

	<p><b>LOUISIANA STATE UNIVERSITY</b> <b>College of Agriculture</b> <b>School of Plant, Environmental, and Soil Sciences</b> <b>AGRO 7076 HTP in Plant Breeding</b></p>	
---	--	---

# Introduction to the HTP course

**Prof. Roberto Fritsche-Neto**

**[rfneto@agcenter.lsu.edu](mailto:rfneto@agcenter.lsu.edu)**

**Baton Rouge, Jan 24<sup>th</sup>, 2024**

# Schedule

Workflow (3 credits)		
Class	Date	Lectures (Wednesdays, from campus) 9:30 -12:30 am
1	Jan 24	Introduction to HTP; Radiation properties and remote sensing
2	Jan 31	Image properties and acquisition; Image matrix practice on R
3	Feb 7	Indices and tools (HW 1)
4	Feb 21	Metashape: orthomosaic, DEM, and cloud points; Shapefiles on R (HW 2)
5	Feb 28	Raster, index, counts, areas, and plant height on R (HW3)
6	Mar 6	Envirotyping (theory) and Practice on R (HW4)
7	Mar 20	Pre-processing images (theory and practice on R and VGG image annotator) (HW 5)
8	April 3	Data mining and modeling (Theory)
9	April 10	Data mining and modeling – Practice on R (HW 6)
10	April 24	Journal Club

**Classes**

**Scripts and data**

**Evaluation**

**Homeworks**

**Journal Club**

[Moodle](#)

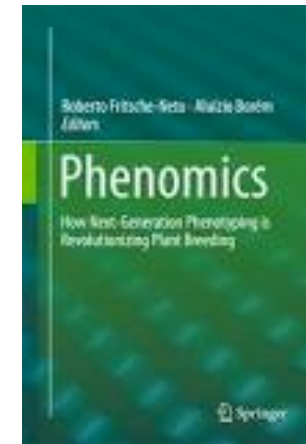
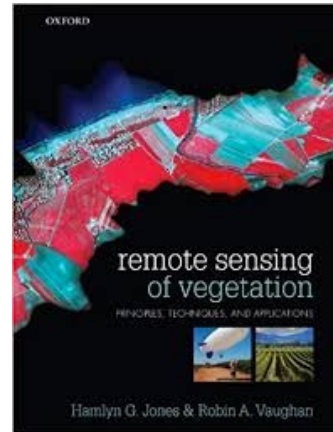
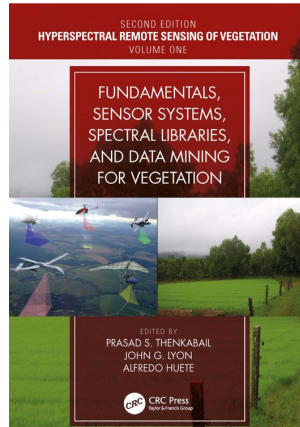
<https://github.com/rfn-qt/>

[Moodle](#)

[4 papers in the end of the course](#)

# Main bibliography and material

- **Books**



- **Softwares**

- QGIS - <https://www.qgis.org/en/site/>
- Metashape - <https://www.agisoft.com/>
- R - <https://cran.r-project.org/>
- Rstudio - <https://posit.co/download/rstudio-desktop/>

- See and install the packages one week in advance based on the script in the next slide

