# 

CORE

## **Ontology-driven International Maize Information System (IMIS) for** Phenotypic and Genotypic Data Exchange

### <u>Rosemary Shrestha<sup>1</sup></u>, Hector Sanchez<sup>1</sup>, Claudio Ayala<sup>1</sup>, Peter Wenzl<sup>1</sup>, and Elizabeth Arnaud<sup>2</sup>

<sup>1</sup>IRRI-CIMMYT Crop Research Informatics Laboratory (CRIL) - Centro Internacional de Mejoramiento de Máiz y Trigo (CIMMYT), Apdo. Postal 6-641, 06600 Mexico, D.F., Mexico; <sup>2</sup>Bioversity International, via dei Tre Denari, 472/a, 00057 Maccarese, Rome, Italy

#### The International Crop Information System (ICIS) Model

- Open-source and open-licensed generic crop information system. Ο
- Developed by the Consultative Group on International Agricultural Research (CGIAR), national agricultural research and extension systems, agricultural research institutes, and private-sector partners.
- Designed to fully document germplasm genealogies with associated meta-data such as passport data and to accurately cross-link germplasm entries with associated experimental observations from evaluations undertaken in the field, greenhouse, or laboratory.

#### The Crop Ontology (CO)

- The CO provides controlled vocabulary sets for several  $\bigcirc$ economically important plant species and facilitates biocurators working in genebanks of plant genetic resources (PGR) and crop breeding data curation and annotation.
- It also facilitates data exchange within and between Ο globally distributed databases, developing userfriendly query interface for searching crop data and allowing researchers to perform comparative studies across crops.





Figure 1. International Crop Information System (ICIS) model for addressing the problem of ambiguous germplasm identification, difficulty in tracing pedigree and lack of integration between genetic resources, breeding, evaluation, utilization and management trait data.

#### The International Maize Information System (IMIS)

- The IMIS is the public maize implementation of ICIS and CIMMYT's flagship Ο germplasm database.
- Contains about 33,000 germplasm ID (GID) entries with thousands of Ο associated data points in hundreds of experimental studies, including many phenotypic observations and growing numbers of genotypic measurements.
- Query interface is available for researchers to search germplasm and Ο associated maize breeding data.
- Many of these data sets are published in IMIS and initiated sharing data with Ο collaborating databases such as MaizeGDB.

The maize-trait ontology is developed based on the traits that are in the Data Ο management system (DMS) trait table of IMIS database along with the traits that are described in maize crop descriptor and GCP datasets.

	CO_322 Ontology Browser It is controlled vocabulary to describe maize traits based on IMIS and GCP datasets.						
Statistics Version: 2.0.2 OLS at EBI Version: 2.0.2 Ontologies Loaded: 50 Terms Loaded: 83308 Last updated: Wed Sep 15 09: 12:06 PHT 2010 See loaded statistics here	CGIAR maize trait ontology agronomic trait physiological trait chlorophyll content dabscisic acid content dabscisic acid to sugar sugar content ducose content osmotic adjustment osmotic potential proline content sucrose content normalized difference ve stomatal conductance leaf temperature stable oxygen isotope e total mineral content total nitrogen content stable nitrogen isotope e total nitrogen isotope e total nitrogen content stable nitrogen isotope e total mineral content the stable nitrogen isotope e mater index canopy temperature mater flowering days to silking male flowering quality trait		Help (show) - User's comm Relations days to silk Term Informa ID: CO_322 Name: day Associated in definition comment preferred name_ related synonym xref_definitior xref_analog xref_analog	ing is_a female flowering tion 2:0000031 Zoom s to silking formation Number of days from sowing to when 50% of the plants silk. This is determined on a daily basis by counting the number of plants with at least one silk extruded divided by the total number of plants in the plot. Unit/precision : 1 day. days to silking DTSLK CO:rs"ISBN:9290431891" TO:0000658 IMIS_TRAITID:2018			
	Legend: is a • develops from	₽ part of ■ other	Paths to Root	Paths to Root:  Child relationships: CGIAR			

Figure 3. A screen capture of the GCP Crop Ontology Lookup Service (OLS) browsing maize trait ontology. The information for the highlighted trait "Days to silking" is shown such as trait ID, trait description, synonym.

