

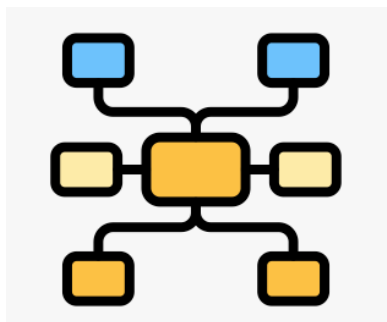


Algorithm 4: preprocess ref-DB (utils.py)

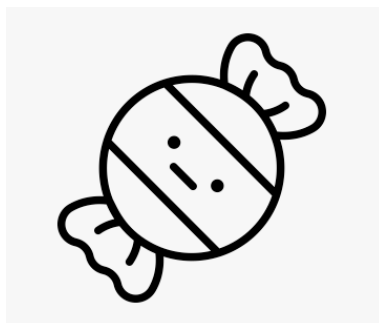
1. run_multisearch
2. remove_corr_organisms_from_ref



Python wrapper:
(make_training_data_from_sketches.py)



YACHT Core **algorithm 2,3**
(hypothesis_recovery_src.py)



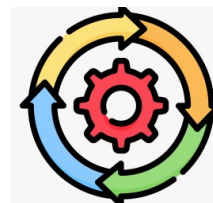
Executive wrapper for CLI
(run_YACHT.py)



Results



Re-format
(standardize_yacht_output.py)



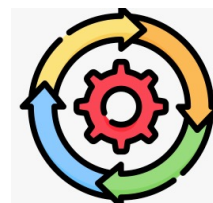
Sourmash wrapper:
(sketch_ref_genomes.py)
(sketch_sample.py)



Build FMH sketches



Download pre-built data



Python wrapper:
(download_default_ref_db.py)
(download_demofiles.py)
(download_pretrained_ref_db.py)

get_organisms_with_nonzero_overlap: alg 3, step 1-2
get_exclusive_hashes: alg 3, step 3
get_alt_mut_rate: not in alg 3, but will be included in results
single_hyp_test: algorithm 2 (alg3 step 6)
hypothesis_recovery: wrapper function for whole algorithm 3