Internet Archive | Open Library

Google Summer of Code 2019

Project Proposal - Building the world's Open Library, together

---

About Me

Personal Information

Name	Shreyansh Jain
Nick	shreyansh23
Email	jain.shreyansh0@gmail.com
Telephone	+91 9131422751
Location	IIT Roorkee, Uttarakhand, India
Time Zone	India (UTC + 5.30)
GitHub	shreyansh23
Typical working hours	UTC 1730 - 0530 / PT 1030 - 2230 / IST 2300 - 1100

University Information

University	Indian Institute of Technology, Roorkee
Major	Electrical Engineering
Current Year	2nd year (2021 expected graduation)
Degree	Bachelor of Technology (B.Tech)

Personal Background

I am a sophomore currently enrolled in Electrical engineering at IIT Roorkee. I have developed a passion for programming and web development in my freshman year and since then, most of my time goes into reading and writing software and contributing to open-source - looking over repositories of the products I use or come across, trying to contribute back to the organisations whose services have been an asset to me. This even includes the web applications that my university undergraduates use. These services are developed and maintained by an open-source organisation IMGIITRoorkee, of which I am a member. I have contributed and built web applications like SHP, FacApp as part of it.

I have the experience of working closely with a team as I am an active member of Information Management Group at IIT Roorkee, a bunch of enthusiasts who manage the institute main website, internet and intranet activities of the university.

I have been an active part of the **Internet Archive** community. I try solving issues on GitHub to get a better understanding of the codebase and the workflow of the organization. In this journey, I learn quite a lot about the projects under the organization. I feel great in interacting with the community. This boosts my confidence and motivates me towards contributing to open-source.

This is the first time I am contributing to an open-source project on such a substantial scale.

Development Skills

- Fluent in Python, Javascript and C/C++
- Proficient in frontend libraries like React and Redux
- Worked with databases like MySQL and Postgres
- Worked on developing RESTful APIs
- Experience in Django and Flask
- Experience in working with Bootstrap and Semantic UI CSS
- Working knowledge of PHP, jQuery, Less, Jest
- Experienced with Docker
- Proficient with Git
- Worked on design tools like Figma and Adobe Illustrator

Projects

Student Home Page (SHP)

- **SHP** is an internal application for the students of IIT Roorkee. It serves as the student profile application for all the (8k+ active) students of IIT Roorkee.
- The app is completely generalised, meaning the entire front end is generated from a JSON object and can be moulded easily to suit the needs of a different university.
- Every Student gets his profile page visible on the iitr.ac.in domain by a self-chosen handle.
- It serves to collect all necessary data that a student requires to show on his home page.
- For the backend, Django and Django REST was used. For the frontend, React was used and styling was done using Semantic UI.

Faculty Profile (FacApp)

- **FacApp** is similar to SHP but it is restricted only to faculty members of IIT Roorkee.
- It collects more of professional data, like a faculty member's education and achievements.
- Like SHP, FacApp is also fully generalised and flexible to modifications for other universities.
- The page generated is visible on the IIT Roorkee website as a faculty member's main page.

Music-Stream-App

- **Stream** is a web app to stream music on a PC, while one can control the song using other PCs connected in a network.
- It is a React-Django application with webpack used as the module bundler for JavaScript code.
- For real-time communication between the Stream and other users, WebSocket connections are make using Django channels.
- The application code can be found here.

College Radio (Wavify)

- **Wavify** is an upgraded version of the Music-Stream-App. It has improved features like support for multiple independent channels like jazz, pop, rock.
- This application was developed in the 48 hours hackathon, *Syntax Error* organized by a campus tech group.
- The code can be found here: backend, frontend.

Motivation

Google Summer of Code is a nice platform to get acquainted with the open-source community and their skilful mentors. It gives professional work experience in a student's university years where one collaboratively builds a product for the welfare of the society. In this process, both the individual and the community improves and grows. The prospect of developing software actually capable of making a difference means a lot to me and I would love to give back to the community in some way.

I had used services of **Internet Archive** before. This includes the media from archive.org, books borrowed from openlibrary.org and going to past using the Wayback Machine. Earlier I had no idea that Internet Archive is the organization behind all these. I first got to know about Internet Archive as an open-source organization when I saw the past years' selected organizations. The projects under the Internet Archive had a familiar technology stack. I decided to contribute to it as a starting point in open-source. Since then I have been contributing to the Internet Archive and my journey so far has been a great one. Every feedback on my PRs (to Open Library or IAUX) and on the slack community channel fills me with motivation. I am fortunate to have met this vibrant community of so many motivated developers and creators. My contributions to **Internet Archive** can be seen here.

