

**nexB/scancode-toolkit**

Project Report  
GSoD 2019

# Reference for Command-Line Options in scancode-toolkit and Reorganize the structure of AboutCode documentation at [aboutcode.readthedocs.io](https://aboutcode.readthedocs.io)

**Ayan Sinha Mahapatra**

## Important Links

My Website	:	<a href="https://ayansinha.dev">https://ayansinha.dev</a>
GitHub Profile	:	<a href="#">AyanSinhaMahapatra</a>
Link to Resume	:	<a href="#">Resume Uploaded at my Website</a> <a href="#">Google Drive</a>
Link to GSoD Proposal	:	<a href="#">Proposal Uploaded at my Website</a>
Link to GSoD Report (Backups)	:	<a href="#">Pull Request</a> <a href="#">Google Drive</a>
Link to Documentation Websites	:	<a href="#">nexB/aboutcode</a> <a href="#">nexB/scancode-toolkit</a>
Link to Blog Post about GSoD	:	<a href="#">Blog Post For GSoD</a>

## Contributions

Documentation development fro GSoD followed this plan generally -

1. Copy Wikis for all [nexB](#) Projects to [aboutcode](#), by converting them to reStructuredText and host them at [aboutcode.readthedocs.io](#).
2. Add comprehensive documentation tests to Continuous Integration @ Travis-CI so documentation builds are error-free and link/style checks are enforced.
3. Prepare and discuss the GSoD project with mentors. Change plans to accommodate feedback throughout the GSoD timeline.
4. Add basic structure of the documentation, restructure already existing into appropriate sections and improve upon them wherever necessary.
5. Add Command Line Options Reference for the develop branch, latest, to be added to the next release.
6. Add some new Tutorials/How-Tos and restructure/improve existing ones [\[WIP\]](#).
7. Add support for first-timers, good first issues, initial setup [\[WIP\]](#).
8. Add support for Plugins and adding plugins via a cookiecutter plugin [\[WIP\]](#).

## Contributions before GSoD Started

### PRs @ nexB/aboutcode

- [\[#17\] Adds Travis-CI for sphinx-build and linkcheck](#)

This adds test scripts to check the build status of Sphinx Documentation along with adding checks for broken links. It also fixes some broken links and updates doc\_maintenance.rst

- [\[#15\] Adds Documentation from Wiki](#)

To the main documentation, this adds Deltacode, Aboutcode Docs and remaining pages of ScanCode Toolkit and ScanCode Workbench from the wiki.

- [\[#13\] Adds GitHub wiki of Scancode-Toolkit and Scancode-Workbench](#)

This adds almost all the Wiki Pages of Scancode-Toolkit (Except the How-To pages) and Scancode-Workbench.

## Contributions in the GSoD Period

### PRs @ nexB/scancode-toolkit

- [\[#1784\] Add Travis CI Test For Docs](#)

This PR adds Travis-CI checks for documentation, similar to that in aboutcode. It also adds the requirements file for ReadTheDocs build and minor doc style fixes.

- [\[#1827\] Add New Issue and Pull Request Templates](#)

Fixes [#1814](#) [Old issue Template]. Adds new Issue Templates and Pull Request Templates with config file for the GitHub welcome app.

- [\[#1812\] Add Documentation from Aboutcode](#)

This PR Migrates scancode-toolkit documentation from aboutcode after adding the final improvements/changes via [nexB/aboutcode#33](#). **This is the most important PR as it adds a major portion of the work done to scancode-toolkit from aboutcode in one PR.**

- [\[WIP\] Add Cookiecutter Plugin and Plugin Docs](#)

This will add docs related to the existing plugin system and how to add plugins easily (by using the cookiecutter project) to extend scan capabilities. Will also add cookiecutter related code necessary to create new plugins.

- [\[WIP\] \[#1847\] Fix remaining documentation Issues](#)

Fixes multiple issues listed under the tag "[Documentation](#)".

