
Attentive Migration of wiki activity pages to git

GOOGLE SUMMER OF CODE 2018 — SUGAR LABS

RUDRA SADHU

Contents

About Me	2
About Project	4
Description	4
Objective	4
Promised Outcome	5
Technologies Used	5
High-Level overview	5
Detailed Technical description	6
Timeline	8
Deliverables	9
Convince us	9
Project and the Community	10
Miscellaneous	12

About Me

- **Name:** Rudra Sadhu
- **Email:** rdrsadhu@gmail.com
- **Sugar Labs wiki username:** rdrsadhu(userpage at [User:rdrsadhu](#))
- **IRC nickname on irc.freenode.net:** rdrsadhu
- **Where am I located? and what hours(UTC) do I tend to work?**
 - I'm located at Kolkata, India (UTC +05:30)
 - Working Hours: Flexible (Whatever works for both me and my mentor)
- **Past experiences with Open-Source projects:**

actively contributing to Open-Source since August, 2017
organisation: [Oppia](#)
firstly, Why Oppia(and now Sugar Labs)?

 - I've always believed, *'there is no process as powerful as education to change the world.'*
 - Mission of Oppia Foundation and Sugar Labs are similar(if not exactly same). Read [here](#).

My first pull request: [Fix #3285: Add share-to-classroom Button](#)
~Happiness is when pull-requests gets merged :)

What else did I chip in?

- while playing around, Found an issue in the application,
Recorded it : <https://github.com/oppia/oppia/issues/3750>
Submitted fix: <https://github.com/oppia/oppia/issues/3750>
~Felt Awesome!
- Created outline for Mathematics Lessons(Geometry)
- 6 other Merged pull-requests. Complete list [here](#)
these are either Code Refactoring (which proves my skill to understand complex codebases, and knowledge of intermediate git workflow- dealing with branches, merge-conflicts etc)
or Trivial Documentation fixes (which confirms my attention to detail)

What did I gain out of contributing to Open Source projects and why I'm again up for it?

- I learned a lot! and more importantly, had fun learning.
- A sense of pride and satisfaction, to be able to contribute for a good cause, which impacts people around the globe.
- Mentions on the [Credits Page](#). Even if I die tomorrow, this will stay on.
- I've got nothing to lose :D

I've tasted failures as well, and I'm not uncomfortable to talk about them. Well, because they often teach us a lot more than successes.

[this](#) pull-request of mine did not get merged.

About Project

- **Name:** Attentive Migration of wiki activity pages to git
- **Description:**
 - At present, the [Activities#Sugar_Activities](#) section in Sugar Labs wiki lists 345 pages.
 - These pages individually serves as kind of homepages(containing all relevant information), for different Sugar Activities.
example:
[Activities/MakeyMakey](#) is the homepage for [MakeyMakey](#) activity.
Whereas, its source-code is hosted at github.com/sugarlabs/makeymakey
 - as Sugar Labs is moving towards GitHub style of development (see [discussion](#))
for any change to a activity,
it gets cumbersome for the developer to update both the GitHub repository and its corresponding wiki-page documentation

thus, it would be beneficial(and more maintainable) in the long run if these 345 wiki pages were embedded only in their corresponding GitHub repositories.
- **Objective:**

Write a program(or script) to migrate all of the above-mentioned wiki-content to their appropriate GitHub repositories in Markdown(or reStructuredText) format, with detailed attention to quality.

- **Promised Outcome:**

100% of the mentioned wiki-pages(from wikitext markup) moved to their GitHub repositories(in Markdown/reStructuredText format).

- Something like github.com/rdrsadhu/portfolio-activity, from [Activities/Portfolio](#)
 - * this is a fork of the original repository [sugarlabs/portfolio-activity](https://github.com/sugarlabs/portfolio-activity)
 - * I'm aware that the converted Markdown syntax of README.md file is not optimal.
 - This was done just to serve as a quick prototype.

