

Git OSS-um: A Tool for Newcomers to Open Source Software



Team Members

- Gary Baker (Team Lead)
 - gpb27@nau.edu
- Van Steinbrenner
- Stephen White

Mentor: Ana Steinmacher

Sponsor: Igor Steinmacher

What is Open Source Software?

- Principles:
 - Transparency
 - Open exchange
 - Collaborative participation
 - Community-oriented development
- Contribution Process



825,000 commits

1 branch

597 releases

∞ contributors

View license

Branch: master ▾


New pull request














Find File

Clone or download ▾

 torvalds Merge tag 'nfs-for-5.1-3' of git://git.linux-nfs.org/projects/trondmy...

Latest commit 14c741d 23 hours ago

 Failed to load latest commit information.

 Documentation	Merge tag 'irqchip-5.1-2' of git://git.kernel.org/pub/scm/linux/kerne...	6 days ago
 LICENSES	LICENSES: Add GCC runtime library exception text	2 months ago
 arch	Merge branch 'x86-urgent-for-linus' of git://git.kernel.org/pub/scm/l...	3 days ago
 block	Merge tag 'io_uring-20190323' of git://git.kernel.dk/linux-block	4 days ago
 certs	kexec, KEYS: Make use of platform keyring for signature verify	2 months ago
 crypto	lib/lzo: separate lzo-rlc from lzo	20 days ago
 drivers	Revert "parport: daisy: use new parport device model"	2 days ago
 fs	Merge tag 'nfs-for-5.1-3' of git://git.linux-nfs.org/projects/trondmy...	23 hours ago
 include	Revert "parport: daisy: use new parport device model"	2 days ago
 init	init/main: add checks for the return value of memblock_alloc*()	15 days ago
 ipc	Merge branch 'work.mount' of git://git.kernel.org/pub/scm/linux/kerne...	15 days ago
 kernel	ftrace: Fix warning using plain integer as NULL & spelling corrections	a day ago
 lib	Merge tag 'kbuild-v5.1-2' of git://git.kernel.org/pub/scm/linux/kerne...	10 days ago

README

Linux kernel

=====

There are several guides for kernel developers and users. These guides can be rendered in a number of formats, like HTML and PDF. Please read `Documentation/admin-guide/README.rst` first.

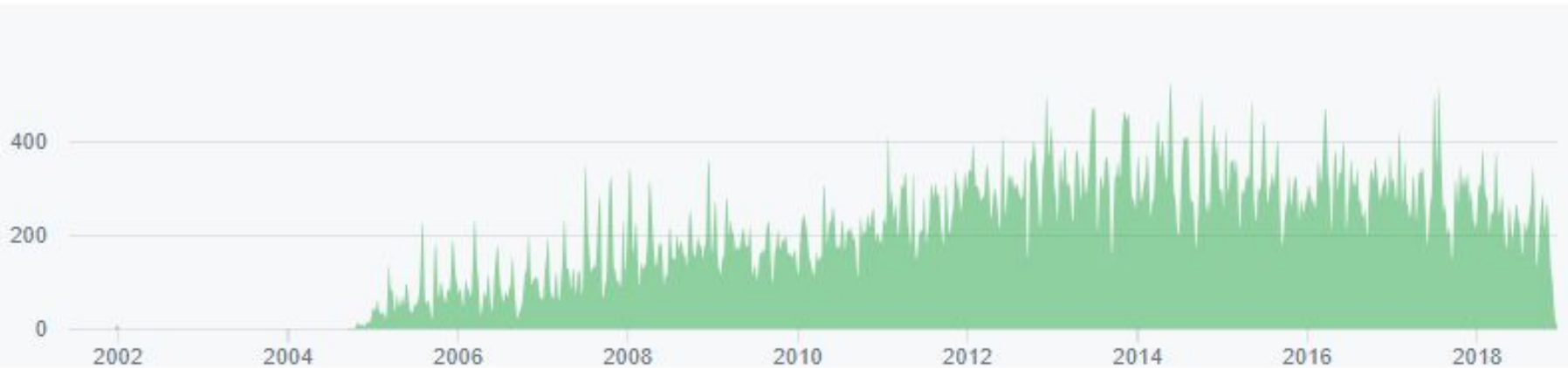
In order to build the documentation, use ```make htmldocs``` or ```make pdfdocs```. The formatted documentation can also be read online at:

<https://www.kernel.org/doc/html/latest/>

There are various text files in the `Documentation/` subdirectory, several of them using the Restructured Text markup notation.

