IRRI-GRIMS: Tool for managing the largest rice collection

Ella Herrera

TT Chang Genetic Resources Center, IRRI

2007 ICIS Developers' Workshop , SPARC, Canada

Rice Science for a Better World

INTERNATIONAL RICE RESEARCH INSTITUTE

- Overview of IRRI-GRIMS
- Features
- Mapping of Genebank concepts/ objects to ICIS
- Advantages of Adopting ICIS Technology
- Current Activities

Overview

- IRRI Genetic Resources Information Management System
  - A data management tool for IRRI genebank that uses ICIS tools and application
  - An overhaul of legacy system (International Rice Genebank Information System/ IRGCIS)
  - Uses GMS, DMS, IMS and GRIMS-specific schema
  - Deployed at IRRI Genebank in January 2007

  - Developing GRIMS required migration of both program codes and data to new programming language and database schema

Rice Science for a Better World
INTERNATIONAL RICE RESEARCH INSTITUTE

Features

- *Incoming* germplasm receipt and registration
- Seed selection for regeneration and characterization
- Seed distribution and seed request tracking (within genebank)
- Seed viability monitoring
- Automated inventory management using barcode technology and digital weighing scale
- Integration with other organizational unit's data within IRRI (i.e. Seed Health, Plant Breeding, and INGER)

Rice Science for a Better World
INTERNATIONAL RICE RESEARCH INSTITUTE

Mapping of Genebank concepts to ICIS

- **Incoming samples**
- Germplasm acquisition method -> Imported

- Imported
  - Germplasm date (GDATE) is the acquisition date
  - Germplasm location is the holding institute -> TT Chang Genetic Resources Center (IRRI)

- Creation of germplasm group and source for collected samples
  - Create a germplasm and attach all the collecting information to it
  - Set this germplasm as the group and source of the sample received by the genebank

Rice Science for a Better World
INTERNATIONAL RICE RESEARCH INSTITUTE

Mapping of Genebank Concepts to ICIS

- Accession
  - Identified through Management Group ID (MGID)
  - Start of management neighborhood where GID=MGID
  - Holding institute as the germplasm location

- Passport Data
  - Descriptive passport data are stored as germplasm ATTRIBUTES
  - Phenotypic passport data are stored in DMS
    - One study corresponds to one planting season
    - A single accession can have multiple characterization data due to multiple field trials
  - * *Publicly available*

Rice Science for a Better World
INTERNATIONAL RICE RESEARCH INSTITUTE

Mapping of Genebank Concepts to ICIS

- Names
  - Management name
    - Seed increase/ regeneration name
  - Donor accession number
    - IRRI received an accession from other genebank
  - Foreign accession number
    - IRRI donated accession to other genebank