#### Implementation of the CO in the IMIS database

- The IMIS traits are mapped to those of maize-trait ontology and linked them to ontology Ο databases by term (trait) identifier.
- Similarly, the scales for trait measurement, locations for germplasm collection [which Ο are listed in the table of Genealogy Management System (GMS) of IMIS] and experimental sites are mapped to the generic ontologies such as scale ontology, gazetteer ontology.

Select your Column neaders below		sample
Find a Column Header here 🤔 Clear 🎤 Factors 🎤 Variates 🎤 Study Features		Headers
Headers A		BreedingProgram_BreedingProgram(BreedingProgram_C)
Season(Season(C)	~	NameOfSeries_NameOfSeries(NameOfSeries_C)
Series_Series(Series_C)		STUDY_Name(NameOfExperiment_C)
STUDY_Name(NameOfExperiment_C)		sname .
VARIETY_Entry_Number(Entry Number_N)		🗐 studyid
VARIETY_GID(GID_N)		GrainYieldKg_Kg(GrainYieldKg)
∃ Fields For: Study Feature (12 Items)		1
sdate		
T shierarchy		
▼ sname		
Sstatus	Add ->	>
V studyid		
stype		
title		
🔄 userid		
∃ Fields For: Variate (14 Items)		
GrainYieldKg_Kg(GrainYieldKg)		
MaturityDays_Days(MaturityDays)		
PlantStandNo_Int(PlantStandNo)		
RankNo_Real(RankNo)	-	



