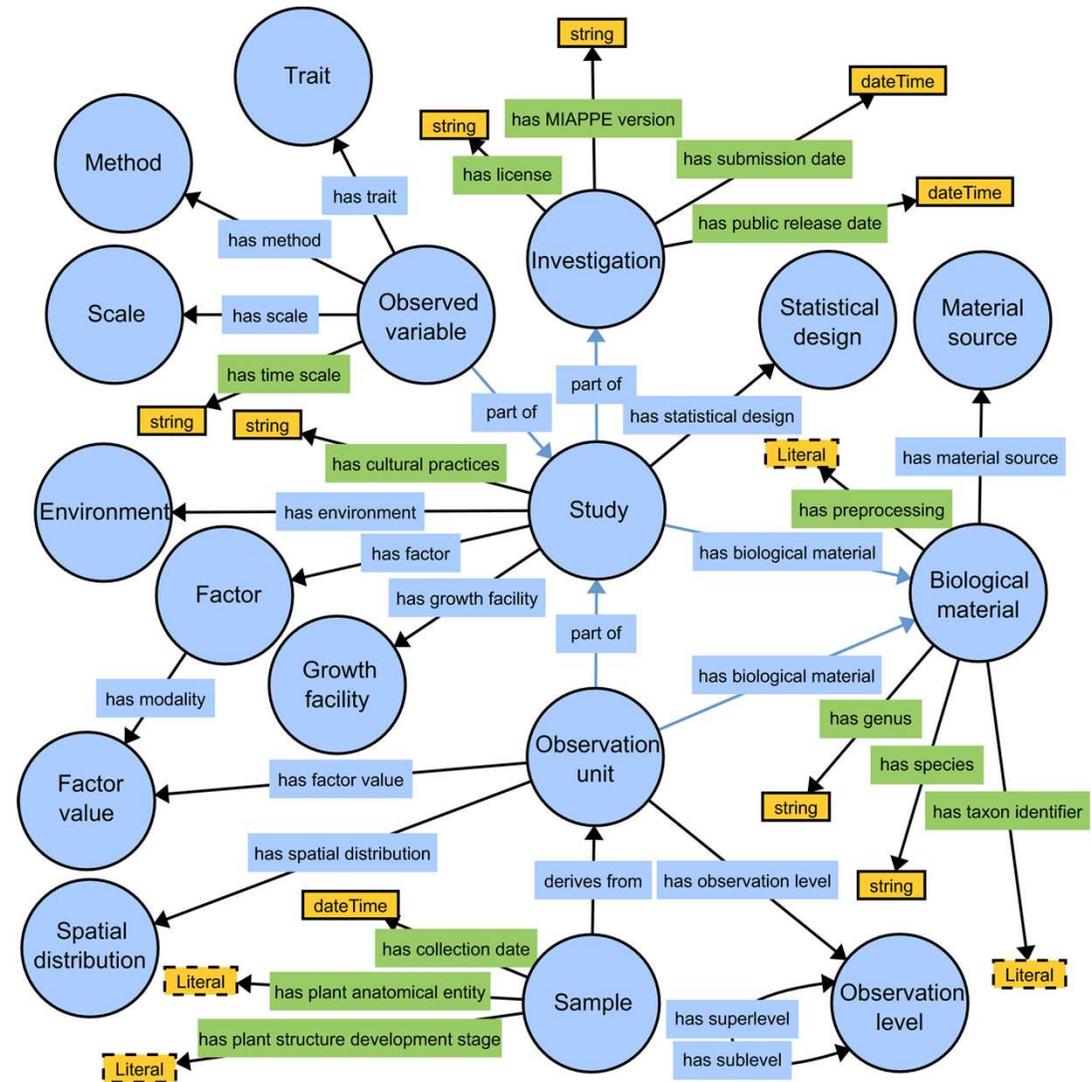
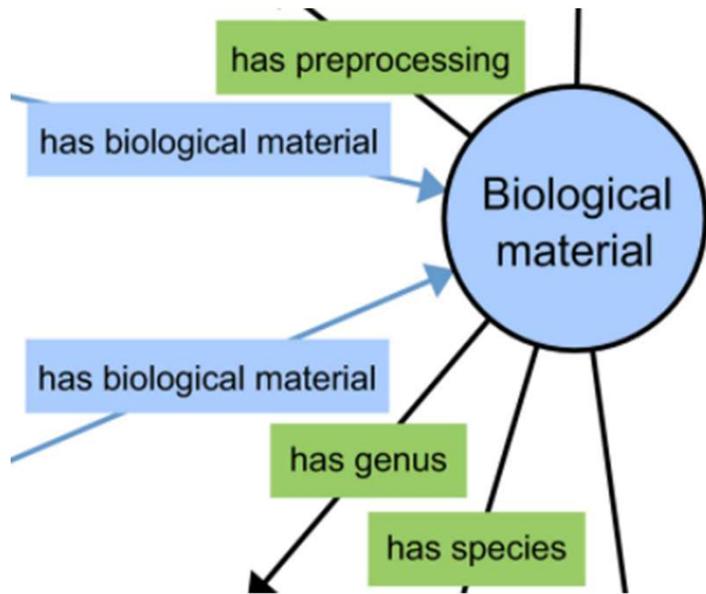
The image shows a complex Excel environment with multiple sheets open. The top sheet, 'Measuring Date', contains a table with columns: ID, Round, Order, Tray, Tray Info, Plant ID, Position, Plant Name, Plant Info, PID, Camera Positor, AREA\_PX, AREA\_MM, PERIMETER\_PX, PERIMETER\_MM, COMPACTNESS, ROUNDNESS, ROUNDNESS2, ISOTROPY, and ECCENTRICITY. The data rows list various plant IDs (e.g., Demo\_01, Demo\_02) and their corresponding measurements.

