Summer of Code

Pradeeban Kathiravelu, Ph.D.

Dept of Biomedical Informatics,

Emory University

Disclaimer

- This presentation is made for GSoC 2023.
  - Based on the past observations.
- Future GSoC rules may change.
  - Pay attention to GSoC official rules.

Intro to me

(a.k.a. My GSoC Journey)

GSoC didn't have limits on how many times one can be a contributor back then.

- 2009: Student/AbiWord
- 2010: Student/OMII-UK
- 2011 – 2013: Mentor/AbiWord
- 2014 – 2015: Student/Emory BMI
- 2016: Mentor/Emory BMI
- 2019: Organization Administrator/Emory BMI
- 2020: Organization Administrator/caMicroscope
- 2021 – Now: Organization Administrator/Emory BMI

Contents

- Why GSoC?
- Before you begin..
- Right Project?
- Apply.
- Code.
- Conclude/Continue.
- What Else?

Why GSoC?

- Code for your preferred open source project for 3 months.
  - And get paid!
  - Open for all the adults (18+) from 2022
    - Not only for students.
  - Must be an ***open-source beginner.***

Stipends

- **Stipends based on your residence location.**
  - Option for:
    - Large Size Projects / Full-time
      - (35 hrs/week) / 350 hours in total.
      - 3000 – 6600$.
    - Medium Size Projects / Half-time
      - (17.5 hrs/week) / 175 hours in total.
      - 1500 – 3300$.

Background Image: http://blog.echoenduring.com/2010/11/23/learning-about-design-from-angry-birds/

Google Summer of Code (GSoC)

- Code for an open source organization for 3 months.
- Have not participated in GSoC as a contributor more than once!
- Google coordinates and rewards you!
  - Getting Accepted.
  - 2 milestones.
    - First Evaluation. (45% of the stipend)
    - Final Evaluation. (55% of the stipend)
      - A completion certificate!

Some statistics of 2019

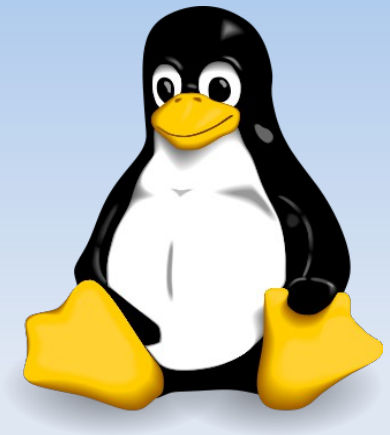
- 201 Organizations
  - 2,815 mentors and co-mentors.
  - 2,066 with an assigned contributor.
- Registered
  - 30,922 contributors, from 148 countries.
- Accepted
  - 1, 276 contributors/projects.
- Completed successfully
  - 1,134.

Success Rate..

is pretty high!

- Most accepted contributors passed the final evaluations
  - Success rate – 88.87% (2019).

Why Google Summer of Code?



What do you need?

Timeline (2022)

- Feb 7th – **Feb 21st**: Organizations apply.
- March 7th: Accepted organizations announced.
- April 4th – **April 19th**: Contributor application period.
- **May 20th**: Accepted contributors announced.

Timeline..

After getting accepted

- May 20thth: Community Bonding Period Begins.
- **June 13th – Sep 5th: Coding Period.**
- July 25th – **July 29th**: First Evaluation.
- Sep 12th – **Sep 19th**: Final Evaluation.
- Extended time for coding.
  - Plan with your mentor if you need additional time to finish your project.
- **Nov 21st**: Contributor “Final” (Extended) Final Submission.

Before you begin..

- Google Summer of Code is all about being Open Source.
- Get your basics and motives right.
- Netiquettes.
- Sign up to the lists.
- Join the relevant communication channel.

Technologies..

- Version Control Systems
  - git, mercurial, ..
- Build Tools
  - Maven, Ivy, Ant, ..
- IDEs (Integrated Development Environments)
  - IntelliJ IDEA, Eclipse, ..
  - Microsoft Visual Studio ..
- Issue Tracker
  - GitHub, Jira, Bugzilla, Trac, ..

Communicating with the team..

and the mentor, over the Internet

- Mailing Lists
  - Dev, User, Commit lists, sub-groups, ..
- Internet Relay Chat (IRC) and Slack channels
- Issue Tracker
- Forums and wiki
- Blogs
- Skype, Personal Mails, conference calls, .. [with the mentors, if that is preferred.]

Network Etiquettes

- Be specific – brief and clear.
- Research (google.. ;)) before asking.
- Be helpful to others.
- Be ethical; respect.
- NO CAPS! (UNLESS YOU ARE SHOUTING!)
- Don't take messages personally.
- Dn't snd ur sms msgs to thrds or lsts.
- Language/English

Proper Addressing..

over the lists/irc/..

- Address the devs and users properly.
  - First Name or Preferred calling name.
  - NO Sir, Madam, bro, sis, pal..
    - Even if you know them, personally.
  - Not too personal.
    - Use “Hi”, instead of “Dear”.
  - Be gender neutral.
    - “Folks” over “Guys and Girls”.
  - Don't misgender – Refer to folks using correct gender.
    - If not known, use “they” instead of “he/she.”

