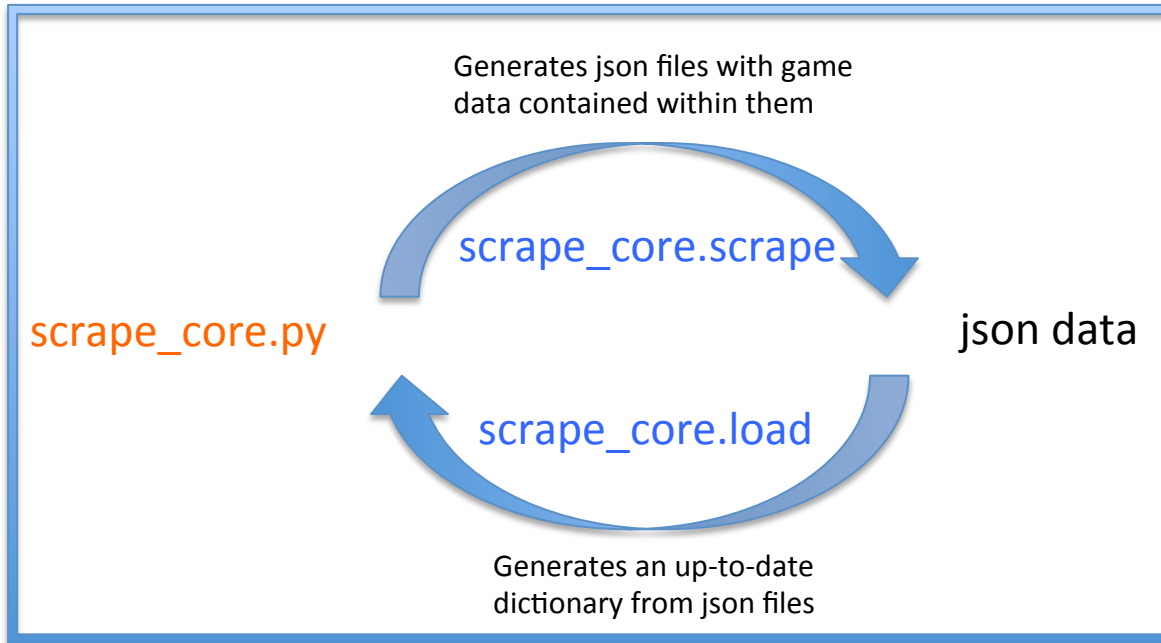


Web Scrape Procedure

Legend

File

Function



`scrape_cron.py`
Runs `scrape_core.py` routine
every xxx hours/minutes

Notes:

- 1) The generation of json files should avoid redundancy as much as possible. Needs a separate function within `scrape_core.py` which checks for json files not already within the current iteration of the database structure.
- 2) Needs tests to make sure it is functioning properly. Create `scrape_tests.py` to run (which will test out the procedure on a small subset of the data; perhaps on a single season or something).
- 3) Are we okay not using a sql database? Do we really want/need it?