Below this, another sheet is visible with a table titled 'Request Data'. The columns are: Date and Time, Type, Fixture ID, and Measurement. The data includes entries for 'Bottom Weight (g)', 'Top Weight (g)', 'Thermo Inspection II', and 'Height (mm)' across various dates in 2020.

The bottom sheet shows a large, dense table with many columns, including 'Session Name', 'Session Recipe', 'Date', 'Time', 'Processing Algorithm', and various 'Mean' and 'StdDev' values for different parameters. The table is highly structured and contains a vast amount of numerical data.

# Getting all parameters described: Yes – it is hard

"Subset of the Plant Phenotyping Experiment Ontology representing the MIAPPE data model."



Papoutsoglou, Faria. "Enabling Reusability of Plant Phenomic Datasets with MIAPPE 1.1." The New Phytologist, vol. 227, no. 1, Wiley, Mar. 2020, pp. 260–73, doi:10.1111/nph.16544. !!! Open Access

# Fighting that dragon ... PHIS to the rescue

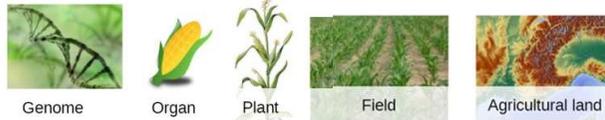
A screenshot of the PHIS website homepage. The browser address bar shows the URL [www.phis.inra.fr/openphis/web/index.php?r=site%2Flogin-as-guest](http://www.phis.inra.fr/openphis/web/index.php?r=site%2Flogin-as-guest). The page title is "Phenotyping Hybrid Information System". The navigation menu includes "Experimental Organization", "Data", "Tools", and "Phis Guest". The main content area features a green background with a large plant stem and leaves. In the center, the word "PHIS" is displayed in white, surrounded by icons representing data (stacks of colored disks) and a molecular structure. Below this, a LEGO knight minifigure is shown holding a sword and a shield. The text "Phis Guest, you are on PHIS" is displayed. At the bottom, there is a Creative Commons license logo and the text "Data are licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License". The footer contains the copyright information: "© INRA MISTEA-LEPSE 2014-2018 (PHIS v.3 - 03rd May 2018) ; © INRA MISTEA - SILEX".

# Now versus PHIS: Using PHIS

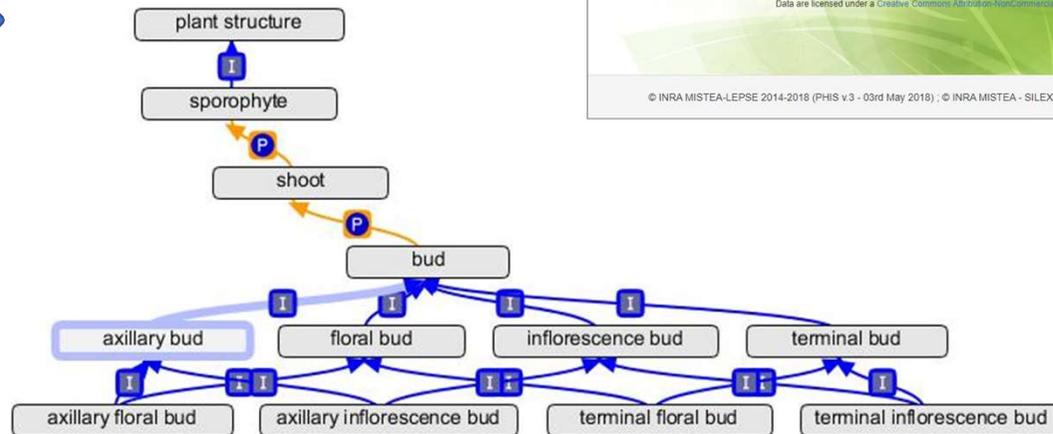
- Web user interface - only a browser is needed