Also, all the current wiki-pages will be deprecated(deleted/moved as desired by the community)

- **Technologies Used:**

- Programming Languages: Python 3.6
- Tools: [git](#), [pandoc](#)
- APIs: [MediaWiki API](#)

- **High-Level overview of Migration Process:**

For each page to migrate

1. Determine to which GitHub repository, it'll be added
2. from wiki.sugarlabs.org/activities/{activity_page}, Get source file(wikitext) which is to be converted(to Markdown/reStructuredText)
3. Get associated extra files(images etc) used in the wiki-page
4. Convert wikitext content to Markdown/reStructuredText format
5. Add converted files to GitHub repository mentioned in Step 1
6. Update all other pages in the wiki, which links to this to-be-deprecated page, with the new GitHub link
7. Deprecate wiki-page and associated extra files from wiki.sugarlabs.org

- **Detailed Technical description of Migration Process:**

1. Query wiki.sugarlabs.org to get list of all pages to convert
2. For each page, research and find out(or discuss as needed with mentor/community) where to place them in GitHub

We have three audiences for documentation,

- a) users; such as children who are learners, and their teachers
- b) integrators; who join many activities into a system,
- c) developers; who develops Sugar and maintains its activities.

Activity pages does not only contains user documentation, but also developer documentation. Some activity pages have both, some have only one or the other.

example:

1. [Activities/TamTam](#) is developer documentation, and should be moved to [tamtam](#)

2. [Activities/Story](#) is user documentation, and should be moved to [help-activity](#)

3. [Activities/Physics](#) is both,
Some decision will be needed on a case to case basis to deal with these pages.

3. For each GitHub repository to push converted files:
 - (a) Fork the repository
 - (b) Pull the forked repository
 - (c) Checkout a new branch for changes
 - (d) For each file to be pushed into the repository:
 - i. Using MediaWiki API, query a specific wiki-page, to get its raw source text(in wikitext format), and all associated files(images etc) used
 - ii. create a safe backup of the original raw wikitext
 - iii. generate one temporary duplicate file(containing raw wikitext), for conversion

iv. using pandoc, convert

to Markdown or reStructuredText?

- for content that is to be in the corresponding source repository in GitHub, use .md
- for user documentation content that is to be in the Help Activity in Sugar, use .rst

v. while the conversion is not perfect (it is difficult or sometimes not-possible, to generate the exact functionalities of a wiki-page into a markdown file. For example, image-galleries)

- modify the temporary wikitext file, with attention to every small detail
 - and convert again
- (e) Test all changes
- (f) Commit all newly generated files
- (g) git push to Forked Repository
- (h) Send a Pull Request to original Repository
- (i) After Pull Request gets merged;
- Update all other pages in the wiki, which links to these to-be-deprecated pages, with the new GitHub link
 - delete contents of original wiki-page, by changing it to a redirect. This way, the page history will be preserved, but we won't be maintaining redundant versions of content.

Important:

- Adding documentation to the activity repositories, may sometimes unnecessarily enlarge the activity bundle size;
in such a scenario, a custom setup.py will be written which excludes the documentation folder.

• Project Timeline:

I will work 45-50 hours a week, and have decided the goals accordingly.

- April 23 - May 13: Community Bonding Period
 - * Research about the variety of content in different wiki-pages
 - * Learn more about syntax of wikitext, different functionalities of MediaWiki API
 - * Learn how to write a custom setup.py to keep the size of activity bundle in check
- May 14 - June 10: Coding Month 01
 - * Week 01: Write and Test Script extensively
 - * Week 02: Complete end-to-end migration of 30 Pages
 - * Week 03: Complete end-to-end migration of 65 Pages
 - * Week 04: Complete end-to-end migration of 100 Pages
- June 11 - July 08: Coding Month 02
 - * Week 05: Complete end-to-end migration of 150 Pages
 - * Week 06: Complete end-to-end migration of 200 Pages
 - * Week 07: Complete end-to-end migration of 250 Pages
 - * Week 08: Complete end-to-end migration of 300 Pages
- July 09 - August 05: Coding Month 03
 - * Week 09: Complete end-to-end migration of 345 Pages
 - * Week 10: Detailed Quality Check of all migrated pages
 - * Week 11: Extensive documentation about the complete project.
 - * Week 12: Buffer time to accommodate any due/extra/unknown work

- **Project Deliverables:**

3 clear quantitatively measurable goals to determine project success/failure status which will help in evaluation

- Month 01: Compete high-quality end-to-end migration of 100 pages
- Month 02: Compete high-quality end-to-end migration of 300 pages
- Month 03: Compete high-quality end-to-end migration of 345 pages

- **Convince us, in 5-15 sentences, that you will be able to successfully complete your project in the timeline you have described. Link to prior work or other resources as relevant.**

what better way to prove that I can do the work, than doing some 'real' work?

