

Generalizable Analysis Pipelines for Handling Large HTA Submissions

Matthew Phelps & Nicolai Skov Johnsen

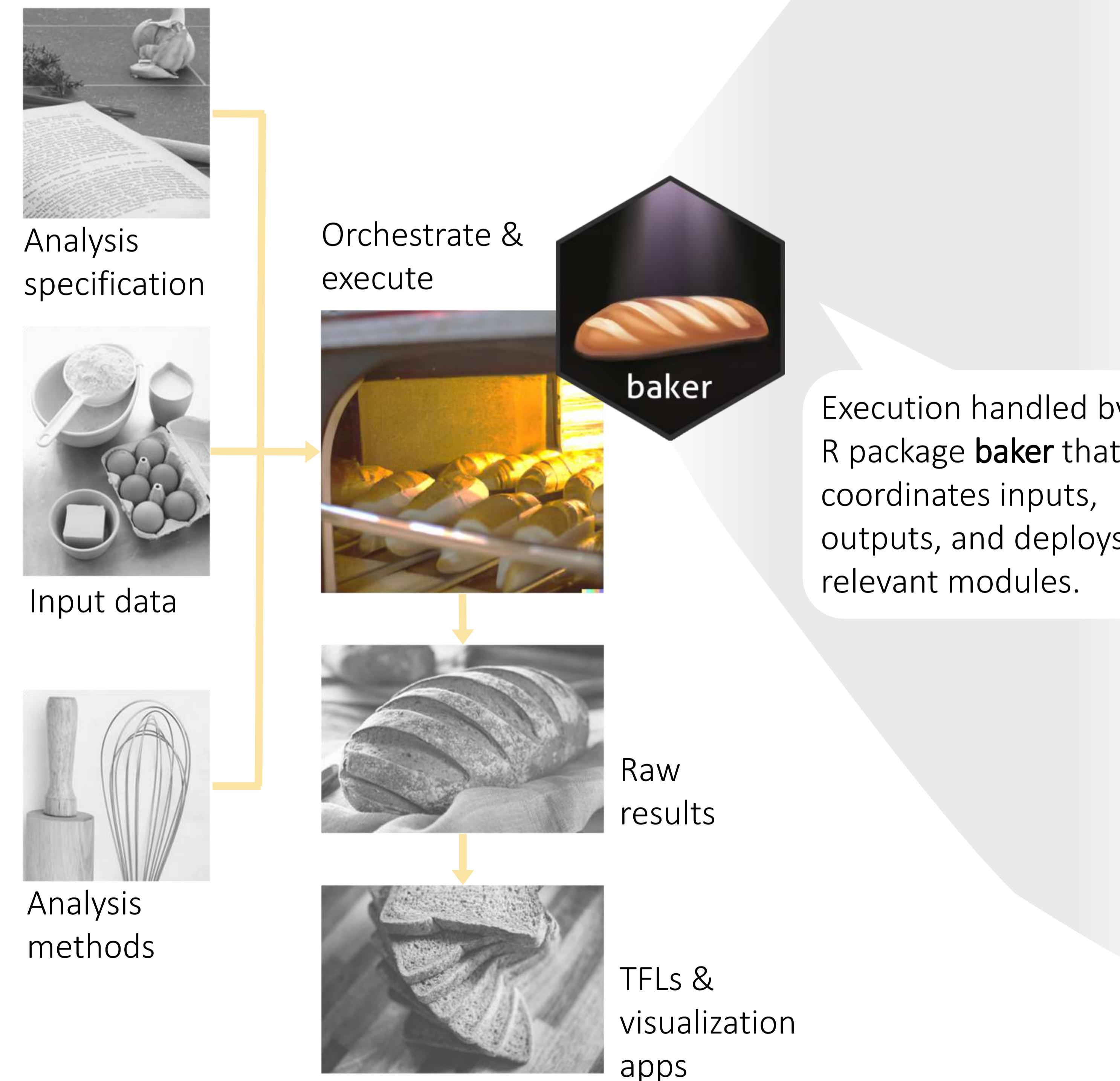
Motivation

Some HTA submissions require thousands of analyses per submission. Producing these is intensive in both programmer and compute time, while often requiring large re-writes of code for each submission despite significant functional overlap.

Solution

R-based modular framework that:

- 1) reduces programmer time with greater code re-use between submissions and increased code quality
- 2) reduces compute time via caching, dependency tracking and parallel computing.



Framework model

Execution stage data model

Implementation modules