Please read the `Documentation/process/changes.rst` file, as it contains the requirements for building and running the kernel, and information about the problems which may result by upgrading your kernel.

GitHub's "Activity" Statistics (Torvalds/Linux)



jacob-and-nathan commented 19 days ago

...

There was a double space.



1



2



Fixed typo: extra space. ...

Verified

ee9ca9c

KernelPRBot commented 19 days ago

...

Hi @jacob-and-nathan!

Thanks for your contribution to the Linux kernel!

Linux kernel development happens on mailing lists, rather than on GitHub - this GitHub repository is a read-only mirror that isn't used for accepting contributions. So that your change can become part of Linux, please email it to us as a patch.

Sending patches isn't quite as simple as sending a pull request, but fortunately it is a well documented process.

Here's what to do:

- Format your contribution according to kernel requirements
- Decide who to send your contribution to
- Set up your system to send your contribution as an email
- Send your contribution and wait for feedback

How do I format my contribution?

The Linux kernel community is notoriously picky about how contributions are formatted and sent. Fortunately, they have documented their expectations.

Firstly, all contributions need to be formatted as patches. A patch is a plain text document showing the change you want to make to the code, and documenting why it is a good idea.

You can create patches with `git format-patch`.

Secondly, patches need 'commit messages', which is the human-friendly documentation explaining what the change is and why it's necessary.

Thirdly, changes have some technical requirements. There is a Linux kernel coding style, and there are licensing requirements you need to comply with.

Both of these are documented in the [Submitting Patches](#) documentation that is part of the kernel.

Note that you will almost certainly have to modify your existing git commits to satisfy these requirements. Don't worry: there are many guides on the internet for doing this.

Who do I send my contribution to?

What does the future look like?

- 1.1 Billion total contributions
- 31 Million People
- 96 Million Repositories
- 200 Million Pull Requests
- The impact of 10%



Our Client

- Dr. Igor Steinmacher, professor at NAU
- Research includes:
 - Support for newcomers in OSS
 - Contributor behavior in OSS
 - Crowdsourcing software engineering



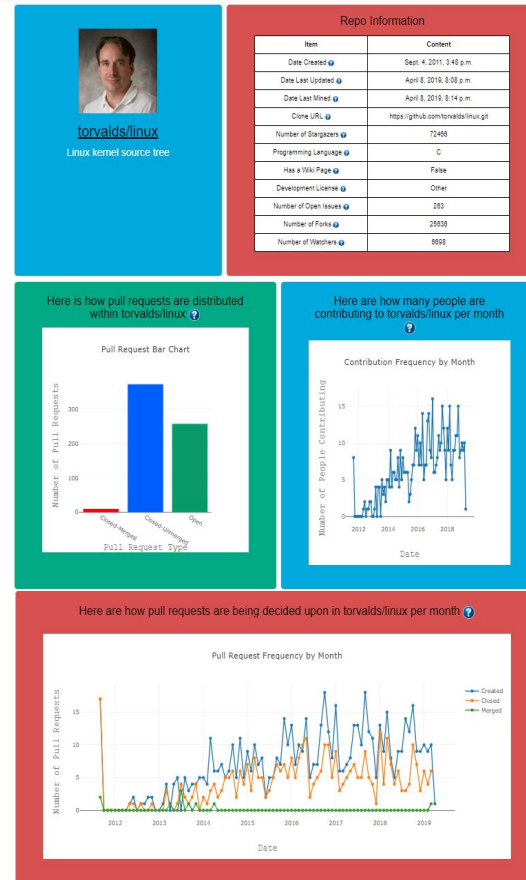
Newcomer Points of Concern

- **Finding a candidate project**
 - Finding your first project to submit a contribution to can be daunting and difficult.
- **Information Overload**
 - Repositories can be very confusing because of the amount of information they give at the start.
- **Unable to Gauge Activity**
 - Finding statistics of activity for repositories is difficult for newcomers not knowing where to look.
- **Unaccepting of New Contributions**
 - Newcomers don't want to look foolish when making their first attempt at contributing.