Find the beta version of proposed script [here](#)

I implemented some basic features of the script, which given any Sugar Activity {pagename} and {forked_repo_url}, migrates the page(without any manual supervision) and commits all generated files into the local copy of forked-repository

- The script is commented to explain what each part is doing
- The script follows, roughly the same workflow stated in Detailed Technical description of Migration Process above

Find few of the migrated pages at [godiard/help-activity/pull/38](#)

I do not have any other commitments during the summer, and will be able to devote complete time & energy to implement this project at the best of its quality.

Moreover, I'm excited to learn and willing to push myself.

Project and the Community

- If the project is successfully completed, what will be its impact on the Sugar Labs community?

– Answer 1: Rudra Sadhu:

- * All information about 1 particular Sugar Activity, will live inside its GitHub Repository, which will help
 - Activity Maintainers
 - New members of the community
 - Users
 - and ultimately everyone :)
- * GitHub is the world's leading software development platform, (check [Stack Overflow 2018 developer survey](#))
 - thus hosting a complete repository with all the documentation, will provide greater exposure and surely help to attract more developers and traffic to the community

– Answer 2: James Cameron(quozl@laptop.org):

”

Originally documentation was separate because we had non-coding developers and tool chains that varied by type of developer. Now we use GitHub the tool chains are combined.

With the project as described, documentation will be concentrated in the source code repository for an activity, reducing ongoing maintenance.

We have less active Wiki contributors than we ever did, and in the current threat environment a Wiki requires significant monitoring and administration;
we recently lost some system administrators and gained new ones;
using GitHub allows us to outsource system administration.”

- Answer 3: Vipul Gupta(vipulgupta2048@gmail.com)

”

The world is changing, technology at present will become obsolete at one point of time. The decision of Sugar Labs to keep with the times, advance, grow, and get new developers associated by migrating its documentation of activities is critical for future opportunities and will help everyone involved.”

- **What will I do if I get stuck on the project and my mentor isn't around?**

- I know where to get information, or ask for help (google, stackoverflow, etc)
- Read documentation of related tools
- Ask other Sugar Labs members via IRC, developer mailing lists
- Ask help from Sugar Labs GSoC coordinator

- **How do I propose to keep the community informed of my progress and any problems or questions I might have over the course of the project?**

- daily updates on IRC
- weekly blog post
 - * about what I did the previous week
 - * what I plan to do the next week
- communication via mailing-lists (as and when needed)
- all the code I write and every Pull-Requests/Issues I open
 - * will be thoroughly commented in an easy to understand format
- Live Record of Activity migration status for all the 345 pages.
 - * I intend to create a wiki-page/git-repo which at any time will reflect the current migration status of all the 345 pages. Something like [this](#).

Miscellaneous

- **Send us a link to a pull request or merge request you have made on a Sugar or Sugar activity bug.**

I went ahead and migrated a few pages out of the 345 listed. Find the pull request at <https://github.com/godiard/help-activity/pull/38>

- **Describe a great learning experience you had as a child.**

Learning how to ride a bicycle surely has been one great experience. Though my father initially helped to ease out the fear; just some weeks later I was alone all by myself, deciding when to speed and how to stop, all without any manual. :D
I did fall a few(read a lot of) times too.

What came out of the process to me, is one very important life lesson:
to keep balance, we must look forward and continue moving.

- **Is there anything else we should have asked you or anything else that we should know that might make us like you or your project more?**

Yes,

- I did learn a lot in the last 6 weeks.
From wikitext markup to Mediawiki API, pandoc, sphinx, reStructured-Text and a bunch of other stuff.
I had not even heard about most of these technologies before coming across this project.

So, *thank you* for the opportunity :)

- I’ve always tried to prove my worth by action.
Meanwhile, I’ll continue to migrate more pages, irrespective of my acceptance into the GSoC program.
I expect my work to be my loudest advocate.
- Personally, the field of education is very close to my heart.
I help out local middle school students to learn basic Computer Science concepts using simple flowcharts and other intuitive stuffs.

and thus for the common goal, I believe I’ll continue to work with the Sugar Labs organization and you all in the coming future.

Drafted with ♥ by another Open Source enthusiast, just like you :)
The proposal is also available at the [Sugar Labs wiki](#).