## PRs @ nexB/aboutcode

- [\[#26\] Adds Documentation Versioning, Command Line Options, Changelog](#)

This adds to the existing documentation a versioning system that is compatible with ReadTheDocs hosting, adds command-line options and updates the documentation for scancode-toolkit so that the documentation is consistent for all the scancode-toolkit versions.

- [\[#27\] CLI Options Reference for v3.1.1](#)

This PR adds to Scancode-Toolkit documentation the entire section of Command Line Interface Options reference pages. Also adds .RST snippets and requirements file for ReadTheDocs build. Some discussions on docs took place here in this thread.

- [\[#29\] Update Docs with changes in Wiki](#)

This commit updates the docs with changes that were made to wikis (Including GSoC 2019 Reports, Roadmap and new How-To's). Also adds more help on printing scan result to stdout.

- [\[#30\] Add doc8 style checks to Travis-CI](#)

This adds to the existing Travis-CI tests, doc8, to enforce style-related issues. For more information refer - [What Is Checked](#)

- [\[#31\] Fix template error](#)

This PR fixes [nexB/scancode-toolkit#1760](#) by updating template.html, which was outdated and using --custom-output threw errors.

- [\[#33\] Add final Changes/Improvements before Migrating](#)

This PR essentially solves all issues mentioned in [#32](#) and adds other final changes before migration. Main

1. Getting Started Section
2. Support for First-Timers
3. Document support For pip install
4. Document python 3 related installations and others
5. Set Up Intersphinx for both Aboutcode and Scancode Docs

- [\[#34\] Delete scancode-toolkit documentation](#)

This PR deletes scancode-toolkit docs as they were migrated via this Pull Request [nexB/scancode-toolkit#1812](#). This also modifies the landing page for scancode-toolkit.readthedocs.io and adds issue/pull-request templates.

- [\[#35\] Add GSoD 2019 Report](#)

This adds a short GSoD report for hosting at aboutcode.readthedocs.io as other GSoC/GSoD reports are also hosted here.

## Long-Running Issues Created

These issues are created as Long-Running Issue Trackers, Roadmaps, and Help Issues to First-Timers so they have a better doc experience and documentation development is carried out smoothly after this GSoD project.

- [\[#1837\] \[Setup\] Commonly Encountered Errors while Setting Up](#)

An issue to help users who are setting up scancode-toolkit for the first time, to debug common errors.

- [\[#1825\] All Good First Issues List](#)

A good first issue means it's recommended for people who haven't contributed to our codebase before but can code in python. This issue lists all such issues to be assigned to interested contributors by the maintainers and links to more detailed documentation guides.

- [\[#1826\] First Timers Only Issues List](#)

A first-timers only issue means it's recommended for people who haven't contributed to our codebase before. This issue lists all such issues to be assigned to first-timers by the maintainers and links to more detailed documentation guides.

- [\[#1813\] Documentation Inconsistencies Tracker](#)

This Issue is to Track all scancode-toolkit Documentation Inconsistencies in One Place. Therefore these are also easy issues for first-timers only.

- [\[#1824\] ScanCode Toolkit Documentation Roadmap](#)

Major documentation sections/features to be added to the scancode-toolkit documentation

## Issues Solved

Minor issues I've solved in scancode-toolkit

- [\[#32\] Scancode-Toolkit Doc Improvements](#)

Multiple tasks were tracked in this issue in terms of improving the Scancode-Toolkit Docs, which were eventually added through [this PR \[#33\]](#).

- [\[#1814\] Old Issue Template](#)

Scancode-toolkit was using an old issue-template and this [PR \[#1827\]](#) solves this issue.

- [\[#1760\] Error in the wiki template file for `--custom-template`](#)

The template in custom output documentation was outdated and hence was giving an error, and this issue was solved by this [PR \[#31\]](#).

## Issues Contributed

Bugs I've reported in scancode-toolkit.