Product: A web application that helps a newcomer to Open Source find a project to contribute to by:

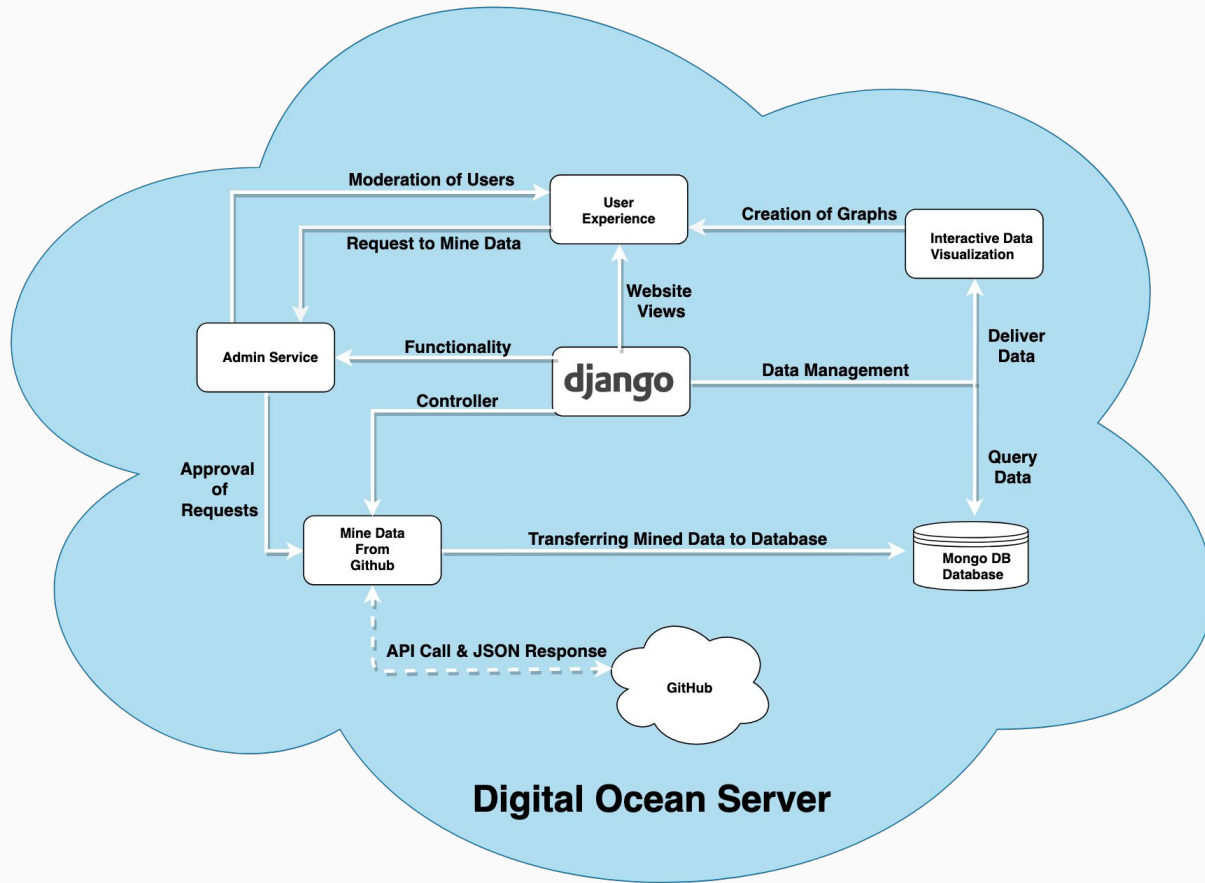
- Displaying simple and easy to read statistics that don't overload the newcomer with information.
- Helping gauge the activity of a repository.
- Showing the acceptance rate of submitted contributions.