I chose this **Open Library** project because it is very well aligned with my interests and also correlated to what I have worked on in the past. Other than that, this project provides a good opportunity to apply my learnings on a practical scale. This is my field of interest and therefore the natural inclination to this project. Moreover, most of the code in Open Library was in python and javascript (both of them are in my favourites). This motivates me to spend my summer working on Open Library.

Pre-GSoC involvement

I have been an active contributor to open-source projects in the past, prior to GSoC. My contributions have helped me gain experience in understanding the flow of any pre-written code at a rapid pace and enabled me to edit and add new features. Contribution to the projects which affects a large community is something which gives an endless source of encouragement to me. Some of my contributions to open-source have been enumerated below:

Contributions to **Open Library | Internet Archive**

I've made a few pull requests to Open Library and gotten most of them merged. I've mostly worked towards improving UI, fixing js/jQuery bugs, and documentation of the project to get an idea of the codebase. Also while exploring Open Library, I reported the bugs or features requests in the form of issues. All of my contributions can be viewed here. Here is a list of my contributions towards Open Library:

- **PR: #1867 (merged):** Update occurrences of .live() method in /templates and /plugins
- **PR: #1931 (merged):** Correct faulty CSS in edit form and make it completely responsive
- **PR: #1940 (merged):** Always enable librarian mode and remove unused code
- **PR: #1941 (merged):** Fix collapsing of read and return button
- **PR: #1942 (merged):** Fix position of Y-Axis label and make graph responsive on the mobile device
- **PR: #1947 (merged):** Add a linting rule for consistent string quotes
- **PR: #1973 (merged):** Fix CSS in cover image modal
- **PR: #1980 (merged):** Fixes js problem of multiple form upload.
- **PR: #1981 (merged)**: Add separate issue template for bug report and feature request
- **PR: #1993 (merged)**: Make facebook and twitter card preview uniform by using only open graph meta tags
- **PR: #1996 (merged)**: Fix bug with occurring on creating a new list.

- **Issue: #1949:** Clean UI: Make the table in loan page look neater
- **Issue: #1970:** Create a new list button not working
- **Issue: #1971:** UI issue in cover image and author image upload modal
- **Issue: #1972:** Multiple UI issues on editions page
- **Issue: #1974:** Separate issue template for bug report and feature request

Contributions to **IAUX | Internet Archive**

Soon as I got to know about IAUX project, I started contributing to it. In the initial phase, my contributions mainly focused on making the project installation more simple and bug-free for the community.
I've made a few pull requests to the project and some of them got merged. Others are in phase improvement and review. Here is a list of my contributions to the project so far:

- **PR: #18 (merged):** Add separate issue template for bug report and feature request
- **PR: #20 (merged):** Add pull request template
- **PR: #22 (merged):** Fix Linux installation issue due to fsevents dependency
- **PR: #23 (open):** Correct test:generate-output command in package.json
- **PR: #51 (open):** Migrate to Babel 7
- **PR: #66 (open):** Fix commands for running ia-prototypes

- **Issues: #16, #21, #41, #42, #43, #53, #65, #68**

- **Documentation:** Helped in documenting the components of Archive.org to the design template library wiki of IAUX.

Contributions to **IMGIITRoorkee**

IMG maintains the official intranet portal, the homepage and the entire technological backbone of IIT Roorkee. Apart from IIT Roorkee related work, IMG targets the developer community of the world by its remarkable project Omniport under the open-source organisation IMGIITRoorkee. Omniport seek to create one true portal for any and every educational institute in the world, catering all its needs. I spent my autumn semester of the second year working on the two of its major applications, SHP and FacApp.

Student Home Page (SHP): SHP is the most generalised student profile app for any institute. I worked on building the application from scratch. See more

Faculty Profile (FacApp): FacApp is the most generalised faculty profile app for any institute. I worked on building the application from scratch. See more

Omniport backend**:** The Django backend of the one true portal for any and every educational institute. Working on SHP and FacApp was possible due to this project. Some of my contributions to this project are listed below.

