

Project Management Data Synchronization

Ian Jaymes Iwata, Anthony Lau, Sean Sunoo

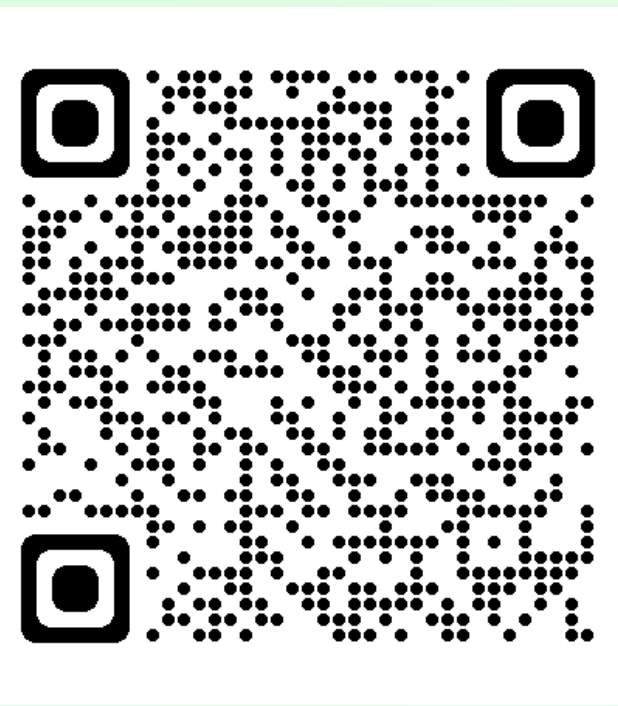
Project Sponsor: Rick Kazman and Carlos Paradis

ICS 496 Capstone Project – Spring 2024

<https://github.com/sailuh/kaiaulu>



INFORMATION & COMPUTER SCIENCES
UNIVERSITY of HAWAII at MĀNOA



What is Kaiāulu?

Kaiāulu is an open-source R package that helps with understanding evolving software development communities, and the artifacts (gitlog, mailing list, files, etc.) which developers collaborate and communicate with. One way it can do this is by downloading data from project manging platforms such as GitHub, JIRA, Bugzilla, etc. for data analysis.

Objective

- Enhance Existing Tools:** Improve functionality of existing downloader and parsers
- Standardize Data Management:** Ensure consistent data organization through uniform file paths and naming conventions among downloaders
- Implement Data Update:** Develop functions to refresh local datasets with new data from project manegement sites
- Schedule Regular Downloads:** Write scripts that utilize Cron jobs to periodically download the latest data

Methodologies

- Agile methodology for flexible and iterative development
- Collaboration through GitHub Issues and Pull Requests
- Bi-weekly developer meetings and weekly sponsor meetings

Roles

- Sean:** Technical/ Testing/ Scribe/ Communication
- Anthony:** Technical/ Testing/ Communication
- Ian:** Technical/ Testing/ Communication

Refresher : : CHEAT SHEET

About

The refresher provide API to synchronize downloader data by rerunning the refresher functions for each respective downloader.

Project Config Setup

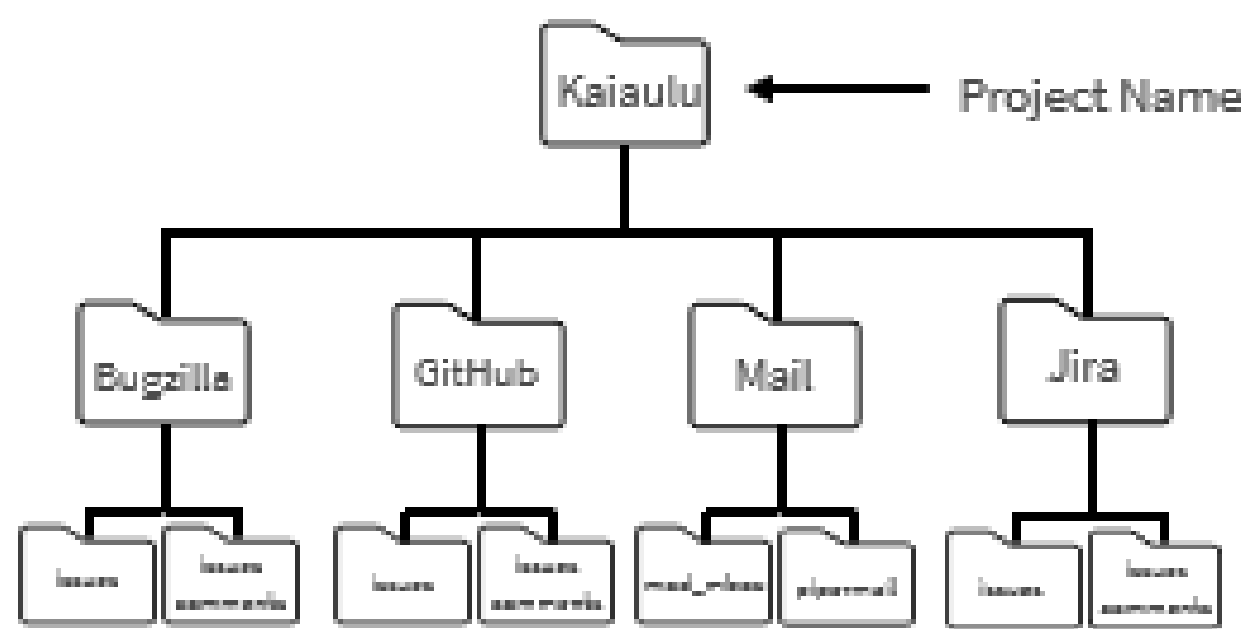
Project configuration files are used to store variables that contain information about a project.