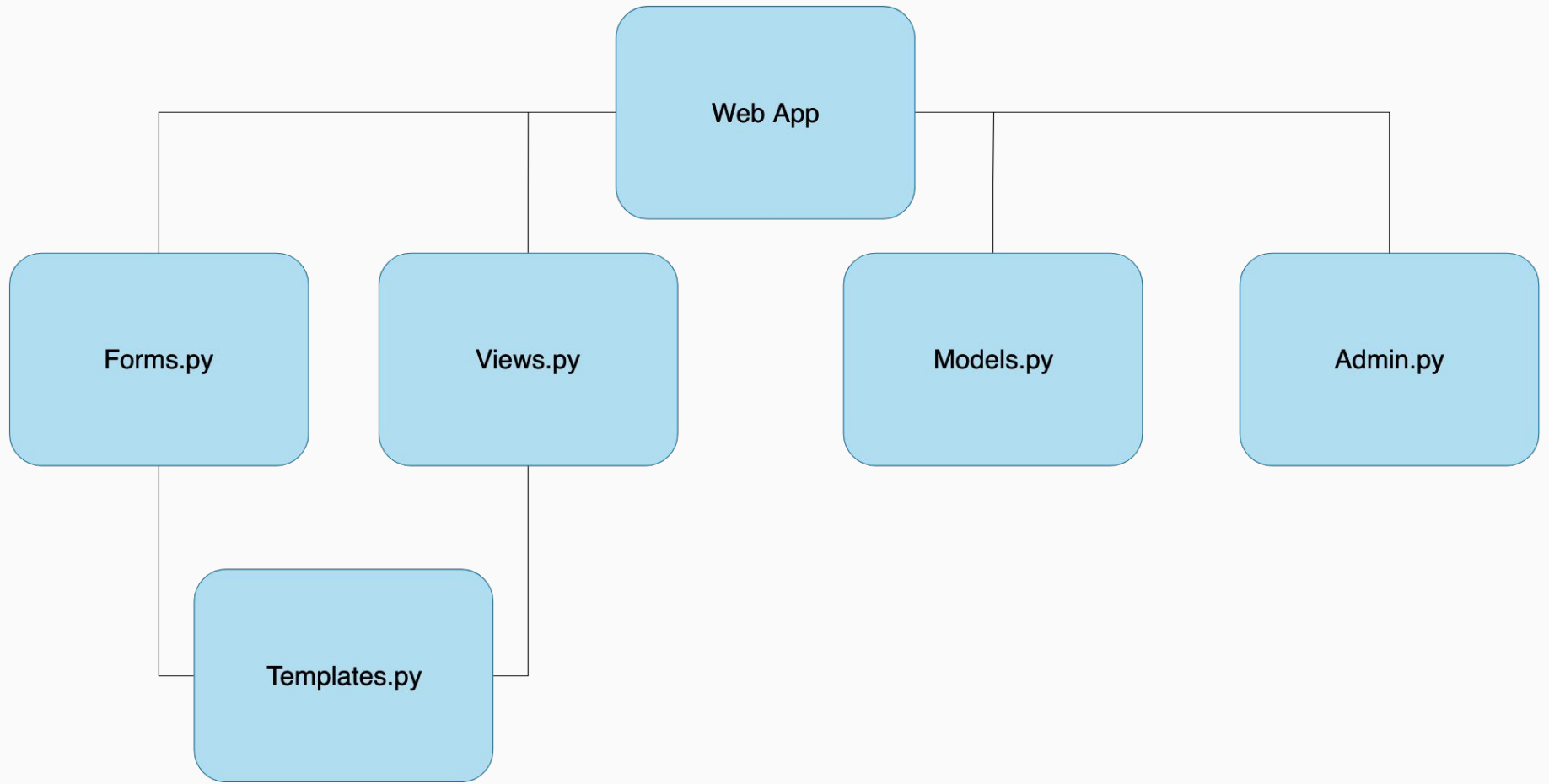


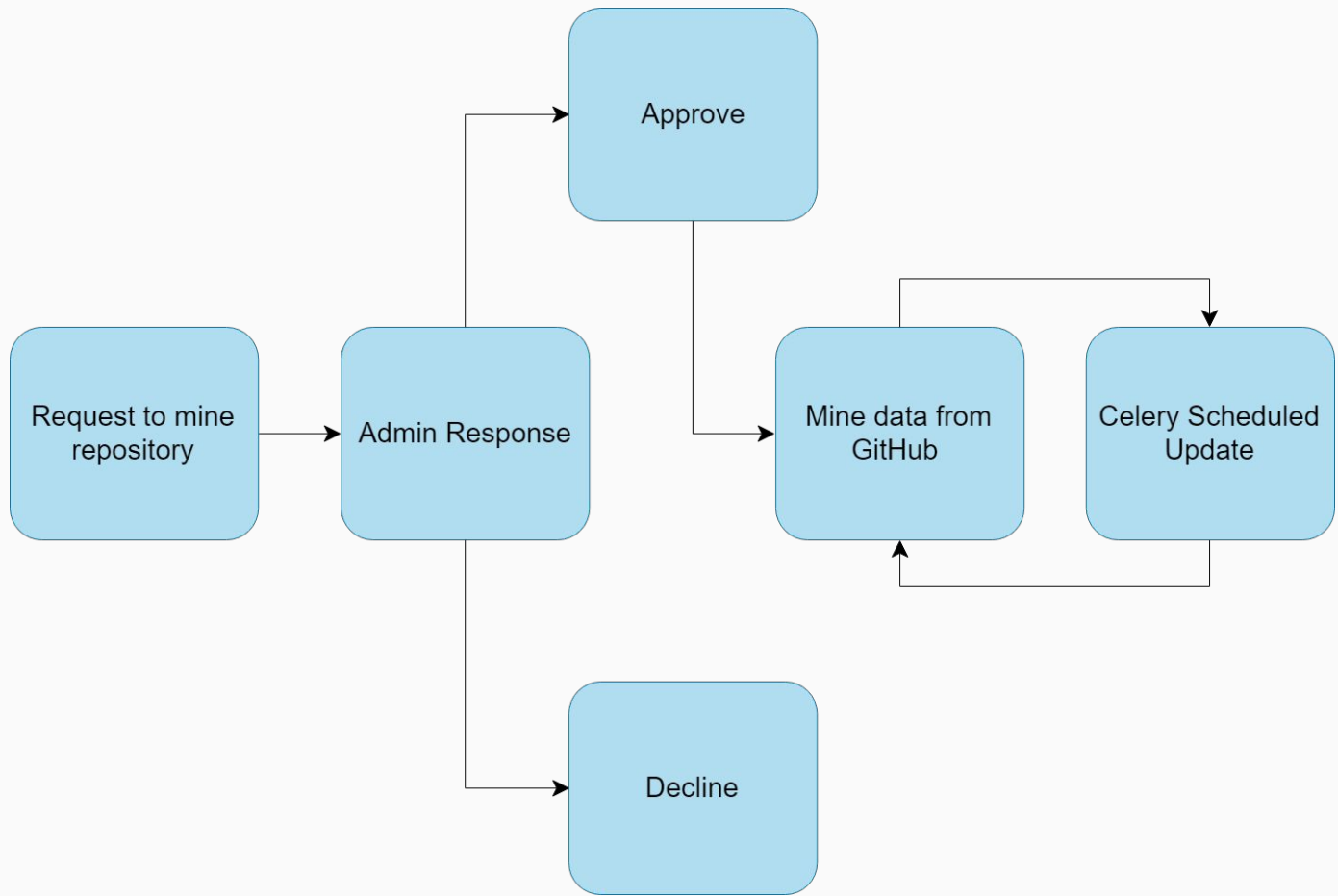
High-Level Requirements

- Login Services
- Mining Data From GitHub
- Data Storage
- Analyze and Visualize Data
- Hosted Web Application









How the app works

- **Use Case:**
 - As a user, I would like to be able to view easily readable data about a repository I am interested in to submit my first Open Source contribution.
- **User Flow:**
 - A user would start at our homepage, go to mined repos page, click on repo and view data on repo page.

