

GitHub Actions

Juan Manuel Perafan

What are GitHub Actions?

- Automation tool integrated into GitHub
- Enables CI/CD (Continuous Integration/Continuous Deployment)
- Automates workflows based on events (e.g., push, pull request)



- Data pipeline execution
- Data quality checks
- Model training and deployment
- Data Ingestion
- Cleaning stale assets
- Update documentation

GitHub Actions Components



Core Components

- **Workflow:** A set of automated tasks - **defined in YAML
- **Event:** Triggers the workflow (e.g., - **push, schedule, manual)
- **Job:** A set of steps running on a - **single runner
- **Step:** An individual task (e.g., run a - **script, install dependencies)
- **Action:** Reusable step or script (e.- **g., checkout code, run tests)

Getting Started

- Create `.github/workflows/` directory - in your repo
- Add YAML file (e.g., `data-pipeline.- yml`)
- Define triggers, jobs, and steps
- Use existing actions (e.g., `actions/setup-python`)
- Test and monitor via GitHub's Actions