Rice Science for a Better World
INTERNATIONAL RICE RESEARCH INSTITUTE

Mapping of Genebank Concepts to ICIS

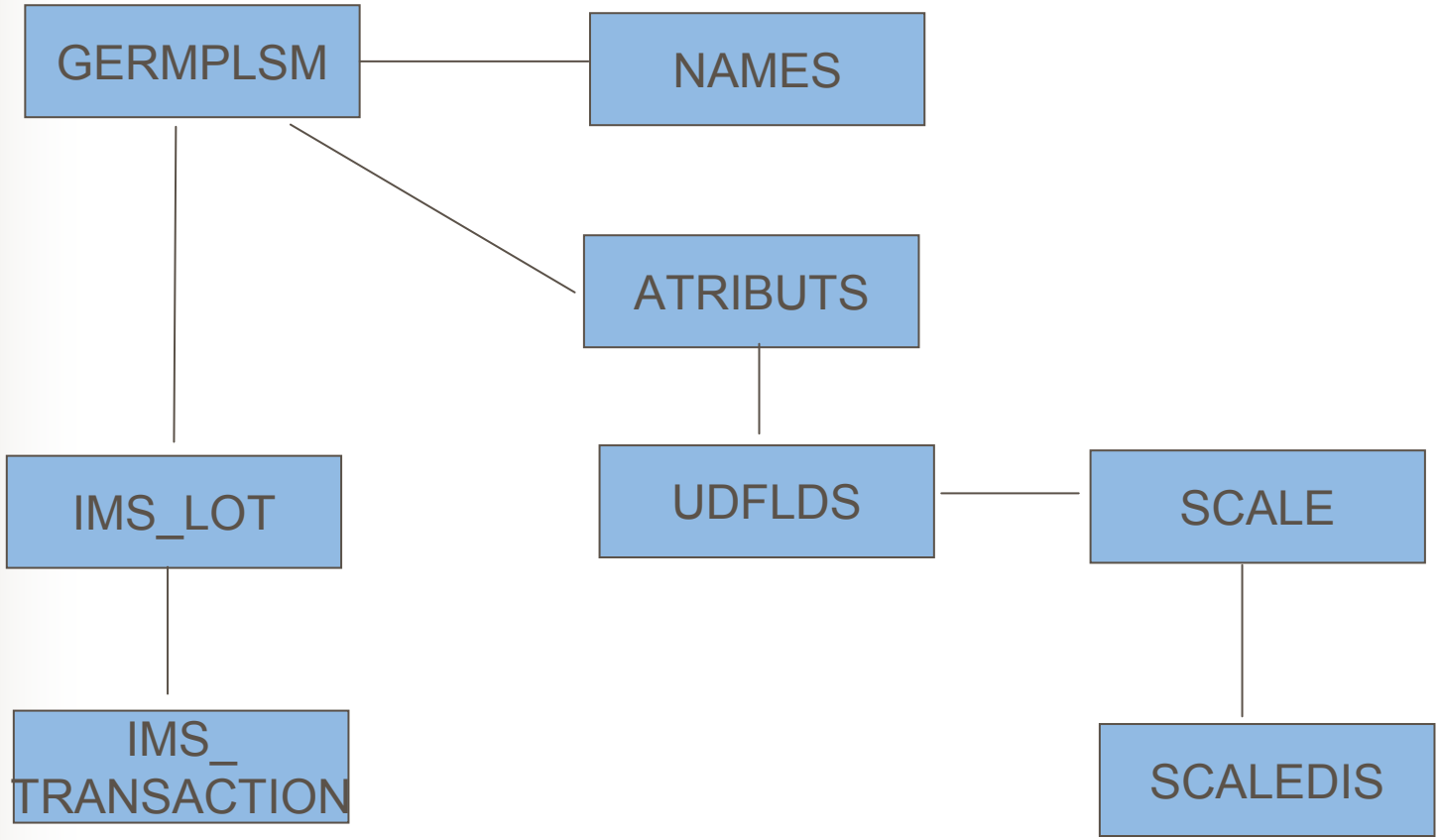
- *What should be the name state/ status of local names in their original form **-or-** of UNICODE names?*
  - NSTAT - Number indicating the storage type and status of the name
  - 1 – Preferred name (must be ASCII) → ***Can we allow unicode names as the preferred one?***
    - ***i.e. Lao script, cultivars with no English names***
  - 10 - UNICODE names which are not preferred
    - Temporarily being used

Mapping of Genebank Concepts to ICIS

- Seed inventory data
  - Seed lot information and transaction data are stored in ICIS IMS_LOT and IMS_TRANSACTION tables
- Distribution data
  - Request details such as the requestor information, plot details, origin, MTA status, etc are stored in DMS

Rice Science for a Better World
INTERNATIONAL RICE RESEARCH INSTITUTE

Relationship Diagram: Names, Passport, Inventory



Valid Values

Rice Science for a Better World
International Rice Research Institute

Advantages of Adopting ICIS Technology

- Genebank data integration to other institutional databases
  - Link to germplasm from other locations
  - Better use of essential information
- Easy access to germplasm generation information
  - Management of genetic identity
  - Tracking of germplasm history
- Open access to genebank data through development of web interface/ search engine