- Multi-scale data



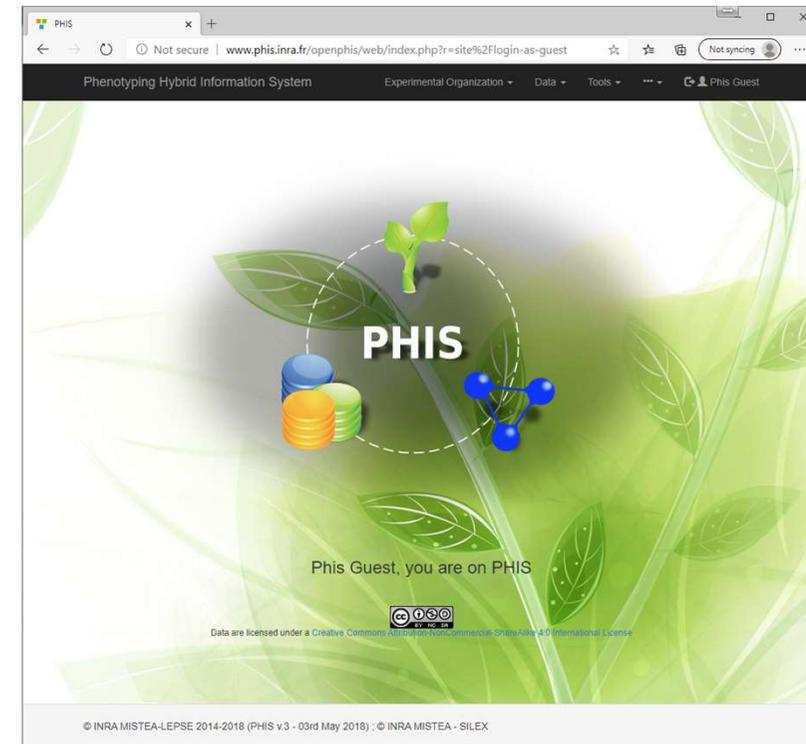
- Ontologies



- Object identification and tracking

- **Example:**

- Take a leaf,
- tie the leaf data to lighting,
- to the hardware, cameras, movement, placement
- to climate data,
- and all the measurements,
- and possible incidents.



<http://wiki.plantontology.org/index.php/Bud>

# Example: Leaf data

URI of leaf:

<mp3:arch/2014/lf/000000056>

URI of plant:

<mp3:arch/2014/pl/000000012>

URI of pot:

<mp3:arch/2010/ca/cabine2>

URI of cabin:

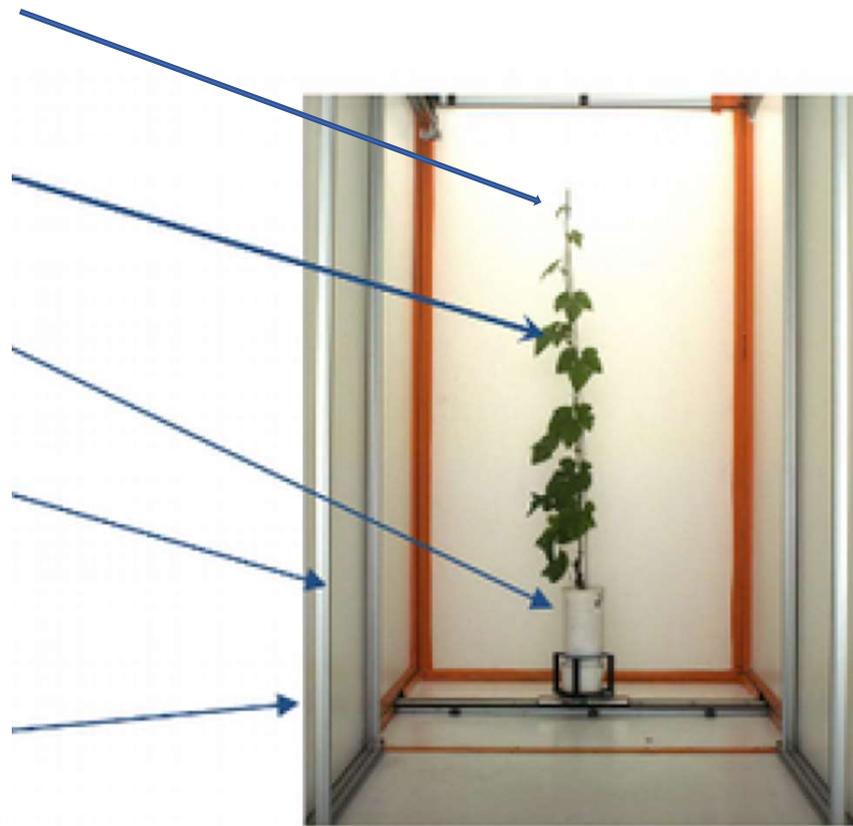
<mp3:arch/2011/ss/cabine2>

URI of camera:

<mp3:arch/2011/ss/00003312>

URI of image:

