**BIEN3.0 database breakaway session – Tuesday**

**Brad’s notes**

Intros:

Aaron –Comp Sci student City College, BSc profgram

Shash – Comp Sci student, Bradley U., MSc program

Bob

A. Comparison: plots vs. species

* two types: specimens, plots
* Specimen example
  + one record, multiple determinations
* Plot example:
  + Location
  + Plot measurements of size, layout
  + Soil measurements
  + Species
    - Species list plus cover & species-level measurements OR
    - Individuals, dbh, height, etc.
  + Discussion: complications of morphospecies, undetermined individuals, unsampled indivivuals (e.g. huge trees and lianas

B. Databases:

* Overview of Vegbank
  + Key features
* Biodiversity data structure:
  + Plot
  + PlotObservation
  + Project
  + PlotInterpretation (vegetayion clasification)
* VegBank ERD
  + Issues:
    - Traits? Not in but could be accommodated
    - Embargoes, fuzzing of data
    - Method metadata
    - Observation table:
      * Plot-level attributes
      * large number of rarely-used fields
      * observationSynonym=redundant entries due to digitization
    - TaxonObservation table:
    - Attributes of individuals or taxa
* Discussions: difference between aggregate and individual observations

Brad & Nick: Reality checkL Goals for today and tomorrow:

* Today: finish overview of vegBank
* Tomorrow:
  + Morning
    - VegX
    - Taxonomy overview
    - User interface examples
  + Afternoon
    - Review of use cases
    - Evaluation of VegX, VegBank: will they accommodate all BIEN use cases?

Bob

* Party
* Reference
* Taxon Concepts

Peter

* Demonstration of Tropicos plot data entry tool

Plan for tomorrow:

* Morning
  + VegX (Nick)
  + VegBranch (Bob)
  + CI-TEAM tool ‘desk(Erik)
  + Bradley University/CTFS overview? (Steve/Shash)
  + Overview: current state of bien2 (Brad/Steve)
  + Taxonomy issues (Brad)
* Afternoon
  + Bien3.0 architecture
  + Compile use cases (Brad)
  + Review of use cases
  + Evaluation of VegX, VegBank: will they accommodate all BIEN use cases?

**Wednesday**

General discussion: change plan to accommodate Aaron’s absence:

Morning:

* State of Bien2
* Overview of proposed Bien3.0 architecture

Afternoon

* Taxonomic issues (brad)
* VegX (Nick)
* TEAM tool (Erik)
* VegBranch (Bob)
* Review of use cases
* Evaluation of VegX, VegBank: will they accommodate all BIEN use cases?

**Overview of BIEN2.0**

Steve: BIEN1 - Specimens, plot into data warehouse

Brad: overview of current bien2.0

Bob: overview of proposed bien3.0 architecture

Nick: must prioritize

1. Core db
   * Challenge of confederation: will the final schema accommodate all potential data. If not, do we:
     + Restrict data sources
     + Expand schema
2. Analytical db
3. Data loaders, schema mapping

Mark: problem of confederation

Core db options:

1. New schema from scratch
2. New schema based on VegX
3. Vegbank verbatim
4. VegBank modified according to use cases

Core db platform:

separate decision

Brad, Mike: prefer new technology

Exchange technology: switch from current xml export/import (from VegBranch) to VegX

Possible spin-off of model for updating VegBank (out of scope for BIEN however)

Data ingest/edit:

* Current model allows only dataset level bulk imports with feedback to data providers
* Should we allow for record-level edit/update by user
* Dataset level ingest

Workflow:

* Build and modify Bien3.0 based on VegBank
* Load directly all existing data (not via VegX)
* Build VegX loaders for ingesting additional data
* Build validation pipelines
  + Taxonomy/TNRS
  + Geovalidation
* Build analytical database

Initial data load:

* Load all existing data directly to VegBien

Versioning:

* Cron refresh of analytical views
* User cites version, date accessed

**Afternoon:**

* Geovalidation steps: how to build into pipeline
* Check: geoplanet (yahoo web service)

VegX overview (Nick)

Review of Use Cases

* compare use cases/requirements with VegBank and VegX
* are requirements met?
* If not, can schemas be modified? Or do we need to chuck altogether?

Consensus:

Both VB and VX will need minor modifications, but nothing major