Required Fields

- jira:**
 - project_key_1
 - domain
 - project_key_name
 - issues
 - issues_comments
 - project_key_2
 - domain
 - project_key_name
 - issues
 - issues_comments

The file "tools.yml" must also be configured. See README.md for more information on 3rd party software dependencies.

Folder Organization



Kaiāulu

Naming Convention

All refresh functions will ensure a **naming convention** of the start and **end datetime** contained in the file

```
refresh_jira_issues()  
github_api_project_issue_refresh()  
refresh_bugzilla_issues_comments()  
refresh_mod_mbox()
```



Refresh Issues and Mail

By calling the refresh functions again, against the same folder that already contains data, and which follows the **naming convention**, additional files will be added to the current date.

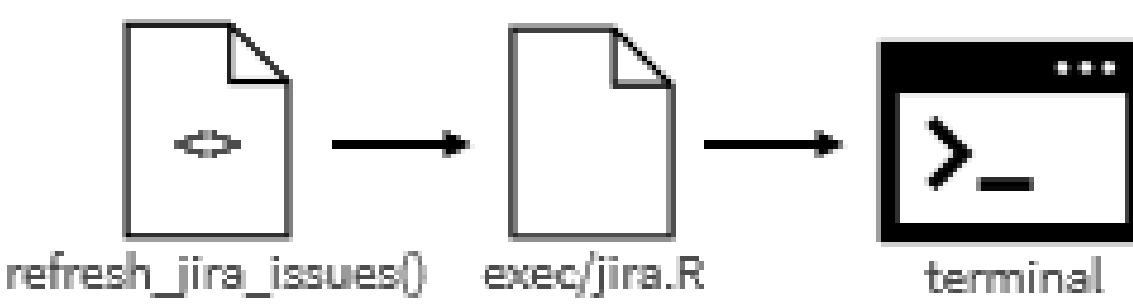
```
refresh_jira_issues(start_datetime,...)  
github_api_project_issue_refresh(start_datetime,...)  
refresh_bugzilla_issues_comments(start_datetime,...)  
refresh_mod_mbox(start_datetime,...)
```



The **highest timestamp** in the suffix of a file name in the data folder is used to identify the file that contains the most recent issue, comment, or email. This is then used as the new highest timestamp for the refresher function.

