GitHub Takedown Script

Zesheng Xing, Joël Porquet-Lupine

Background

- GitHub is popular space for version control and project showcase
- Students uploaded class materials and homework solutions
- Instructors spend time and efforts to track them and take them down
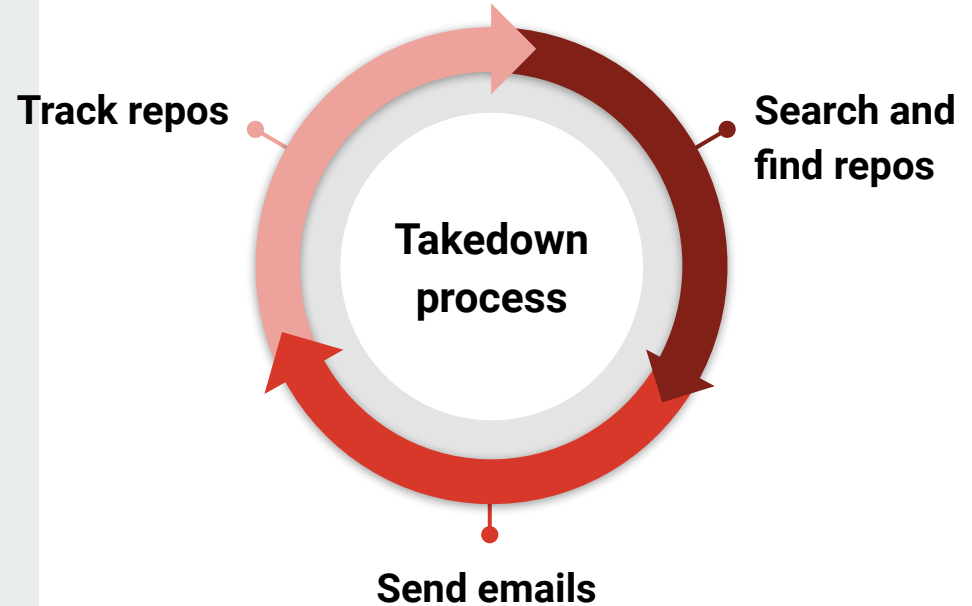- Students get in trouble with SJA

Task

- Find target repositories
- Send emails to repo owners
- Track their status

Design

- Search repos: GitHub APIs
- Send emails: smtp clients
- Track repos: yaml/json files

Track repos

Search and
find repos

Takedown
process

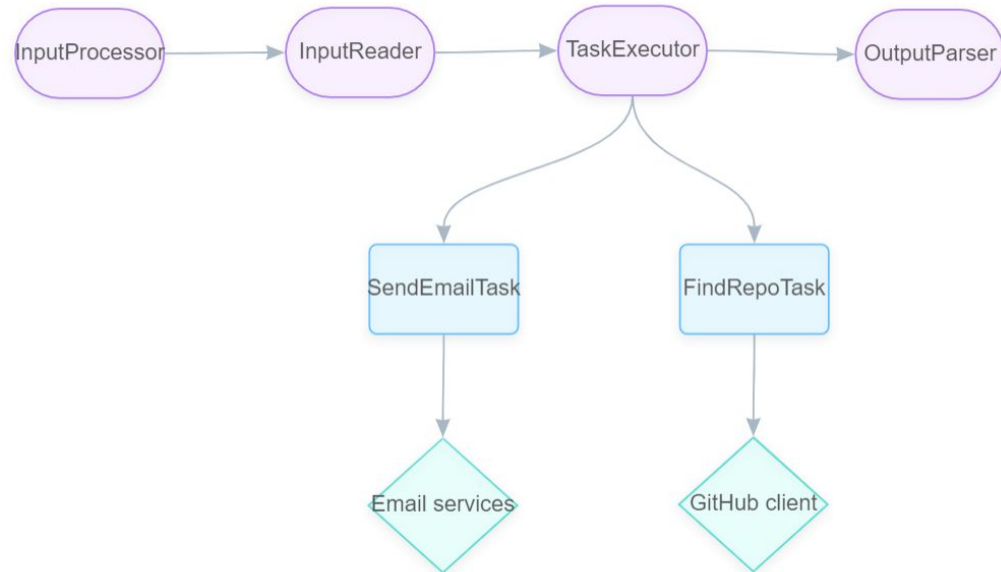Send emails

Package Architecture

Both package and software are provided.

takedown

- client
    - GitHub
- controller
    - InputProcessor
    - InputReader
    - OutputParser
    - TaskExecutor
- task
    - FindRepoTask
    - SendEmailTask

Package Architecture (cont.)

Optimization

- Cache network response
  - reduce requests to individual information
- enable configuration files
  - reduce command line length

=====================================

- Multiple threads

Live Demo

PyPi, GitLab, GitHub

Limitation

- Heavily rely on GitHub APIs
  - rate limit
  - search restrictions
- Not too much options for emails
- Hard to do e2e tests
  - controllers are all tested

Questions, Comments