Genebank Needs

- Development of web interface for more effective data sharing and exchanging

  Web interface that :

  - Searches for passport and phenotypic data
  - Has an option to query only the studies that came from Genebank
    - Create a STUDY TYPE that corresponds to genebank characterization data (?)
  - Handles seed request

Current Web Search

Passport data - Characteristics of collection site

Morpho-agronomic data

- Anther length (mm) - wild
- Apiculus color at post-harvest
- Apiculus color at reproductive
- Auricle color at vegetative
- Awn color
- Awn length (mm) at anthesis - wild
- Awn presence at reproductive
- Blade color at vegetative
- Blade pubescence at vegetative
- Basal leafsheath color at vegetative
- Collar color at vegetative - cultivated
- Chromosome number - wild
- Culm angle at reproductive
- Culm diam. of basal internode at reproductive
- Culm length at reproductive
- Culm number (code) at reproductive - cultivated
- Culm strength at harvest - wild
- Culm strength at reproductive - cultivated
- Date at first flowering
- Date of seeding
- Distance (mm) panicle to spikelet at 7 DAA - wild
- Endosperm type
- Flag leaf angle at 7 DAA - wild
- Flag leaf angle at reproductive - cultivated
- Grain length (mm)
- Grain thickness (mm) - wild
- Ligule pubescence - wild
- Ligule shape
- 2nd leaf length (cm) - wild
- Leaf length - cultivated
- Lemma and palea color at anthesis - wild
- Lemma and palea color at post-harvest
- Lemma and palea pubescence
- Leaf senescence
- Leaf texture at 75 DAG - wild
- 2nd leaf width (cm) - wild
- Leaf width - cultivated
- 2nd leaf ligule length (mm) at 7 DAA - wild
- Maturity - cultivated
- Node color
- Panicle axis at reproductive
- Texture of panicle axis at 7 DAA - wild
- Panicle exsertion at reproductive
- Panicle length at post-harvest
- Panicle length at reproductive - wild
- Population composition - african cultivated
- Panicle shattering
- Panicle threshability
- Panicle type
- Rhizome/stolon formation - wild
- Seed coat color
- Scent - cultivated

ate
collected
tics
ristics
nation
osition
ity
(sq. meters)

Rice Science for a Better World
INTERNATIONAL RICE RESEARCH INSTITUTE

STEP 2 - Select conditions

Germplasm name list

Browse... Send File Pls. limit your entries to 1000!

Accession list

Browse... Send File Pls. limit your entries to 1000!

Specify conditions

Passport data- Basic

- ☐ Acquisition date
- ☐ Donor
- ☐ Source country
- ☐ Seed Source (country)
- ☐ Seed Source (station)
- ☐ Previous designation

Passport data - Collecting team

- ☐ Collecting institute
- ☐ Collector(s) Firstname
- ☐ Collector(s) Lastname
- ☐ Sender's Firstname
- ☐ Sender's Lastname
- ☐ Species
- ☐ Scientific name
- ☐ Classification of O. sativa samples
- ☐ Variety name
- ☐ Collection number
- ☐ Funding agency

STEP 3 - Specify conditions

Rules for specifying conditions

1. Your condition must of the form **<OPERATOR> <VALUE>**.
2. Default OPERATOR is = when omitted.
3. Valid OPERATOR are:
   { = , > , < , >= , <= , LIKE , IN, BETWEEN }
   a. Format for IN : **IN (<VAL1>,<VAL2>,...)**
   b. Format for BETWEEN : **BETWEEN <VAL1> AND <VAL2>**
4. VALUE should NOT be enclosed in quotes.
5. Click here to view expected VALUES for each descriptors.