<mp3:arch/2015/im/000000564>



# Example: Leaf data

www.phis.inra.fr/openphis/web/index.php?r=measure-online%2Findex&dp-2-page=7

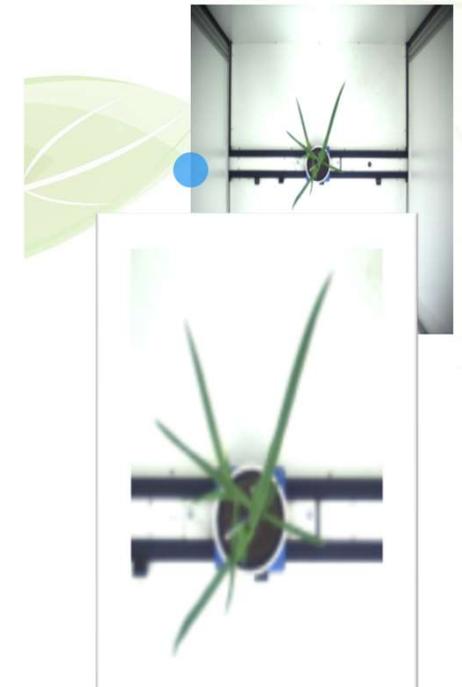
Phenotyping Hybrid Information System

Experimental Organization Data Tools Phis Guest

m3p/arch/2017/c17001685	1693/CRAZI/ZM4112/WWW/PSI_100/Loop5/ARCH2017-03-30	m3p./ARCH2017-03-30	m3p./variable/v000028	side120	2017-04-13 10:23:12.55884	1206.96268463	
m3p/arch/2017/c17001685	1693/CRAZI/ZM4112/WWW/PSI_100/Loop5/ARCH2017-03-30	m3p./ARCH2017-03-30	m3p./variable/v000028	side90	2017-04-13 10:23:12.058812	1270.45767581	
m3p/arch/2017/c17001685	1693/CRAZI/ZM4112/WWW/PSI_100/Loop5/ARCH2017-03-30	m3p./ARCH2017-03-30	m3p./variable/v000028	side180	2017-04-13 10:23:13.557897	1175.0858041	
m3p/arch/2017/c17001685	1693/CRAZI/ZM4112/WWW/PSI_100/Loop5/ARCH2017-03-30	m3p./ARCH2017-03-30	m3p./variable/v000028	side150	2017-04-13 10:23:13.057869	1204.42786431	
m3p/arch/2017/c17001685	1693/CRAZI/ZM4112/WWW/PSI_100/Loop5/ARCH2017-03-30	m3p./ARCH2017-03-30	m3p./variable/v000028	side240	2017-04-13 10:23:14.558955	1295.06046653	
m3p/arch/2017/c17001685	1693/CRAZI/ZM4112/WWW/PSI_100/Loop5/ARCH2017-03-30	m3p./ARCH2017-03-30	m3p./variable/v000028	side210	2017-04-13 10:23:14.057926	1256.43103921	
m3p/arch/2017/c17001685	1693/CRAZI/ZM4112/WWW/PSI_100/Loop5/ARCH2017-03-30	m3p./ARCH2017-03-30	m3p./variable/v000028	side300	2017-04-13 10:23:58.656162	1230.51050794	
m3p/arch/2017/c17001685	1693/CRAZI/ZM4112/WWW/PSI_100/Loop5/ARCH2017-03-30	m3p./ARCH2017-03-30	m3p./variable/v000028	side270	2017-04-13 10:23:58.164134	1316.80428851	
m3p/arch/2017/c17001685	1693/CRAZI/ZM4112/WWW/PSI_100/Loop5/ARCH2017-03-30	m3p./ARCH2017-03-30	m3p./variable/v000028	top0	2017-04-14 16:21:11.623347	2171.88572168	
m3p/arch/2017/c17001685	1693/CRAZI/ZM4112/WWW/PSI_100/Loop5/ARCH2017-03-30	m3p./ARCH2017-03-30	m3p./variable/v000028	side330	2017-04-13 10:23:59.156191	1199.84422994	
m3p/arch/2017/c17001685	1693/CRAZI/ZM4112/WWW/PSI_100/Loop5/ARCH2017-03-30	m3p./ARCH2017-03-30	m3p./variable/v000028	side30	2017-04-14 16:21:11.866361	1849.42819607	
m3p/arch/2017/c17001685	1693/CRAZI/ZM4112/WWW/PSI_100/Loop5/ARCH2017-03-30	m3p./ARCH2017-03-30	m3p./variable/v000028	side0	2017-04-14 16:21:11.623347	1596.58061838	

5919ca6380ccaa6202f8dc38

Object Uri	m3p./arch/2017/c17001685
Object Alias	1693/CRAZI/ZM4112/WWW/PSI_100/Loop5/ARCH2017-03-30
Experiment ID	m3p./ARCH2017-03-30
Variable Uri	m3p./variable/v000028
Date	2017-04-14 16:21:11.623347
Value	2171.88572168
Unit Name	pixels
View	top0
Image Uri	m3p./arch/2017/c17001605361
Binary Image	http://web.supagro.inra.fr/phis/data/phenoarch/binaries/AF



Source: The publicly accessible PHIS try-out installation at:

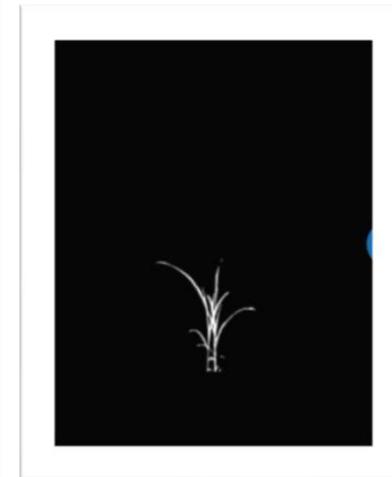
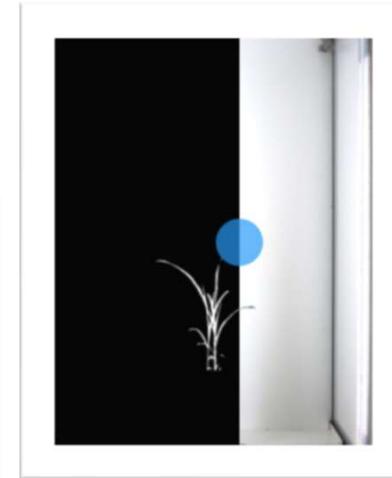
<http://www.phis.inra.fr/openphis/web/index.php?r=event%2Findex&page=6&per-page=5>

# Example: Leaf data: Images

...php?r=measure-online%2Findex&dp-2-page=7

ation System Experimental Organization Data Tools Phis Guest

93/CRAZI/ZM4112/WW/PSI_100/Loop5/ARCH2017-03-30	m3p:/ARCH2017-03-30	m3p:/variable/v000028	side120	2017-04-13 10:23:12.55884	1206.96268463	
93/CRAZI/ZM4112/WW/PSI_100/Loop5/ARCH2017-03-30	m3p:/ARCH2017-03-30	m3p:/variable/v000028	side90	2017-04-13 10:23:12.058812	1270.45767581	
93/CRAZI/ZM4112/WW/PSI_100/Loop5/ARCH2017-03-30	m3p:/ARCH2017-03-30	m3p:/variable/v000028	side180	2017-04-13 10:23:13.557897	1175.0858041	
93/CRAZI/ZM4112/WW/PSI_100/Loop5/ARCH2017-03-30	m3p:/ARCH2017-03-30	m3p:/variable/v000028	side150	2017-04-13 10:23:13.057869	1204.42786431	
93/CRAZI/ZM4112/WW/PSI_100/Loop5/ARCH2017-03-30	m3p:/ARCH2017-03-30	m3p:/variable/v000028	side240	2017-04-13 10:23:14.558955	1295.06046653	
93/CRAZI/ZM4112/WW/PSI_100/Loop5/ARCH2017-03-30	m3p:/ARCH2017-03-30	m3p:/variable/v000028	side210	2017-04-13 10:23:14.057926	1256.43103921	
93/CRAZI/ZM4112/WW/PSI_100/Loop5/ARCH2017-03-30	m3p:/ARCH2017-03-30	m3p:/variable/v000028	side300	2017-04-13 10:23:58.656162	1230.51050794	
93/CRAZI/ZM4112/WW/PSI_100/Loop5/ARCH2017-03-30	m3p:/ARCH2017-03-30	m3p:/variable/v000028	side270	2017-04-13 10:23:58.164134	1316.80428851	
93/CRAZI/ZM4112/WW/PSI_100/Loop5/ARCH2017-03-30	m3p:/ARCH2017-03-30	m3p:/variable/v000028	top0	2017-04-14 16:21:11.623347	2171.88572168	
93/CRAZI/ZM4112/WW/PSI_100/Loop5/ARCH2017-03-30	m3p:/ARCH2017-03-30	m3p:/variable/v000028	side330	2017-04-13 10:23:59.156191	1199.84422994	
93/CRAZI/ZM4112/WW/PSI_100/Loop5/ARCH2017-03-30	m3p:/ARCH2017-03-30	m3p:/variable/v000028	side30	2017-04-14 16:21:11.866361	1849.42819607	
93/CRAZI/ZM4112/WW/PSI_100/Loop5/ARCH2017-03-30	m3p:/ARCH2017-03-30	m3p:/variable/v000028	side0	2017-04-14 16:21:11.623347	1596.58061838	



