



Our Constraints

We're based on volunteers, that come to hack for a few hours a week.

All tools we choose must be:

- **Easily installable**
(portable, no complex setup)
- **Maintainable**
(volunteers come and go)
- **DRY**
(no boilerplate!)
- **Low footprint**

We need a new ETL

Existing solutions were too much:

- Complex set-ups
- Too flexible
 - code based
 - unstructured - relies on conventions
 - no resource guarantees
- Too powerful
 - Too many features lead to steep learning curve
- Optimized for a different working point.

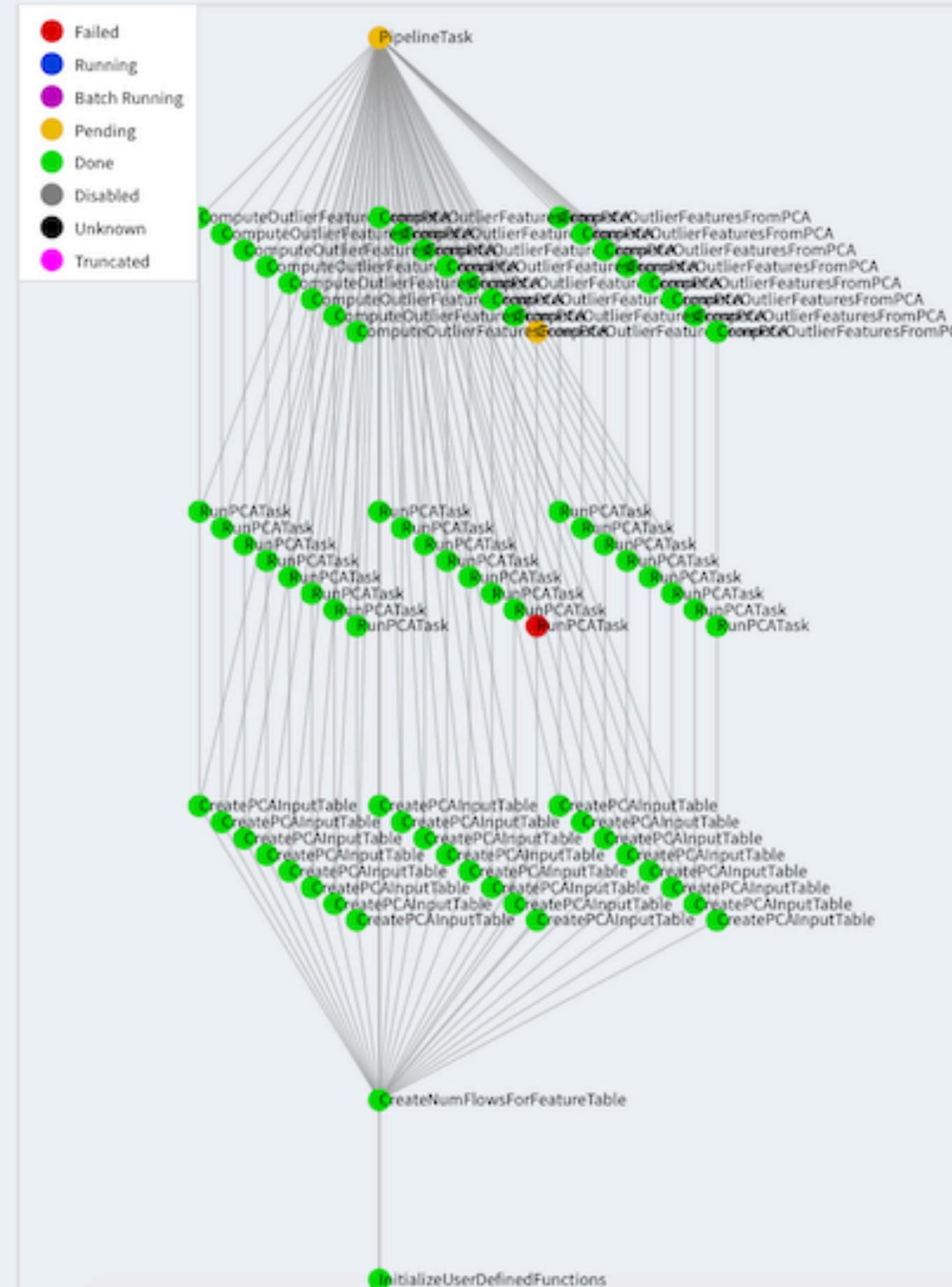
'few sources - deep processing'

VS.

'many sources - shallow processing'

PipelineTask(date=2017-02-07)

Dependency Graph



Our Constraints

We're based on volunteers, that come to hack for a few hours a week.

All tools we choose must be:

- **Easily installable**
(portable, no complex setup)
- **Maintainable**
(volunteers come and go)
- **DRY**
(no boilerplate!)
- **Low footprint**