Example:

Variety name	=MALAGKIT
Apiculus color at post-harvest	020
Apiculus color at reproductive	052
Auricle color at vegetative	000
Grain length (mm)	> 5
Grain thickness (mm) - wild	<=6
Grain width (mm)	<7

Maturity - cultivated	>120
Seed coat color	=080

Next step

Data Query results

June 5, 2007

IRGC Acc.	Variety name	Species name	ACQ_DATE	DONOR_CODE	ORI_COUNTRY	PREV_NAME	SENDER_LNAME	SENDER_FNAME	VARLINE_TYPE	COLL_INST
53592	AC2864	O. SATIVA	05-MAR-80		IND		ROY	J. K.		
48977	ADONG	O. SATIVA	20-APR-79		MYS		WADA	G.		
79802	ADONG	O. SATIVA	29-DEC-89		MYS		VAUGHAN	C/O DUNCAN A.	T	MARDI
79803	ADONG	O. SATIVA	29-DEC-89		MYS		VAUGHAN	C/O DUNCAN A.	T	MARDI
79804	ADONG	O. SATIVA	29-DEC-89		MYS		VAUGHAN	C/O DUNCAN A.	T	INTERNATIONAL RICE RESEARCH INSTITUTE (IRRI)
79804	ADONG	O. SATIVA	29-DEC-89		MYS		VAUGHAN	C/O DUNCAN A.	T	ARC
60201	ADONG HITAM(H)	O. SATIVA	23-MAR-82		MYS		WADA	G.		
71481	ADONG PULUT	O. SATIVA	11-APR-85	ARC 1240	MYS		TSEU	CHRISTOPHER	T	AGRICULTURE RESEARCH CENTRE, TUARAN, SABAH
71481	ADONG PULUT	O. SATIVA	11-APR-85	MARDI NO. 5582	MYS		TSEU	CHRISTOPHER	T	AGRICULTURE RESEARCH CENTRE, TUARAN, SABAH
71481	ADONG PULUT	O. SATIVA	11-APR-85	ARC 1240	MYS		TSEU	CHRISTOPHER	T	AGRICULTURE RESEARCH CENTRE, TUARAN, SABAH
71481	ADONG PULUT	O. SATIVA	11-APR-85	MARDI NO. 5582	MYS		TSEU	CHRISTOPHER	T	AGRICULTURE RESEARCH CENTRE, TUARAN, SABAH
71481	ADONG PULUT	O. SATIVA	11-APR-85	MARDI NO. 5582	MYS		TSEU	CHRISTOPHER	T	INTERNATIONAL RICE RESEARCH INSTITUTE (IRRI)

Genebank Needs

- ICIS module that handles germination result/ viability data
  - Viability and seed inventory data always go together (i.e. selection of planting material for regeneration and distribution)
  - Should it be a variate?
  - Another object in IMS?
  - Or a field in IMS_LOT?

Rice Science for a Better World
INTERNATIONAL RICE RESEARCH INSTITUTE

Other Genebank Concepts not yet in ICIS

- Taxonomy
  - Status: Uses SCALEDIS
- Mission
  - Status: Stored as an attribute value
    - Funding agency, country, start date, end data

Current Activities

- Test and implement full migration of GRIMS backend database from Oracle to PostgreSQL
- Use of hand-held device for field trials
- Improve the data quality of genebank data in IRIS especially for the germplasm group and source information
- Integration of documents and images to IRIS
- Integration of Herbarium Database to IRIS
  - Inherits MGID's passport data

Rice Science for a Better World
INTERNATIONAL RICE RESEARCH INSTITUTE

Acknowledgement

- Dr. Ruaraidh Sackville Hamilton
- Dr. Thomas Metz
- Corina Habito
- Adel Alcantara
- Pola de Guzman
- Soccie Almazan
- Ato Reaño

Thank you!

Rice Science for a Better World
INTERNATIONAL RICE RESEARCH INSTITUTE