5919ca6380ccaa6202f8dc40

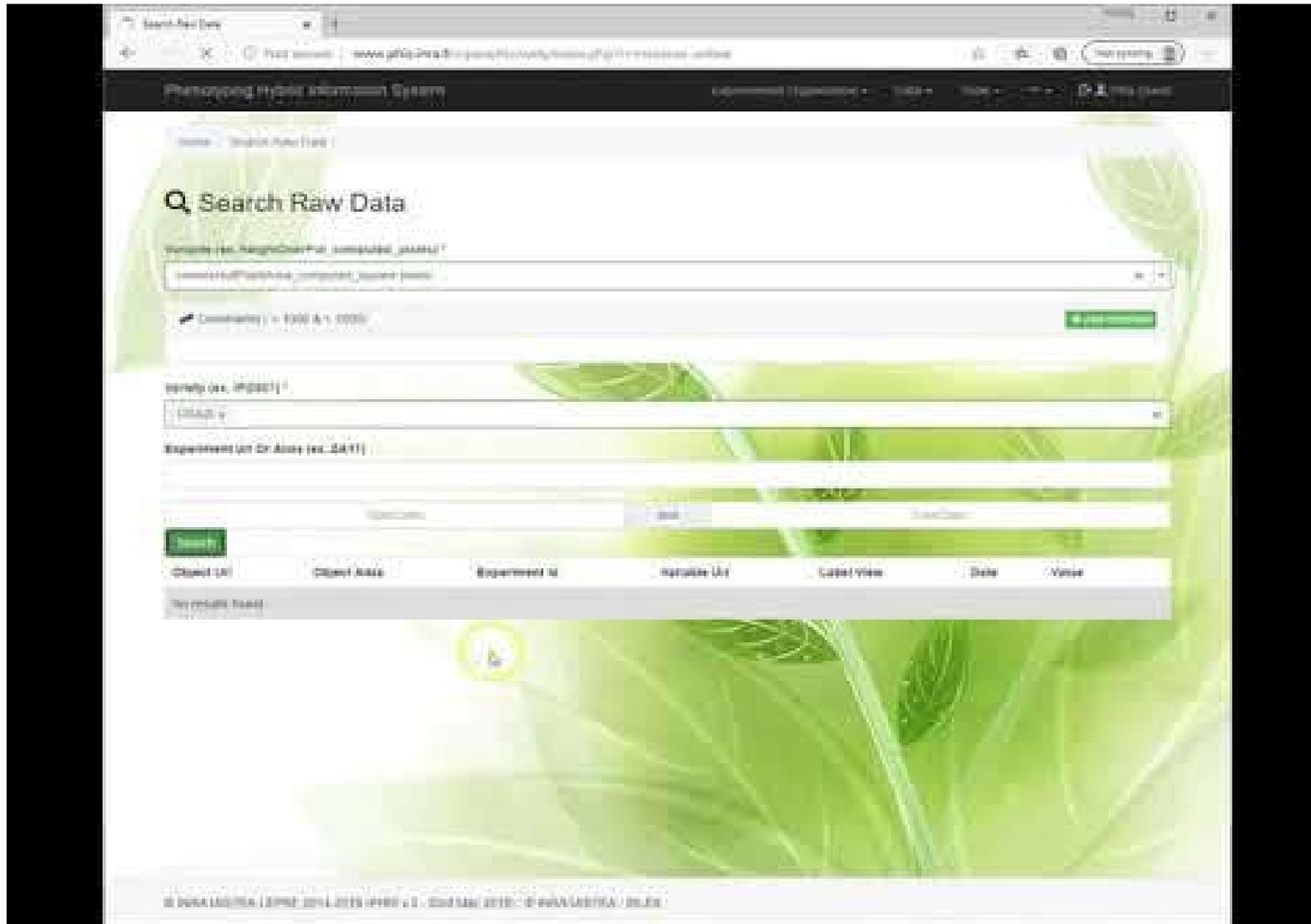
Object Uri	m3p:/arch/2017/c17001685
Object Alias	1693/CRAZI/ZM4112/WW/PSI_100/Loop5/ARCH2017-03-30
Experiment ID	m3p:/ARCH2017-03-30
Variable Uri	m3p:/variable/v000028
Date	2017-04-14 16:21:13.366446
Value	1869.70724881
Unit Name	pixels
View	side210
Image Uri	m3p:/arch/2017/c17001605369
Binary Image	http://web.supagro.inra.fr/phis/data/phenoarch/binaries/AR



Source: The publicly accessible PHIS try-out installation at

<http://www.phis.inra.fr/openphis/web/index.php?r=event%2Findex&page=6&per-page=5>

# Example: Leaf data: Images



Video at:

<https://www.youtube.com/watch?v=klhlunwo47c&feature=youtu.be>

Source: The publicly accessible PHIS try-out installation at

<http://www.phis.inra.fr/openphis/web/index.php?r=event%2Findex&page=6&per-page=5>

# Example: Events during a trial

Phenotyping Hybrid Information System

Experimental Organization ▾ Data ▾ Tools ▾ ... ▾ Phis Guest

Home / Events

## Events

Date Range

Start Date and End Date

Refresh

Showing 26-30 of 31,151 items.

