Issues with CVS plot data.

1. locality\_\_@DwC\_\_@vegpath.org: In some cases (eg 040-04-0144) this is actually presenting “Location Narrative”, and not “Author Location” as needed, whereas in other cases where Location narrative is not populated then we see the Author location data as needed. This should be strictly “Author Location”. **[Seems easy to fix]**
2. I do not know the definitions of the fields georeferenceProtocol\_\_@DwC\_\_@vegpath.org & “geovalid\_bien”. However, they were blank for all records and I suspect they relate to geovalidation, in which case they should not be blank. **[Seems simple, but not sure. Need a report from Aaron. I think this is a problem with other datasets as well.]**
3. It was good to see communityConcept.name\_\_@VegX\_\_.communityDet@vegpath.org populated, but some critical associated fields were not present, including the Community Code and the fit and confidence values. Community code is very important and is needed for both VegBank and CVS. **[At a minimum we need the CommunityCode from CVS, and preferably Fit and confidence. In VegBank Fit and confidence are in the CommunityInterpretation table. Not sure if CommunityCode is to be found in VegBank – need to ask Michael.] duplicat**
4. Morphospecies have non-alphabetical characters scrubbed out. For example Hypericum [graveolens + mitchellianum] is rendered as Hypericum [graveolens mitchellianum]. Those extra characters in the morphospecies need to be retained. This may be a problem with all morphospecies in BIEN. **[Seems easy]**
5. For every record “identifiedBy\_\_@DwC\_\_@vegpath.org” has the value “Robert Peet” and “dateIdentified\_\_@DwC\_\_@vegpath.org” has a value of “10/1/2008”. I think currently you are using the person who contributed the dataset (Peet) and the date of the contribution of the dataset (2008). I need clearer definitions of these two fields, but it appears they refer to the person who Identified the plant and on which date. These fields do have a home in CVS and we need to point you to these. If for some reason the field is blank the default should be the person who collected the plot and the date of the collection. **[Seems easy]**
6. I see many data lines duplicated for no obvious reason **[?Symptom of deeper problem?]**
7. Do were really want to discard all soil data? **[I prefer to retain, though this is a bit awkward in the VegBank model. Discuss.]**
8. Cover values are given as the midpoint of a range with no indication of the range of the bin? A cover value of 0.505 seems very precise, but it is really the bin #2 in the CVS scale corresponding to 0.1-1% cover. This needs to be indicated in some way. It may be a problem with all the cover values in BIEN. **[Need to discuss.]**
9. The admittedly weird modules and subplots of the CVS protocol are very badly handled. In the current version it is not uncommon for a species to have on the order of 15 records recorded for a single plot. Some of these have different numbers of individuals and some have difference cover values. Unfortunately, all are recorded as having the areas of the full plot, but a module is usually smaller than a plot , say .01 rather than .1 ha. The area associated with the species and cover or count can refer to the module or the full plot. In addition there can be separate records for different size classes of trees with no indication of either the size or the area associated with the record. This we need to discuss so that you understand the structure of the data and whether it all needs to be retained.

**[This is very complicated. We can simplify by ignoring subplot data. We will need to walk through this with Aaron, Michael and myself, and possibly Brad and Brian]**

1. **[This is new. I observe that some, but not all, plots in VegBank dump do not have the plot name populated (**[**plotName\_\_@VegX\_\_.plot@vegpath.org**](mailto:plotName__@VegX__.plot@vegpath.org)**). I have not seen this problem in CVS, but then I have seen only a small dump of data]**