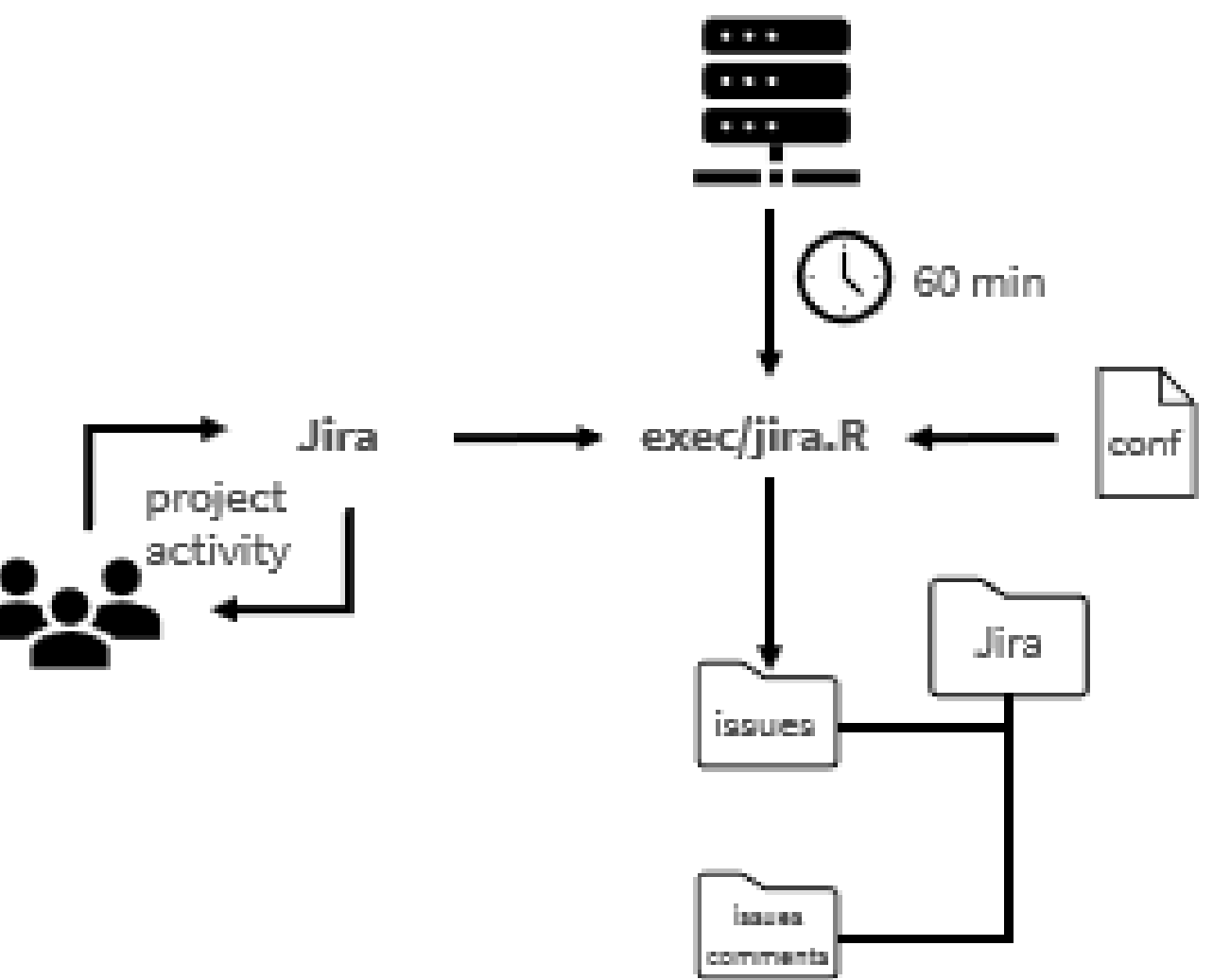
Refresh Script

Executable scripts for refresh functionality callable from the command line allows integration with Cron jobs to automate data synchronization.



Monitoring System

The monitoring system utilizes Cron jobs to call refresh scripts every x minutes (E.g., 60 minutes). The refresher script is called regularly to ensure that local data is always up to date without manual intervention.



Solution

- Reworked file path storage and configuration files
- Parse most recent data from file paths
- Create functions that call various APIs to download data created/ updated only after that date
- Automate downloading of new data via scripts using Cron jobs
- Improve function and notebook documentation

Challenges

- Understanding scope of documentation requirements
- Getting acclimated to coding environment
 - Understanding various functionalities and purposes of Kaiāulu
 - Installing and running 3rd party packages
 - Acquiring proficiency in the R language
- How we overcame: Working closely with the sponsor and helping each other in meetings

Learnings

- Team coordination and communication leading to clearly defined (though evolving) requirements
- Understanding depth and requirements of documentation, keeping in mind users who may lack familiarity withs software
- Writing effective code that fits within the bounds of requirements but is tangible to future contributors

Conclusion

Kaiāulu is an open-source software that will be continuously updated in the future. Our overarching goal was to add refresh functionality to functions that download project data and streamline organization and naming conventions of files that were implemented incongruently by many groups in the past. Future improvements for the project include improving code logic, automating downloaders via scripting and adding more analysis to data already downloaded.