#	Type	Description	Concern	Alias	Date of Event
26	Event	Harvesting of plot	diaphen:/2017/o17000118	118/DZ_PG_74/ZM4397/WWW/2/DIA2017-05-19	2017-09-22T12:00:00+01:00
27	Event	Harvesting of plot	diaphen:/2017/o17000119	119/DZ_PG_62/ZM4399/WWW/2/DIA2017-05-19	2017-09-22T12:00:00+01:00
28	Event	Harvesting of plot	diaphen:/2017/o17000120	12/DZ_PG_10/ZM4365/WWW/1/DIA2017-05-19	2017-09-22T12:00:00+01:00
29	Event	Harvesting of plot	diaphen:/2017/o17000120	120/DZ_PG_62/ZM4393/WWW/2/DIA2017-05-19	2017-09-22T12:00:00+01:00
30	Event	Harvesting of plot	diaphen:/2017/o17000121	121/DZ_PG_19/ZM4367/WWW/2/DIA2017-05-19	2017-09-22T12:00:00+01:00

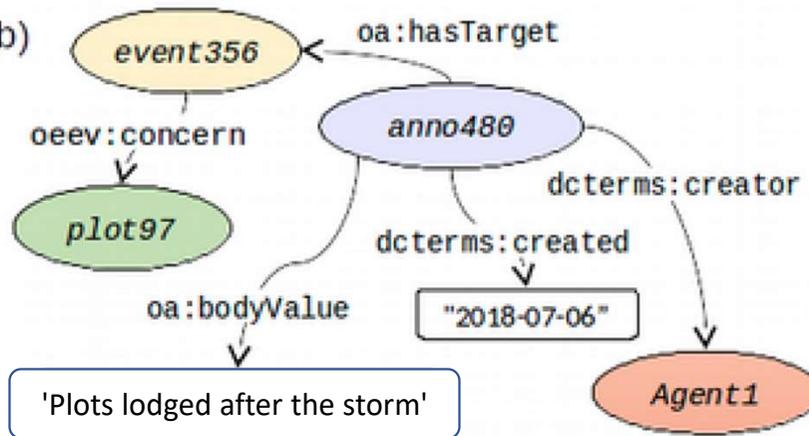
- Event
- Displacement
- Move from
- Move to
- Facility management
- Calibration
  - Automatic calibration
  - Manual calibration
- Installation of devices
- Operability
  - Corrective maintenance
  - Repair
  - Evolutionary maintenance
  - Servicing
  - Checking
  - Cleaning
- Restart
- Start
- Stop
- Scientific object management
  - Adding product
  - Fertilization
  - Clipping
  - Detasseling
  - Harvesting
  - Irrigation
  - Loading
  - Observed phenology
  - Potting
  - Sampling
  - Sowing
  - Staking
  - Thinning
  - Treatment
    - Curative treatment
    - Preventive treatment
- Trouble
- Breakdown
- Dysfunction
- Incident
  - Lodging
  - Pest attack
  - Pot fall
  - Stuck plant

# Example: Events during a trial ... Lodging

(a)



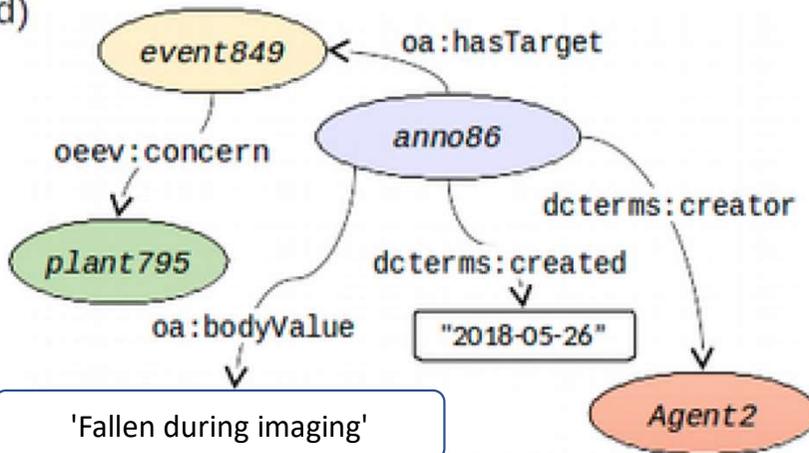
(b)



(c)



(d)



# Some definition of terminology is needed

- **PHIS = P**henotyping **H**ybrid **I**nformation **S**ystem

- **Hybrid:**

- Not as in a plant hybrid, but as in **combining different databases**

- **Information System:**

- "Information systems are interrelated components working together to **collect, process, store, and disseminate** information to **support decision making, coordination, control, analysis, and visualization** in an organization." [\*]

- **Ontology-driven** information system

- **Ontology:** "Onto" = Existence, being real + "Logia" = Study (science)

- "In the world of computer science and bioinformatics, an ontology is **a representation of a body of knowledge that exists about objects in a domain of interest, the categories** of those objects, and **their inter-relationships**. The categories of objects, or "classes," are a conceptualization of the knowledge that exists about the objects" [\*\*]
- "The term ontology originates from philosophy, where it refers to a unique description of the universe or 'things that are'. In modern information technology, especially in the Semantic Web, **an ontology is a model of** (some aspects of) **the world, which introduces key vocabulary** (such as **concepts and relations**) of a target domain and their meanings." [\*\*\*]

- **Ontology-driven:** It is all centered around the ontology. The ontology decides the structure. Makes the data interoperable (you can understand and use my data).