- **PR: #37 (merged):** Add missing camel case parser in drf.py

- **Issue: #42:** Add a validator on Period Mixin itself, raising validation error if the dates overlap.

Project Introduction

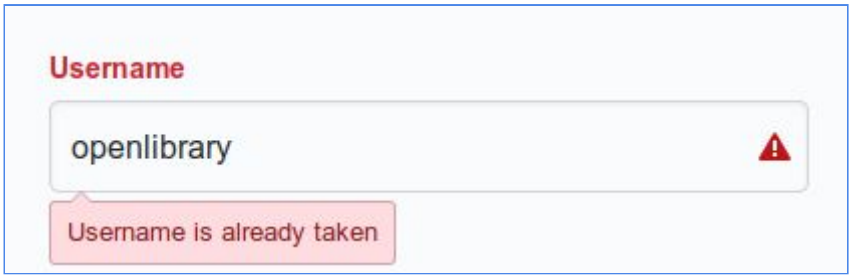
Mentors: Michael E. Karpeles

Open Library is a digital library run by Internet Archive which loans over a 2M ebooks each year to over 2M patrons around the world. For GSoC, I would like to invest my summer in working on *Improvising experience for the user and Implement a User Feedback System for Open Library.* UX is important because it tries to fulfil the user's needs. It aims to provide positive experiences that keep users loyal to the product or brand. I found the UX at Open Library broken at some parts. I would love to improve it. A User Feedback System (**UFS**) will an extremely important tool for a website of such a large user base.

Detailed Description

These are the potential ideas on which I would like to work on. I tried to keep my scope to improving user experience and to UFS, which are my targets for the project.

- Sign-up Flow:
  The current sign up flow is of an old type with no dynamic validation for the unique username. I wish to add a feature for checking availability by custom form validations. While registration when the user chooses a username (screen name), custom validations will be checked on each keypress. Based on the response the username availability will be checked and the message will be shown alongside. The implementation is similar to GitHub's checking for the username.



Courtesy: GitHub

- Implement the feedback system:
  In my journey of using Open Library on a daily basis, I felt at many places that I need to give feedback to maintainers. Sometimes editing the content isn't enough. Many times user want to tell the maintainers. I wish to implement a feedback system in the Open Library. One counter-argument for this idea might be that this would increase load for the maintainers. I am planning to introduce this by having the features of tags. The tags will be maintainer created types like "*book*", "*work*", "*login*", etc. The benefit of this **tag + comment** based feedback system would be huge for the maintainers. This would work as analytics and will immediately indicate what should be the priority issue/feature that the developers should work on.

There is an immense scope of for this feature. An example can be "*creating an admin restricted view which gives the result of the feedback system in a graphical form*", "making the comments public visible", etc. I would love to work on this under the guidance of the project mentors.

- Minimizing the support emails by writing detailed documentation:
  This includes both User and Code documentation. My target would be making documentation so rich such that all developer/API targeted queries will go through GitHub. Hence there will be a reduction in support emails. I always contribute to the wiki whenever I find some mistake or improvement.

Bucket list:

Apart from the project ideas discussed above, I would also love to work on more ideas if the time allows. If I ran out of time in the GSoC period, I would still love to work on them post-GSoC. Some of them are:

- Follow User feature:
  Adding a feature to follow a user. There will be a page where the user can find a list of people he/she wants to follow. Further, many more features can be added using the API for the book lists by the user. Also, there will be a separate view/component showing the *followers* also a separate view/component for users *following*.

- Add email notification:
  Get notified of the "want to reads" of the user you follow. Integrate Open Library with sendmail feature of web.py.

- 100% Mobile Responsive:
  This includes making every page mobile and tablet responsive also decreasing the loading time for Open Library to improve the user experience.

- Another broad work in my bucket list is solving the issues from GitHub issue tracker. I like to solve issues. I find challenging to read the code and debug the problem. So I would like to my time solving them.

Research

A list of web services and documents that I studied in the course of planning the project implementation:

Open Library Wiki

A very comprehensive wiki that helped me get deeper insights into the Open Library.

Infogami Developer Tutorial

The Open Library code is a little hard for first-timers. The Developer Tour has detailed information about the Infogami which powers Open Library.

Open Library Tour

The Open Library cannot be said as *just a digital library* since it has a lot more features in it. Some examples can be list creation, book borrowing, advanced book search. In my initial time to the project, this tour helped me understand all the features which Open Library provides.