Report Format Builder Query Builder Query Results Hele

ews	A Cut Copy Clipboard	alibri B <i>I</i> <u>U</u>	A - A		E E PT →	Refresh All *	New ∑ Totals Save ♀♀ Spelling Delete ♀   More ♀ Records	A A Filter	Selection - Advanced - Toggle Filter t & Filter	Size to Switch Fit Form Windows * Window	Image: Base Seplace       Find       Image: Base Seplace       Image: Base Seplace	
Access 0	Objects 💿 «	TRAI	FID ◄	TRNAME -	Ontology	A SCTYPE	- SCNAME -	Scales OLS	•			
bles	*		2012 Con	nmonRust1_5	CO_322:000085	i C	(1-10)	CO_042:0000	446			
BIBREF	FS		2010 Bro	wnSpot1_5	CO_322:000008	i C	(1-5)	CO_042:0000	457			
	of Traits scales complete		2036 Gra	yLeafSpot1_5	CO_322:000088	C C	(1-5)	CO_042:0000	457			
	<i>c</i>		2017 Cun	vulariaLeafSpot1_5	CO_322:000008	) C	(1-5)	CO_042:0000	457			
			2048 Pha	esLeafSpot1_5	CO_322:000009	) C	(1-5)	CO_042:0000	457			
DATA_I	N		2066 Tars	spotComplex1_5	CO_322:000009	C	(1-5)	CO_042:0000	457			
DATA_1	Ţ		2042 Leat	fBlightTurcicum1_5	CO_322:0000092	2 C	(1-5)	CO_042:0000	457			
DATAT	TR		2114 See	dlingDiseasePer	CO_322:000009	C	Percentage	CO_042:0000	139			
DMSAT	TTR		2129 Viru	usMalDeRioCuartoPer	CO_322:0000094	C	Percentage	CO_042:0000	139			
	IDS		2125 Viru	usMaizeRayadoFinoPer	CO_322:000009	5 C	Percentage	CO_042:0000	139			
			2124 Viru	usMaizeMosaicPer	CO_322:000009	5 C	Percentage	CO_042:0000	139			
EFFECI	1		2126 Viru	usMaizeRoughDwarfPer	CO_322:000009	7 C	Percentage	CO_042:0000	139			
FACTO	DR		2128 Viru	usMaizeStripePer	CO_322:0000098	3 C	Percentage	CO_042:0000	139			
IMIS TR	RAITS		2069 Viru	usMaizeDwarfMosaicPe	CO_322:0000099	) C	Percentage	CO_042:0000	139			
INSTITU	UT		2070 Viru	usMaizeStreak1_5	CO_322:000010	) C	(1-5)	CO_042:0000	457			
LEVEL_	C		2127 Viru	usMaizeStreakPer	CO_322:000010	) C	Percentage	CO_042:0000	139			
LEVEL	N		2130 Viru	usSugarcaneMosaicPer	CO_322:000010:	C	Percentage	CO_042:0000	139			
			2104 Tryp	otophanPer	CO_322:000010	C	Percentage	CO_042:0000	139			
LEVEL	.'		2099 Lysi	never	CO_322:0000104		Percentage	CO_042:0000	139			
LEVELS	5		2123 Tryp	otophanPer_Endosperm	CO_322:0000100		Real	CO_042:0000	446			
OBSUN	NIT		2068 Tryp	otopnanPer_wholeGrai	CO_322:000010		Percentage	CO_042:0000	139			
OINDE	EX		2044 Lysi	nePer_wholeGrain	CO_322:0000108		Percentage	CO_042:0000	139			
Paste E	Errors		2095 Lysi	never_Endosperm	CO_322:000010		Percentage	CO_042:0000	139			
DEDSO	MIS		2093 Iron	1_ppm	CO_322:0000110		PPIVI	CO_042:0000	257			
PERSO			2131 ZINC	c_ppm	CO_322:000011		PPIM	CO_042:0000	237			
REFLIN	IKS		2000 Sele	toinBor	CO_322:0000110		Borcontago	CO_042:0000	120			
REPRES	STN		2052 Prot	teinPer WholeGrain	CO_322:000012		Percentage	CO_042.0000	129			
SCALE			2005 PT0	teinPer_Endosnerm	CO_322:000012		Percentage	CO_042.0000	139			
SCALE	CON		2107 Pro	VitaminA ugPerGram	CO_322:000012		Percentage	CO_042.0000	139			
SCALE	DIS		2108 Pro	VitaminA_pgr crorum	CO_322:000012	5 D	Mm	CO_042:0000	003			
SCALET	TAR		2040 Larg	eGrainBorer1 5	CO 322:000012	C C	(1-5)	CO_042:0000	457			
SCALE			2002 Aph	nid1 5	CO 322:000012	C C	(1-5)	CO 042:0000	457			
SIEFFE			2046 Mai	zeWeevil1 5	CO 322:000012	C	(1-5)	CO 042:0000	457			
STUDY			2071 We	evilTotNum	CO 322:0000130	) D	Real	CO 042:0000	446			
TGROU	UP		2072 We	evilWtLossPer	CO 322:000013	с	Percentage	CO 042:0000	139			
TMETH	IOD		2037 GS(	NDVI)	CO 322:0000132	2 C	Integer	CO 042:0000	447			
TRAIT			2064 Stor	matalConduc	CO_322:0000134	D	Real	CO_042:0000	446			
VADIAT	TF		2039 Infr	aredThermo	CO_322:0000142	2 D	Real	CO_042:0000	446			
, VARUAI			2087 Grai	inColor1 5	CO 322.000014	i C	(1-5)	CO_042:0000	457			

Figure 4. The Data Management System (DMS) table of the IMIS showing mapping traits to the maize\_trait ontology and also IMIS trait scales to the ICIS scale ontology. The curation is expended for linking phenotypic data to genotypic data by mapping the IMIS traits with molecular data such as QTL to the maize\_trait ontology

#### REFERENCES

View Edit

- Bruskiewich R, Senger M, Davenport G, et al. 2008. The Generation Challenge Programme Platform: Semantic Standards and Workbench for Crop Science. International Journal of Plant Genomics Article ID 369601, 6 pp.
- McLaren CG, Bruskiewich RM, Portugal AM et al. 2005. The International Rice Information System. A Platform for Meta-Analysis of Rice Crop Data. Plant Physiol. 139:637-642.
- Shrestha R, Arnaud E, Mauleon R, et al. 2010. Multifunctional crop trait ontology for breeders' data: field book, annotation, data discovery and semantic enrichment of the

