

Vancouver Domain Modeling and Data Templates Workshop report

Contents

- [1 Synopsis](#)
- [2 Activities](#)
- [3 Summary of Outputs](#)
- [4 Proposed presentation of Platform prototype – GCP annual meeting Sept 2005](#)
 - [4.1 CMTV Task 28-1&9 Guy Davenport/Andrew Farmer](#)
 - [4.2 GCP Query Builder Task 28-2 Richard Bruskievich](#)
 - [4.3 HPC Structure Task 27 Reinhardt Simon](#)
 - [4.4 DIVA - Task 28-4 Reinhardt Simon](#)
 - [4.5 Mondrian Demonstration –Task 28-6 Edwin Rojas](#)
 - [4.6 Gene Catalogue Task 28-10 Richard Bruskievich \(Liaison to Tasks 31 & 32\)](#)
 - [4.7 LIMS demonstrations Task 28-7&8 Akin, Jayshree and Maha](#)
 - [4.8 Template Demonstration Task 25 Guy Davenport](#)

Synopsis

Two day meeting involved the Principals of the five GCP domain model editorial teams, plus the Task Leader of the Data Template Task.

Activities

- Presentation of the data template task outputs
- Review and harmonization of GCP domain models

- Discussion of implementation strategy for domain model in GCP platform (Java) software
- Identification of delivery targets for Rome meeting

Summary of Outputs

Guy Davenport presented brief overview of data template task

Domain modeling status discussed

Core domain models (Richard Bruskiwich, IRRI) reviewed the core "metadata" model design and philosophy

Each editorial team presented their current models

Passport (Tom Hazekamp, IPGRI)

Germlasm/Phenotype/Genotype (Richard Bruskiwich, Graham McLaren and Thomas Metz, IRRI)

Mapping (Manuel Ruiz, CIRAD)

Location/Environmental Data (Reinhard Simon)

Functional Genomics (Masaru Takeya)

One-on-one meetings between Richard and each editorial team leader were held to move theme specific models to use metadata core.

Richard refined models during meeting and posted refined models to CropForge CVS

Implementation strategy for GCP domain model driven platform discussed

Guy Davenport and Reinhard Simon proposed use of Eclipse RCP and Plugin framework

Guy converted GCP platform layers into plugins (posted to CVS)

Identification of delivery targets for Rome meeting

GCP team present decided to attempt a prototype integration of CMTV, ICIS (QueryBuilder), DIVA and possibly some selected Java-based genomics tool(?) into a single platform prototype using the GCP domain models, Eclipse RCP, ISYS and other pertinent technologies, in time for the Rome meeting. In addition, other GCP tasks and platform projects may present their wares. The following section of this page outlines a proposed target set of

demonstrations.

Proposed presentation of Platform prototype – GCP annual meeting Sept 2005

The objective of the presentation will be to show integration of platform components using the GCP data model connecting to several data sources from different institutes and following a single biological use-case across different data domains using the different applications.

CMTV Task 28-1&9 Guy Davenport/Andrew Farmer

Demonstrate the ability of CMTV to access map data from CMAP, TROPGENE, GBROWSE using the GCP data model. Musa maps from CMAP and TROPGENE can be used, but rice maps from CMAP and GBROWSE should also be shown to demonstrate the generality of the architecture to different CGP partners.

The application should be able to export a list of markers to the GCP query-builder.

We should also aim for a Java Webstart installable package.

GCP Query Builder Task 28-2 Richard Bruskwiech

Given a list of markers, return all germplasm which has been genotyped at one or more of the markers. Given the GIDs return phenotype/genotype/passport data Export GIDs and genotyping data to HPC Export location data to DIVA

HPC Structure Task 27 Reinhardt Simon

Run Structure on the HPC using data exported from the query-builder. Export results the DIVA for spatial correlation of Structure groups by location

DIVA - Task 28-4 Reinhardt Simon

Given Structure output and LocIDS use DIVA to plot locations and genotype classes. Starting from DIVA, use GIDs to query the GCP databases and retrieve related phenotype or allele data (additional attributes) for plotting.

Mondrian Demonstration –Task 28-6 Edwin Rojas

Gene Catalogue Task 28-10 Richard Bruskiwich (Liaison to Tasks 31 & 32)

From CMTV markers and maps, identify a genomic map interval of loci (perhaps, syntenic region in rice) to obtain a list of candidate gene loci. Given this list of genes, filter out based on gene annotation: GO terms, gene family membership, perhaps some significant functional genomics (e.g. gene expression) experimental results. Perhaps send log ratio of intensities of gene expression back to CMTV as a feature plottable on the original (or cross-linked) map(?).

LIMS demonstrations Task 28-7&8 Akin, Jayshree and Maha

Template Demonstration Task 25 Guy Davenport

Loading genotype information into Germinate