Introduction to the Sweetpotato (SP) Database Infrastructure
Purpose & Goals

• Efficient storage of all SP breeding data
• Trial data, phenotypic and genotypic information, metadata
• Incorporate analysis tools
• Easy querying: Slice data by year, location, program, etc.
• Enable new, genome-based breeding methods
Rony Swennen’s banana breeding notebooks on their way to Belgium
1980s: The Personal Computer Revolution
Excel-based data management

1. Experiment planning & Data Capture
2. Data Analysis
3. Storage
Data Management with Excel

**Plusses**
- Easy to use
- User has complete control

**Minuses**
- Sharing Data
- Combining Data
  - Data Analysis difficult across sheets
- Data Integrity
- Becomes difficult to manage for large datasets
- Difficult to manage genotyping information
Internet
Data management using the “cloud”

• Most data is managed using the web
  – Youtube for videos
  – Google for documents
  – Flickr, iCloud etc. for photos
  – Twitter for status updates
  – etc.

• What about breeding data?
Data Infrastructure

Main Datastore
• https://sweetpotatobase.org/

Analysis (integrated with sweetpotatobase)
• http://hidap.sgn.cornell.edu/

Data Collection
• Android FieldBook App
• AccuDataLogger
• The Breeding API
  • Application Programming Interface
• Standardized way to exchange breeding data between applications over the internet
• http://brapi.org/
Database Requirements

Standardization of

- Trait dictionary and measurement procedures
  - Ontologies (Reinhard)
- Naming of plant accessions
Genotyping data

- Can be extremely voluminous data
  - (GBS, etc)
- Must be integrated with phenotypic data
- For Cassavabase, already 1.5 billion data points
- Hard to manage without a database
Data Security on the Web

• Data protected by logins
• Different levels of user access privileges
• Regular on-site and off-site backups
Open Data Policy

- Data on Sweetpotato-base are open
- Need a user account to download
- Downloads are tracked by account
Breeding Tools
Plan and create sweetpotato breeding trials.

Upload Accessions
Make a Cross
Create a Trial
Manage Trials

You have already accepted the website usage policy. Thank you.

Site News
Phenotype Demo
A presentation covering how phenotyping and genotyping sweetpotatoes can be done.

Featured Publication
“Cloud” approach

Plusses

- Easy to use through a web browser
- No software installation necessary
  - Software is continuously updated
- All data automatically integrated
  - Query over several years, locations, etc. possible
- Integrates phenotypic with genotypic data

Minuses

- Internet needs to work
Asante!