

Improving Software Engineering in Academia

Samuel Grayson

Name	Domain	Distributed	Task over- head	Data par- allel	Provenance	Dependencies	Delete un- used mod- ules	Correct wrt changes in de- pen- den- cies
funnies		Distributed						
Luigi		Distributed						
Reena	Computational sci- ence	Distributed						
Yadage	Computational sci- ence	Distributed						
Snakemake	Computational sci- ence	Distributed						
Popper	Computational sci- ence							
Nix	Package man- age- ment	Distributed						
Bazel	Build sys- tem	Distributed						
Buck	Build sys- tem	Single server						
Make	Build sys- tem	Single server						

Name	Domain	Distributed	Task over- head	Data par- allel	Provenance	Dependency time	Delete un- used prod- ucts	Correct wrt changes in de- pen- den- cies
Scons	Build sys- tem	Single server						
Rake	Build Sys- tem	Single server						
Shake	BuildSystem	Single server						

- **CWL**: Just a language for going to other workflow managers.
- **Apache Flink**: Stream oriented.
- **Apache Airflow**: Too complex, designed for big-data.