

Software Development in OCR-D

by Paul Pestov & Mareen Geestmann

(Göttingen State and University Library)

—

OCR-D Coordination Project

OCR-D Kick-off Workshop 30.07.2021

Scope and Intention

Our main tasks:

Ensure and increase quality and stability

Ensure the continuity of the software

Our approach:

Consolidate best practices

Write and implement concepts and guidelines

Involve you in the process

Who is our Target Group?

- [OCR-D/core](#)
- Module projects
- Implementation projects
- Any upcoming project in the eco-system

Recap Phase II: Best Practices

- Work on GitHub
- [Apache 2.0 license](#)
- Set up CI/CD
- Write unit tests
- Use [semantic versioning](#)
- Release early, release often

Contributions

- Provide contribution guidelines ([example](#))
- Create a steering committee/team of collaborators?
- Potential activities:
 - Help new users and contributors
 - Provide technical guidance
 - Facilitate discussions
 - Review and merge pull requests

Code Review Checklist

Goal

create common ground between reviewers and contributors

Code Review Checklist

| Tier | Description |
|-------------|--------------------|
| 1 | Project specific |
| 2 | Clean Code |
| 3 | Maintainability |

Tier 1 - Requirements

- Does the feature's functionality meet the (IP's) requirements?
- Are the interfaces implemented according to the [specifications](#)?

Tier 2 - Reusability

YAGNI and DRY

Tier 3 - Documentation

- Try to explain *why* things happen **instead of *what*** happens
- [README](#), [specs](#), [API Reference](#)

Tier 3 - Testing

- Does the feature have **unit tests**, as a minimum?
- Are the tests executed automatically by a CI/CD runner?

The screenshot displays a CI/CD runner interface. At the top, there is a green button with a white checkmark and the text "Success". To its right, the text "test-pull-request" is visible. Below this, a section titled "Jobs" contains a list of four test jobs, each with a green checkmark icon, a job name, and a number:

| | |
|---------------|------|
| test-python39 | 3265 |
| test-python38 | 3267 |
| test-python37 | 3266 |
| test-python36 | 3268 |

Code Review Checklist

Tier 1

Project specific

Tier 2

Clean code

Tier 3

Maintainable

Requirements

Reusability

Documentation

Compatibility

Readability

Testing

Design patterns

Security

Performance

Error handling

Feedback

- Pull request for contribution guidelines in [ocrd-website](#)
- Biweekly Open Tech Call

thank you!