1. Use ReproZip to store data sets and algorithms in a virtual machine.
2. It is easy to replace data sets. If interfaces are preserved, then should be possible to change versions of libraries as well.
3. Be able to rerun the algorithm with ability to handle errors gracefully, e.g. be able to say “can’t handle these inputs”. Choosing these intelligently might be a topic of collaboration with Sarah. We might be able to put constraints on the program, e.g. must handle missing data without blowing up but with some informative error method. Original program must have the property that it can handle missing data in any of the source files but points out this missing data.
4. Name for this: Version Flexible Software.

Data sets or underlying software components change. What do I do now?

Results that appear in some paper X

Data Sets, e.g. particular version of annotations of genes or proteins in some organism, particular GO hierarchy

Custom algorithm that may depend on components, e.g. of scikit learn or Blast or whatever.