

VegBank - Bug #771

Loader rectification tools:Picklists: closed (and open?)

11/13/2002 01:51 PM - Michael Lee

Status:	New	Start date:	11/13/2002
Priority:	Normal	Due date:	
Assignee:	Michael Lee	% Done:	0%
Category:	userLoad	Estimated time:	0.00 hour
Target version:	3.0.0	Spent time:	0.00 hour
Bugzilla-Id:	771		
Description			
Develop rules for rectification			
Related issues:			
Is duplicate of VegBank - Bug #769: [676] :Loader rectification tools:Picklis...		Resolved	11/13/2002
Blocked by VegBank - Bug #770: [673] :Loader rectification tools:Picklists-b		Resolved	11/13/2002

History

#1 - 01/14/2003 12:33 PM - Michael Lee

I think that this bug deals with the mapping of one set of values into another set of values. I'm not sure if we want the user to be able to do this from data on upload (only applicable to NPS and XML -- incl. TurboVeg -- VegBranch takes care of this on its own), or if we are to specify values ahead of time (i.e. and XML doc mapping values onto closed list values). The former sounds a bit challenging (they could just use VegBranch), and the latter sounds fairly reasonable (we already are using a similar XML doc to validate values). The former feature is probably for version 2.0.0; the latter (which this bug represents, unless otherwise changed later) could be for version 1.0.0.

#2 - 04/19/2003 07:07 AM - Robert Peet

Not sure what this bug refers to.

Step one is creation of the picklists. I think these are now stable and we can mark this part as complete.

Step two is to make certain that plots being submitted have enteries match the appropriate picklists. NOte sure if this is in place yet, but it should be Gabe's job, so I will reassign this bug to Gabe.

#3 - 08/21/2003 02:53 PM - Gabriel Farrell

This is the creation of web ui that allows users to correct fields in the data they submit when vegbank has a problem with a particular data field(s).

So a user is loading a plot and they have a rocktype in their data that vegbank doesn't recognize. A screen pops up asking the user to add one of the acceptable rocktypes.

The picklists are defined in the database schema most easily accessible from java in the datamodel classes e.g. for a list of acceptable rocktypes org.vegbank.common.model.Plot.getRocktypePickList().

I am unsure if this also refers to doing the same for unrecognized plant or communities.

#4 - 08/26/2003 10:43 AM - P. Anderson

Some points:

- the same rock types, plants, and communities can be named differently
- typos are inevitable
- an interactive rectifier tool is possible
- the tool can make a best guess using a phonetic code (pronunciation)
- guess with a SQL LIKE '%abbreviation%' query too

We could generate phonetic keys from the names of our VegBank data (rocks, plants, etc.), then simply search that index for any mismatched input to make a guess.

SOUNDEX is a SIMPLE technique devised by Donald Knuth FOR FILING NAMES by the way they are PRONOUNCED rather than the way they are spelled.

METAPHONE, an algorithm described by Lawrence Philips in the December 1990 issue of Computer Language, works BETTER than Soundex because it knows the basic rules of English pronunciation but it uses more storage and computation.

Jakarta has open source Java code for both algorithms.

This tools is totally possible but definitely requires some work to be done right.

#5 - 08/27/2003 05:53 PM - Robert Peet

The above sounds like a great solution. We need something like this. How much time will it take to implement, and what additional help is needed?

#6 - 11/21/2003 11:49 PM - Michael Lee

- Bug 769 has been marked as a duplicate of this bug. ***

#7 - 11/21/2003 11:50 PM - Michael Lee

- Bug 770 has been marked as a duplicate of this bug. ***

#8 - 03/27/2013 02:15 PM - Redmine Admin

Original Bugzilla ID was 771