

The heterogeneity of observational papers makes every attempt to write a uniform catalog almost impossible. My aim is to build a new catalog selecting the best targets whose datasets were included in at least one of the three major exoplanets online databases: NASA Exoplanet Archive ([www.exoplanetarchive.ipac.caltech.edu](http://www.exoplanetarchive.ipac.caltech.edu)), Exoplanet Orbit Database ([www.exoplanets.org](http://www.exoplanets.org)), and Exoplanet Encyclopaedia ([www.exoplanet.eu](http://www.exoplanet.eu)).

I wrote a Python code that collects and selects the most precise measurement for all interesting planetary and orbital parameters, taking into account the presence of multiple aliases for the same target present in at least one of the databases. For each parameter, the code stores the corresponding reference paper link. For this reason, when the merging process is completed it could be possible to have a final dataset for each target which is not necessarily composed by consistent measurements. It is however not essential for our statistical purposes.

The code is able to download the source files from the three catalogs by use of VO ConeSearch connections and it retrieves the compulsory user preferences through a Graphic User Interface, which allows to choose the mass/radius range selection. It is also able to generate automatic plots that are commonly used in the exoplanetary community, but the user can retrieve and manipulate data at will.

Some of the problems I encountered while building this code were the following:

- There are different aliases by which the same planet is known. It would then be a problem to identify the same target and to merge the various datasets. I am currently trying to solve this problem by establishing a VO query to Simbad, as Marco suggested.
- I also needed to manually rename most of the columns in the existing catalogs, so that the code could effectively interpret them as the same quantity. I acknowledge the fact that this may be done differently, now that I know what a TAP service is and how to use it.
- The catalogs make dramatically different choices when inserting a new target in the database: some may choose to exclude the target if it is not characterised by precise measurements, or to add calculated values without proper flags to confirm that operation. This is beyond any computational effort I could make and could be solved only by developing a new, uniform, exoplanet database, with as less constraints as possible — or at least with an easily customisable capacity to filter data and adjust them to one's needs.