Fake Data Generators

Ruben Jacobo, Waylon Ho

Project Sponsors: Carlos Paradis and Rick Kazman

https://github/sauluh/kaiaulu

ICS 496

Fall 2023

Introduction

What is Kaiaulu?

Open Source Package designed to provide a means for analyzing a variety of software development stacks (git log, mailing list, issue trackers, source code, etc.) Kaiaulu API returns tables, which allows for simple data exploration and integration from various sources.

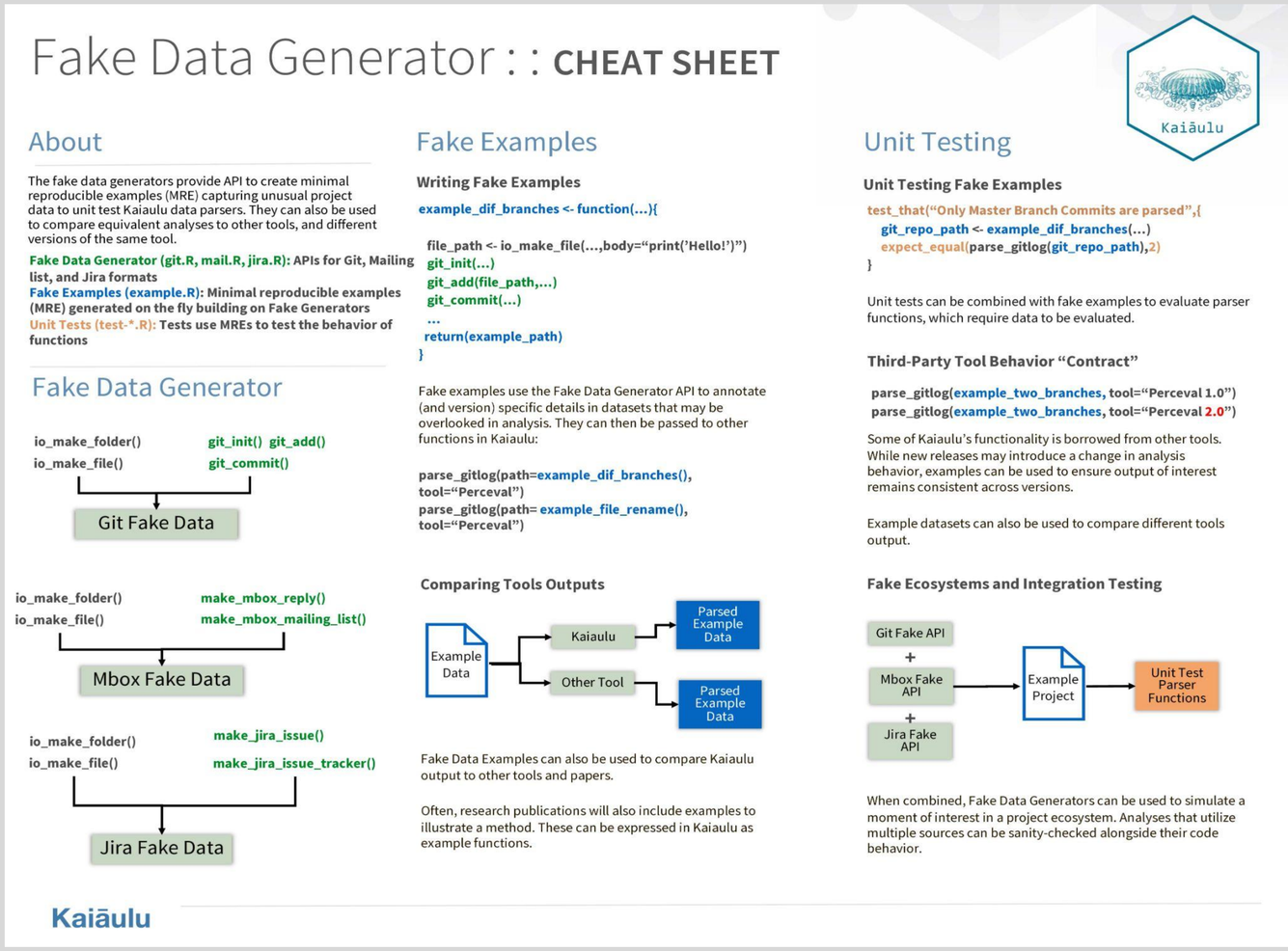
Methodology

Leveraged **Agile** methodologies to:

- Streamline our software development process
- Emphasize close collaboration
- Adaptively plan out new implementation
- Iterative delivery on deliverables

Technology Used

- RStudio IDE and R Language
- Testthat
  - Unit Testing Capabilities
- Perceval
  - Third party data retrieval tool used in conjunction with parser functions



Solution

Tasks Accomplished

- Developed and implemented specialized functions for generating fake data in three critical domains: Git, Mailbox, and Jira files.
- Created realistic yet artificial datasets that mimic real-world scenarios for testing and validation purposes.
- Enhanced the robustness of existing parser functions by subjecting them to a wide array of simulated real-world scenarios.
- Implemented an architecture of Fake Data creation that allows for a.) On-the-fly data generation b.) Examples that capture edge situations in real world dataset and c.) their associated unit tests.
- Crafted a succinct and visually intuitive one-page cheat sheet encapsulating key concepts and usage scenarios of the fake data generators.

Challenges

- Adapting to RStudio and the intricacies of the R language.
- Delved into Kaiaulu's existing codebase, ensuring a comprehensive understanding before making impactful contributions.
- Effectively communicating progress and challenges of weekly deliverables.

Learnings

- Mastered effective communication with our sponsor, translating requests seamlessly into code that aligns with their needs.
- Recognized the critical role of communication with team members and the sponsor/client, pivotal for project success within the specified timeline.

INFORMATION & COMPUTER SCIENCES

UNIVERSITY *of* HAWAIʻI *at* MĀNOA