A DOT NO PROVIDENT	MEETITOP DE LOS DOB 43 JE	MEETING DEL DE DOB 43 INT	MEETINGS OF LOS DOD 43 IS1	10515	1170001 8
neer i -ma-emysiology	MESTIZUS DE DPS POB 4315	MESTICOS DE LES POBIASTIST	MESTICOS DE LPS POB 43151	1015	1173001.0
MMYT-HQ-Physiology	MESTIZOS DE LPS POB 43 IS	MESTIZOS DE LPS POB 43 IS1	MESTIZOS DE LPS POB 43151	10515	1184836.7
MMYT-HQ-Physiology	MESTIZOS DE LPS POB 43 IS	MESTIZOS DE LPS POB 43 IS1	MESTIZOS DE LPS POB 43 IS1	10515	1190505.0
MMYT-HQ-Physiology	MESTIZOS DE LPS POB 43 IS	MESTIZOS DE LPS POB 43 IS1	MESTIZOS DE LPS POB 43 IS1	10515	1207140.5
MMYT-HQ-Physiology	MESTIZOS DE LPS POB 431S	MESTIZOS DE LPS POB 43 IS1	MESTIZOS DE LPS POB 43151	10515	1236435.8
MMYT-HQ-Physiology	MESTIZOS DE LPS POB 43 IS	MESTIZOS DE LPS POB 43 IS1	MESTIZOS DE LPS POB 43151	10515	1237967.4
MMYT-HQ-Physiology	MESTIZOS DE LPS POB 43 IS	MESTIZOS DE LPS POB 43 IS1	MESTIZOS DE LPS POB 43151	10515	1243740.3
MMYT-HQ-Physiology	MESTIZOS DE LPS POB 431S	MESTIZOS DE LPS POB 43 IS1	MESTIZOS DE LPS POB 43151	10515	1254960.5
MMYT-HQ-Physiology	MESTIZOS DE LPS POB 43 IS	MESTIZOS DE LPS POB 43 IS1	MESTIZOS DE LPS POB 43 IS1	10515	1255769.3
MMYT-HQ-Physiology	MESTIZOS DE LPS POB 43 IS	MESTIZOS DE LPS POB 43 IS1	MESTIZOS DE LPS POB 43 IS1	10515	1258816.8
MMYT-HQ-Physiology	MESTIZOS DE LPS POB 4315	MESTIZOS DE LPS POB 43 IS1	MESTIZOS DE LPS POB 43 IS1	10515	1259812.1
MMYT-HQ-Physiology	MESTIZOS DE LPS POB 4315	MESTIZOS DE LPS POB 43 IS1	MESTIZOS DE LPS POB 431S1	10515	1262233.6
MMYT-HQ-Physiology	MESTIZOS DE LPS POB 43 IS	MESTIZOS DE LPS POB 43 IS1	MESTIZOS DE LPS POB 43 IS1	10515	1313423.5
MMYT-HQ-Physiology	MESTIZOS DE LPS POB 43 IS	MESTIZOS DE LPS POB 43 IS1	MESTIZOS DE LPS POB 43 IS1	10515	1315180.1
Page 1 of 15	>>1 O Save as CSV				Displaying results 1 - 25 of 36

Free theme builder license

Next Step >> Retrieve Data

Hide/Show Column

Figure 2. A screen captures of the International Maize Information System (IMIS) query interface for searching genotypic or phenotypic data. In the figures, query steps and result for the trait data "grain yield" are shown in detail.

literature. AoB Plants. plq008.

#### **AVAILABILITY**

- The ICIS information is available at http://www.icis.cgiar.org. Ο
- The IMIS query interface is available at http://imis.cimmyt.org/ for searching maize Ο data.
- The GCP crop ontology and platform are described on the pantheon website at http://pantheon.generationcp.org.
- The GCP Lookup Service is available at http://cropontology.org to browse ontologies. Ο
- CropForge software project management site is available at The Ο http://cropforge.org/projects/gcpontology for the collaborators and users.