Project Benefits and Goals

For developers and designers:

- Reduced load on GitHub issue tracker due to analytics in the Feedback system.
- The developers will get to know the priority issues automatically

For Archive.org users:

- An improved User Experience
- An attractive Signup page
- A brand new Feedback system to report query, (especially useful those who are unfriendly with GitHub)

The community is made by the developers and its users. So with their benefits are also attached to the combined benefits of its developers and the users. So in a way, the **Open Library** community will be benefitted as well.

Measuring the impact

The project target is to *Improve User Experience* and to *Implement UFS.* The improvement in User Experience can be measured by the amount of time a single user spends on openlibrary.org website. This is may be related to the traffic openlibrary.org will get. Which in turn can be directly related to the Alexa rank of the website.
The UFS will also be very helpful for the developers as well for the users. For developers, it will reduce the load as it will automatically suggest the priority issue based on the labels assigned by user feedbacks. It would be beneficial for the users because they will get an easy way to submit their feedback. There will be no need for the user to even leave the website.

Proposed Timeline

This is the timeline I will stick to when working on Open Library during GSoC.

Pre-GSoC period | Community Bonding Period

Interact with the mentors of the project and set up feedback loops.
Continue to refine the plans for the project in consultation with the mentors.
Get involved with the community which, after all, is what this period is for.

Phase 1: Form validation and start Feedback system

Week 1-2 | May 27 - June 10

Discuss the flow of improved sign up with proper dynamic validations.
Implement the signup discussed.
Improve the CSS of the new validation fields implemented

Week 3-4 | June 11 - June 24

Discuss the features to be included in the feedback system.
Discuss the tags to be included and the flow of the feedback.
Create wireframes and mockups for feedback view.

Phase 2: Complete Feedback system

Week 5-6 | June 25 - July 8

Create the API for the feedback system.
Create the frontend that displays feedback forms and multiple tag selection fields.
Discuss the features and frontend for the analytics of the admin view.

Week 7-8 | July 9 - July 22

Create the mockups the analytics page.
Create the frontend for the analytics view.
Rigorous testing of the feedback system and analytics.

Phase 3: Implementing follow user feature

Week 9-10 | July 23 - August 5

Develop APIs following user.
Create the mockups for *user followers* and *user following view*.
Associate list created with the users followed.

Week 11-12 | August 6 - August 19

Discuss the features of the email notifications with mentors.
Integrate email notification with the followings

Week 12-13 | August 20 - August 26

Buffer period to catch up on any pending work, else:

Make bug fixes and code quality refactorings.

Add any small features that might've been overlooked.

Add easter eggs as a special give to the community.

Post GSoC period | finishing touches

Continue to work on the project, finishing items that have been put on an if condition in the last two weeks.

Improve the codebase in aspects that will only come to light when integrations are getting merged.

Distant future | long-term commitment

In the distant future, I would love to work with Open Library / Internet Archive as they build projects which truly help large scale community. Truly, Internet Archive as a highly active and promoting community. I would love to work more projects if I can.

Improve the codebase in aspects that will only come to light when integrations are getting merged.

Availability

My vacations start on 5th May 2019 and end on 11th July 2019. The official GSoC period is from 6th May 2019 to 26th August 2019. I can easily devote **50-55 hours per week** until my college reopens and **45-50 hours per week** after that. I am highly productive on weekends and intend to complete most of my work before my college reopens.

Other than this project, I have no other commitments during the vacations in the summer. I shall keep my status posted to all the community members on a weekly basis and maintain transparency in this project. Moreover, most of my work hours largely coincide with the mentor's work hours. So this way it would be very comfortable for both us to communicate and this will speed up the code review workflow.

After GSoC

Are you interested in working with Internet Archive after GSoC ends?

Yes. I'm interested as I always was. I have been contributing to **Internet Archive** for over three months now and I've really loved the experience. I got familiar with the community and I believe I have learned a lot interacting with the prospective mentors. I feel this kind of mentorship is necessary and has contributed to my growth. I will be an active member of the community and keep contributing. My motivation would always be that I had been able to contribute to something big and widely in use. This gives me a lot of satisfaction. I would also love to develop my mentorship skills so that I can give back to this community by nurturing the future generations of contributors.

* * * * * * * *