Mailing lists

- Post only to the relevant list/channel.
- Check the mail archives or channel logs first.
- No [URGENT]/[IMPORTANT] tags.
- No unnecessary attachments.
- No Cross Posting.
- Don't hijack threads.
- Don't post off-topic.

IRC/Slack Etiquettes

- Join the relevant channels.
- Be an observer first.
- Refer to others using their irc/slack nick to get attention.
  - But avoid misusing.
    - No @channel or @here.
- Don't expect immediate replies; wait.
- Discuss in the channel. Minimize direct messages to mentors, unless they suggest otherwise.

Find a mentoring organization..

- Have a look at the list of past GSoC.
- 201 in 2019!
- New Organizations.
- Google as the mentoring organization.
- Introduce GSoC to an organization.

Find THE right project..

- Go through the organizations' project ideas list.
- Check for the previous years' GSoC ideas from the potential organizations.
  - until the accepted organizations are announced and their ideas lists are published.

Project Ideas List

- Project ideas by the organizations are intentionally left generic enough.
  - Open for interpretation by the contributors.
  - Clarify with mentors on assumptions and chosen approach.

[4] Creating shareable "albums" from Niffler data sets

Mentors: Judy Gichoya (judywawira -at- emory.edu) and Pradeeban Kathiravelu (pradeeban.kathiravelu -at- emory.edu)

Overview: Niffler is a framework to retrieve DICOM images from PACS real-time as a DICOM stream as well as retrospectively. Images can be retrieved from a PACS via Niffler in real-time (via Niffler meta-extraction module) or on-demand (via Niffler cold-extraction module). However, these downloaded data sets remain in the local environments such as a research server or a cluster where Niffler is run from. To use this data, researchers must identify certain subsets of data. This can be achieved by querying the retrieved data. For instance, Niffler stores the metadata of the data retrieved in real-time in a Mongo database. By querying the metadata, subsets of images can be identified. However, currently Niffler does not possess the ability to create such "albums" from a set of DICOM images retrieved by Niffler, and share with other users.

Present Status of the work: Currently, Niffler does not have the ability to select subsets of images or create albums. We are sharing images through other orthogonal approaches (via rclone, for example).

Proposed Methodology: There are several approaches to implement such albums feature. One approach is to using Kheops to provide an interface to create and view the albums. MEDIator can be extended and incorporated to Niffler to create subsets and share the images via a unique URL as well.

Benefits: The proposed feature will make the images retrieved by Niffler accessible by more researchers for their experiments, by replacing the current manual efforts of data sharing. Moreover, Kheops natively integrate with OHIF Viewer. As such, images retrived by Niffler can be viewed through OHIF Viewer, by creating albums with Kheops.

Deliverables: An approach to creating shareable datasets from the DICOM images retrieved by Niffler. It could be adopting existing frameworks such as MEDIator and Kheops and scripts and integration code with those frameworks or an entirely new module to Niffler for this feature. However, students are encouraged to use Kheops or alternatives, rather than reinventing the wheel (unless there is a convincing reason).

Required Skills: Python and Java.

Code Challenge: A demonstration of potential integration of Niffler with such existing frameworks. The proposed frameworks are samples only. The students may choose their own.

Source Code: https://github.com/Emory-HITI/Niffler/

Slack room: gsoc-emory-bmi.slack.com niffler

Effort: Half-time

Get to know more..

about the projects

- Talk to the mentor(s)
  - Assigned by the organization
  - for each project idea.
- Mailing lists and archives.
- Issue Tracker
  - Open issues or tickets
    - New features/enhancements (RFE)
    - Bugs (easy/difficult and normal/critical)

What makes you
special?

- Experience
- Your interests and motivation
  - Pick something you really enjoy doing.
- Opportunities
  - What makes you the right person?

Are you willing

to contribute further?

- Willingness
  - To contribute to the community
  - Beyond the time frame of GSoC.
- We want committers and long-term collaborators!
  - Not just contributors for the summer!

Experience

- Language
  - Java, Python, Javascript, Go, C++, C, ..
  - Not much time to learn a new language (?)
- Prove It!
  - Patches.
  - Code challenges.
  - Assist other new potential contributors!!!
  - Project expertise
    - Bug reports and fixes.
    - Go through the archives, wikis, and web sites.

Opportunities..

- Project that matches your previous work experience.
- Choose the right project.
- Timezone Difference
  - Use it effectively
- Multiple Applications
  - Up to 3!!! But only one accepted per contributor.
- Preferences!

Communicate

- Communicate early – but meaningfully.
  - Not necessary to send a message just to say ‘hi’.
- Communicate often.
- Ask **meaningful** questions.
- Answer others' questions!
- Mentors are here to help!

  (respect)

Avoid lazy-sounding questions

- I am new to GSoC. Please help me get started.
  - Follow the GSoC guidelines and organization's docs.
- Where to submit the proposal? When is deadline?
  - General GSoC details are in the GSoC guidelines. Use mentor's time and help more productively.
- I like to do project idea #4. Please help.
  - Unclear what help you expect.
- Can you explain the coding challenge?
  - Do you want us to paraphrase the whole coding challenge? What do you not understand?