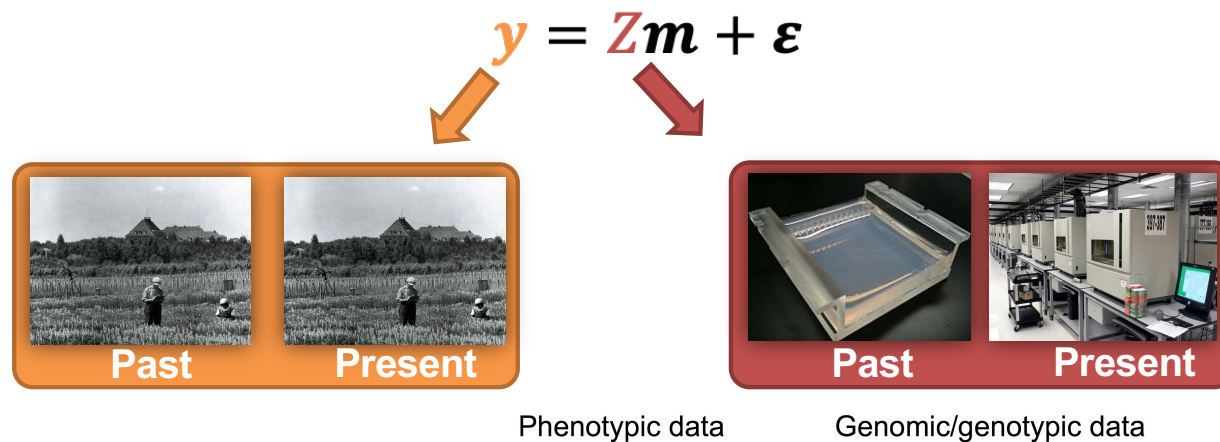
# R packages and material

- Create a folder on your PC and inside it download the materials
- Download the datasets a week before (26 Gb of data!!!):
- Also, install R (version 4.2.3), then, these packages running this code:
  - `install.packages("imager", "raster", "devtools", "magick", "foreach", "doParallel", "sp", "tidyr", "data.table", "rlas", "lidR", "spdep", "pak", "BiocManager", "devtools", "superheat", "caret", "caretEnsemble", "h2o", "Cairo", "reticulate", "keras", "tensorflow", "abind", "ggplot2", "terra", "sf")`
  - `BiocManager::install("EBImage")`
  - `pak::pkg_install("TiagoOlivoto/pliman")`
  - [https://cran.r-project.org/src/contrib/Archive/rgdal/rgdal\\_1.6-7.tar.gz](https://cran.r-project.org/src/contrib/Archive/rgdal/rgdal_1.6-7.tar.gz) (install manually this package)
  - `devtools::install_github("andersst91/UAStools")`
  - `devtools::install_github("rfn-qt/SoilType")`
  - `devtools::install_github('allogamous/EnvRtype', force=TRUE)`
  - Install Tensorflow via <https://tensorflow.rstudio.com/install/>
  - For Mac users follow: <https://medium.com/gft-engineering/macbook-m1-tensorflow-on-jupyter-notebooks-6171e1f48060>

# Plant breeding as a solution

- It is the genetic improvement of plants for human benefit
- Two sources of information - **Genotype and Phenotype**
- Connect genotype to phenotype - **heritability**
- Understand the underlying genetics of traits
- Improve the rate of genetic gain
- Develop cultivars more efficiently
- It necessitates very large sample (**population**) sizes
- Improve the probability of selecting favorable recombinants
- **Challenge:** collecting phenotypic data on large populations

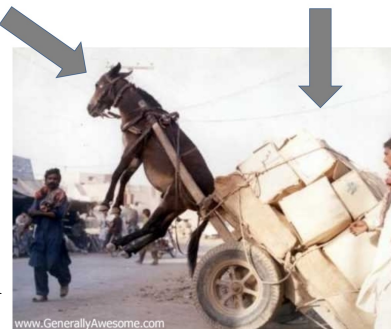
# Phenotyping vs. genotyping



**Phenomix (HTP)**

**Current challenges:**

1. Validation of methods
2. Optimization of information



# Plant Phenotyping

- **Classic**
- Time intensive
- Limited amount of data collected
- Small population sizes - **single time point data - Quality?**
- Trait bias - **visually scored**
- Expensive?
  
- **High-Throughput (HTP)**
- Efficient
- Massive amount of data collected - **Large population sizes**
- Multiple time point data
- Latent traits – **machine scored**
- Lower cost?

# Bottleneck



- How do we increase VG, accuracy, and response to selection?
- How can we accurately evaluate thousands of genotypes?

- **Effect on Breeder's equation**

$$RS = \frac{i \cdot r \cdot Vg}{T}$$

- Reduce phenotyping costs
- **Challenges:** build the structure, algorithms, and pipelines



# Applications



**Corn  
Stand  
RGB**

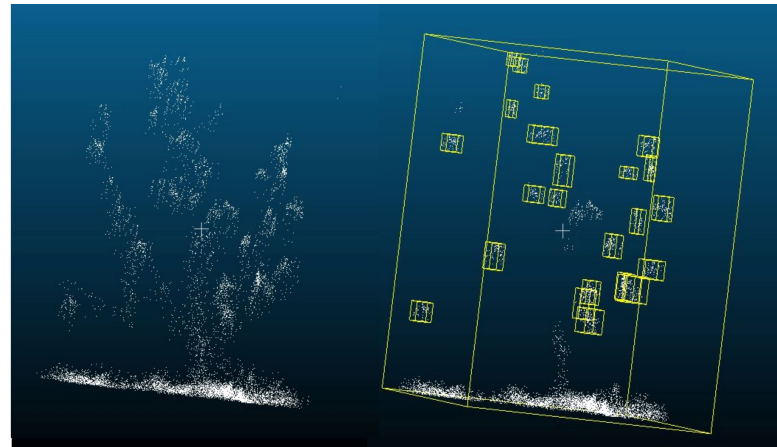
<u>Machine</u>	<u>Manual</u>
2.25 seconds per plot	5.75 seconds per plot
Requires 1 person	Requires 9 people

<u>Machine Average</u>	<u>Manual Average</u>
2%	5.6%
Range 0% - 5% on plot basis	Range 0% - 26% on plot basis

Source: Gore (2017)

## Light Detection And Ranging (LIDAR) phenotyping of defoliated cotton plants



# HTP is an interdisciplinary field

- Computing languages
- Artificial intelligence
- Physiology
- Crop development
- Quantitative genetics
- Remote sensing
- Image and spectral wavebands
- GPS or GNSS
- <http://thetruesize.com/>