... "combining different databases"

... "introduces key vocabulary that we share"

... "support decision making, coordination, control, analysis, and visualization"

... "a representation of a body of knowledge"

... "their inter-relationships" so we can easily use each other's data

[\*] Excerpted from *Management Information Systems*, twelfth edition, Prentice-Hall, 2012.

[\*\*] Cooper, Jaiswal. "The Plant Ontology: A Tool for Plant Genomics." *Plant Bioinformatics*, vol. 1374, Springer New York, 2016, pp. 89–114, doi:10.1007/978-1-4939-3167-5\_5.

[\*\*\*] Pan, Aßmann. "Ontology-Driven Software Development." *Ontology-Driven Software Development*, 1. Aufl., Springer-Verlag, 2013, doi:10.1007/978-3-642-31226-7.

# Is it worth the investment of your time?

- Funding agencies may demand that the data you produce is FAIR
- Making your data FAIR may take considerable effort on your part
- PHIS may help to:
  - Connect your data with other datasets and research
  - Help to manage your data
  - Assign ontologies to your data
  - Store and share your data
- It will be a REALLY big investment of your time, the first time around.
  - You will have to get to know the new system, to a certain extent
  - You will have to learn how to format your data to the system
  - You will have upload/integrate your data to the PHIS server
  - You will have to quality control your data now living in the system
- There are altruistic reasons for joining. More well-annotated data can advance research for us all.
- But, what's in it for you? ...



**Other than INRAe in France, who are using PHIS?**

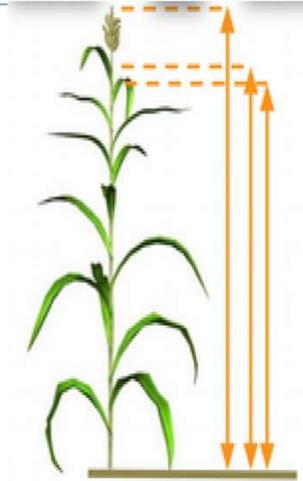
- Wageningen, Netherlands
- Canberra University, Australia
- Tokyo University, Japan
- Louvain-la-Neuve, Belgium

# For example: Perhaps avoid these common mistakes systematically?



## Some common mistakes we do

- **Metadata in file names** (not standardized, very often not machine readable, reduces metadata quantity and quality)
  - 2017-Paris-Syrah-irrig-goblet.csv
  - 2017-St-Paul-Merlot-guyot-not\_irrig.csv
- **Variable naming**  
same name for several variables, not well defined, no ID, no schema,...
- **Data are stored on personal computer**
- **Unstable files** (machine incompatible organisations)
- **Ambiguous ID**
- **Context, faults are not described**
- **No data links**
- **Missing data representation**
- **No license,**
- ... etc



Plot566  
in 2016

Plot566  
in 2017



# Our cooperation with INRAe

- PHIS was created by INRA (now INRAe) and the Montpellier SupAgro research institution
- INRAe has made PHIS freely available as open-source software
- We are enlisting INRAe's development team for support
  - Limited hours of paid support
  - For effective installation, configuration and testing of our PHIS
- Allows for us to gain experience and configure PHIS for our use case
- Experience and familiarisation will allow us to support others

# How we are going to implement and test a Nordic PHIS installation

- Install the PHIS platform with consulting from INRAe
- Test with small dataset
  - A sub-set of our potato phenotyping dataset
- Test with full dataset
  - Our full potato phenotyping dataset
- Evaluate/status
- Add other datasets
  - Data interoperability and cooperation
  - Would you be interested in being part of this test phase?

## The test dataset, Potatoes:

- Same genotypes have been analyzed in all three institutions:
  - SLU – Biotron
  - UHEL - NaPPI
  - UCPH - PhenoLab
- Allows for
  - shared analysis
  - unification of ontologies
  - setting up storage structures.

# When can you use the Nordic PHIS as a normal resource/infrastructure?

- We will need to see the pilot project be a success
- Then we will need some funding, and some commitment from participating universities, to make it a reality
- We will need to develop policies and protocols:
  - For welcoming new users
  - For supporting users (both technical/IT and ontology issues)
  - For those of us working with it > ensure long term system accessibility/viability
  - For "spreading the word" = Growing the user base
  - For upgrading and expanding the system (incl. hardware). E.g. as new users and thus expectedly much more data will be added

## Links etc.

- [www.inrae.fr/en](http://www.inrae.fr/en)
- [www.inrae.fr/en/centres/occitanie-montpellier](http://www.inrae.fr/en/centres/occitanie-montpellier)
- [phis.inra.fr](http://phis.inra.fr)
  
- PHIS is part of the OpenSILEX collaborative meta-project developed at MISTEA joint research unit (INRA - SupAgro).
  
- The current version of OpenSILEX PHIS is presented on the webpage [opensilex.org](http://opensilex.org) / [opensilex.github.io](http://opensilex.github.io)

Thank you.

Now, questions?