- [\[#1778\] "pip install scancode-toolkit" gives an error](#)
- [\[#1759\] Error in `summary-by-facet` Plugin](#)
- [\[#1758\] Error in `filter-clues` Plugin](#)
- [\[#1790\] Travis-CI jobs on queue for a long time](#)

# Project in Details

## Main Features of this GSoD Project

1. **Restructured Documentation hosted at ReadTheDocs:** Outdated Documentation from GitHub Wikis was updated, improved and converted to .RST, to be hosted at `aboutcode.readthedocs.io`.
2. **Documentation Tests and Roadmaps:** Documentation tests were added to check Documentation Builds, Internal/External Links, and Style Standards. These tests were added in both `aboutcode` and `scancode-toolkit` and already existing/new documentation was made compliant of these tests.
3. **Command Line Options Reference:** The entire Command Line Options were listed, explained, and their application elaborated from scratch. This is the main goal and the most important addition to the `scancode-toolkit` documentation.
4. **Migration of Scancode-Toolkit docs into its own Repository:** After adding the `scancode toolkit` docs to `aboutcode` and reviews, this section was migrated to its own repository, to be distributed with `scancode-toolkit` and with local builds enabled.
5. **[WIP] Improve/Add to Existing Tutorials/How-Tos:** Added two major Tutorials and other Tutorials/How-Tos were also updated to be consistent with the latest `scancode` versions and more Tutorials/How-Tos can be added later to better help the users.
6. **[WIP] Add Plugin Support:** A cookiecutter based plugin creation system has to be added, and relevant docs for this process and for the plugin architecture has to be added.
7. **[WIP] Solve Other Documentation Issues:** Solving multiple issues listed under the tag "[Documentation](#)", through [this PR \[#1847\]](#).



## Deviation from the [Original Project Proposal](#)

### Additions

- Comprehensive Plugin Support through a cookiecutter plugin and documentation help.
- Getting Started Sections and other help for beginners.
- Elaborate the JSON Model [Not a part of the GSoD Project]

### Omissions

- The “**explanation**” section. (How Scancode Works)

## Work Remaining

This is work that is in progress currently and will be completed (by me) after the official GSoD period ends.

### 1. Improve/Add to Existing Tutorials/How-Tos

Already added two major Tutorials on How To Scan and How to Visualize Scan Results, as these are the most important to the new users. Other Tutorials/How-Tos were also updated to be consistent with the latest scancode versions and restructured as required. Now more Tutorials/How-Tos can be added later to better help the users.

### 2. Add Plugin Support

Plugins can help extend scan capabilities for different user requirements and currently this is a complex process, without any comprehensive documentation (only this [tutorial](#) was added). A cookiecutter based plugin creation system has to be added, and relevant docs for this process and for the plugin architecture has to be added.

### 3. Solve Other Documentation Issues

Solving multiple issues listed under the tag "[Documentation](#)", through [this PR \[#1847\]](#).

# About my time in GSoD @ aboutcode

## Challenges

- Applying for GSoD and preparing the Proposal.
- Learning more about Git and OpenSource Project Workflows.
- Working with the Developers to make sure the Documentation is complete, easy to access, easy to understand.
- Making sure my plans are in line with what the developers/maintainers have in their plan.
- Understanding Code from Scancode-Toolkit.
- Testing Scancode-toolkit for different scans and adding the documentation requirements with scancode-toolkit [Work Remaining].

## Why I'm Thankful

- Support from my Mentors, and from everyone in the Organization
- This Opportunity to contribute to Open Source Software
- This Opportunity to learn more about Documentation/User Experience and Creating Documentation for different requirements/users.
- This Opportunity to learn more about OpenSource Project Workflows and glimpse into Contributing/Maintaining Code.
- This opportunity to learn more about Code Development, Continuous Integration, Testing, and Issue Tracking.

## Acknowledgments

I would like to thank [Dennis Clark](#), [Philippe Ombredanne](#) and [Steven Esser](#) for their help throughout my GSoD project, from helping me prepare my proposal and project to constructive criticism, nit-picky reviews, conference calls to discuss project roadmaps and to explain to me when I was stuck in any problem. Their support and positivity helped me through this project and made me enjoy every bit of it.