Avoid vague/generic questions

- I cannot compile Bindaas. Please help.
  - Where exactly did you get stuck? What is your OS? Explain the execution environment.

Ask clear specific questions

- I followed the project idea and achieved the below as a demo. However, the last step is not clear. Does that mean I should clear the contents first?
  - Shows effort and progress.
- I compiled Bindaas on Ubuntu 20.04 with Oracle JDK 1.8. However, it fails with some error messages. The full logs are attached here as a text file. What went wrong?
  - Sufficient details included. Even works as a bug report.

Be Known..

- Be heard!
  - Quality over quantity.
- Be visible!
- Be responsive!
- Be quick!

Apply

- Register as a contributor for GSoC.
- Use the project's wiki for draft proposal
  - if applicable.
- Apply on Google's system.
  - Can edit later, until *the last minute!*
  - Get the mentors' opinions and improve.
- Check often for the mentors' comments
  - attend to them.

Proposal

How to impress the mentors/developers?

- Stick to the organization's template.
- Abstract.
- Introduce yourself properly.
  - Focus on the relevant facts.
  - Why do you fit? Your skill sets.
  - List of the patches (if any) you have submitted.

Propose..

- Project Goals
  - Proves you got them correct.
- Deliverables
  - Code, Documentation, test cases, ..
- Description
  - Benefits to the organization and other projects.
  - Can also be given along with the timeline.

Propose..

- Timeline
  - Finer details.
  - Break up to periods of 3 - 4 days or a week.
  - Testing takes time.
  - Don't be over-optimistic.
  - Require 35 or 17.5 hours of work per week based on the project type (large vs medium size projects).
- Links
  - References and additional details.

Application Template

- Name:
- Email:
- Project Title:
- Synopsis:
  - A short description of your project.
- Benefits to the organization/project
  - and/or other project(s):
- Deliverables:
  - Quantifiable results.

- e.g.: "*At the end of my project, Niffler Metadata Processing Pipeline will be 3 times faster*."

- Project Details:
  - A more detailed description of your project:
- Project Schedule:
  - How long will the project take?
  - When can you begin work?
  - Do you know of any planned absences or other major conflicts
    - summer classes, vacations, etc.

- Bio:
  - Who are you?
  - What makes you the best person to work on this project?
- Additional Requirements:
  - Patches / Specific requirements for the project.
- Further Related Information

After the submission..

- Don't go invisible!
  - Evaluation is still going on.. ;)
- You may be asked to give more details
  - Bug fixes / pull requests.
  - Demos and screenshots.
- You may even start coding on your project!
  - Especially if you didn't apply for multiple projects.
- Be motivated.

Got Selected? \yay/

- Don't Panic.
- You have one more month
  - to mingle with the developers and the code base.
- More one-to-one interaction with the mentor(s) once accepted!
- Keep in touch with the developers.

Community Bonding Period

- Learn the project
  - Go through the code base
  - Documentation.
    - Coding styles and coding guide lines.
- Communicate often
- Understand the project idea more.
  - Come up with a design.
  - Start with simple fixes and incremental development.

Coding..

Easiest task of all.. ;)

- Commit often, if given committership.
  - Send frequent pull requests (daily ?) otherwise.
  - Meaningful Commit messages.
- Get feedback from the mentor(s).
- Keep the community/mentor updated
  - Daily (?).
- Plan for the mid and final evaluations early with the mentor.

Conclude/Continue..

- Final Submission – a link showing your GSoC work
  - If a new project/module, can be a link to the source repo.
  - Or a web page listing/pointing to your exact commits.
- Don’t miss any of the evaluations.
  - Both mentors and contributors must submit evaluations.
  - All 2 of the evaluations!
- Focus on becoming a committer in the organization
  - if not already given committership.
- Keep contributing.

What else?

- More FOSS?
- Annual?
- Stipend?
- Student?
- Country/Location?

More Open Source

programs/contests..

- Season of KDE (SoK)
- Outreachy
- Google Season of Docs (GSoD) – for documentation.
- Hyperledger Mentorship Program
- The X.Org Endless Vacation of Code (EVoC)

Some links

- GSoC home page
- Google Open Source Blog on GSoC
- Proposal [Pradeeban]
  - GSoC 2015 Emory BMI
  - GSoC 2014 Emory BMI
  - GSoC 2010 OMII-UK/OGSA-DAI
  - GSoC 2009 AbiWord

Are you ready?

- Past successful proposals - Wikis, blogs, ..
- Ideas list:
    - Emory BMI
- Apache Software Foundation
  - More slots and more choices.
    - Tomcat, Derby, Axis2, and more ..
- Join the projects' mailing lists and channels.
  - Apache CXF, ...

For more Information ..

- Join the GSoC mailing lists.
- Check the GSoC official site and blog.
- Join your local GSoC Google Group
  - For e.g., Group for Sri Lankan contributors: http://groups.google.com/group/gsoc-srilanka


Questions?..