Welcome to Git OSS-um!

You want to contribute, we want to help!

Product Description

Git OSS-um is a tool to help newcomers to open source software find the best project to make their first contribution.

Enjoy and Happy Coding!

We hope our product helps you make a successful first contribution!

Mined Repos (16)

■  blockchain/blockchain-wallet-v4-frontend

■  django/django

■  elixir-lang/elixir

■  google/gumbo-parser

■  jabref/jabref

■  kriadmin/30-seconds-of-python-code

Compare Repos

Filters

Search


Search For Repositories

Languages

- C (1)
- C++ (1)
- Elixir (1)
- HTML (1)
- Java (1)
- JavaScript (2)
- Python (5)
- R (1)
- Ruby (1)

Min pull requests





Open Source Newcomer Issues:

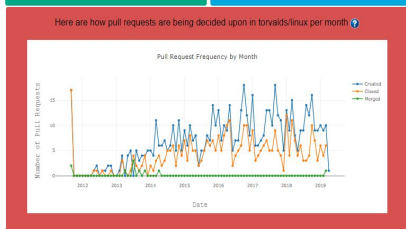
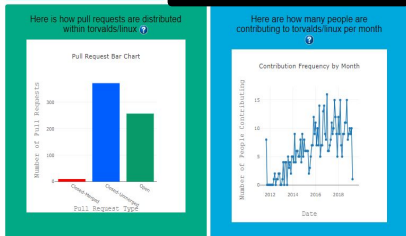
- Finding a Project 














torvalds/linux

Linux kernel source tree

Repo information	
Item	Content
Date Created 	Sept. 4, 2011, 3:48 p.m.
Date Last Updated 	April 18, 2019, 1:56 p.m.
Date Last Mined 	April 18, 2019, 3 p.m.
Clone URL 	https://github.com/torvalds/linux.git
Number of Stargazers 	7368
Programming Language 	C
Has a Wiki Page 	False
Development License 	Other
Number of Open Issues 	262
Number of Forks 	25822
Number of Watchers 	6708



Item	Content
Date Created 	Sept. 4, 2011, 3:48 p.m.
Date Last Updated 	April 18, 2019, 1:56 p.m.
Date Last Mined 	April 18, 2019, 3 p.m.
Clone URL 	https://github.com/torvalds/linux.git
Number of Stargazers 	73684
Programming Language 	C
Has a Wiki Page 	False
Development License 	Other
Number of Open Issues 	262
Number of Forks 	25822
Number of Watchers 	6708

Open Source Newcomer Issues:

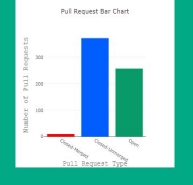
- Information Overload 



torvalds/linux
Linux kernel source tree

Repo information	
Item	Content
Date Created ip	Sept 4, 2011, 3:40 a.m.
Date Last Updated ip	April 8, 2018, 9:02 a.m.
Date Last Commit ip	April 8, 2018, 9:19 a.m.
Clone URL ip	https://git.kernel.org/torvalds/linux.git
Number of Stars ip	73408
Programming Language ip	C
Has a Readme Page ip	False
Development License ip	Other
Number of Open Issues ip	235
Number of Pulls ip	2808
Number of Followers ip	8858

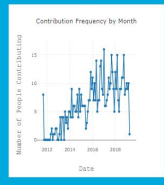
Here is how pull requests are distributed within torvalds/linux [?](#)



Number of Pull Requests

Pull Request Type

Here are how many people are contributing to torvalds/linux per month [?](#)

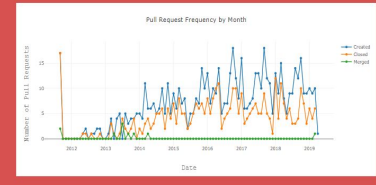


Contribution Frequency by Month

Number of People Contributing

Date

Here are how pull requests are being decided upon in torvalds/linux per month [?](#)

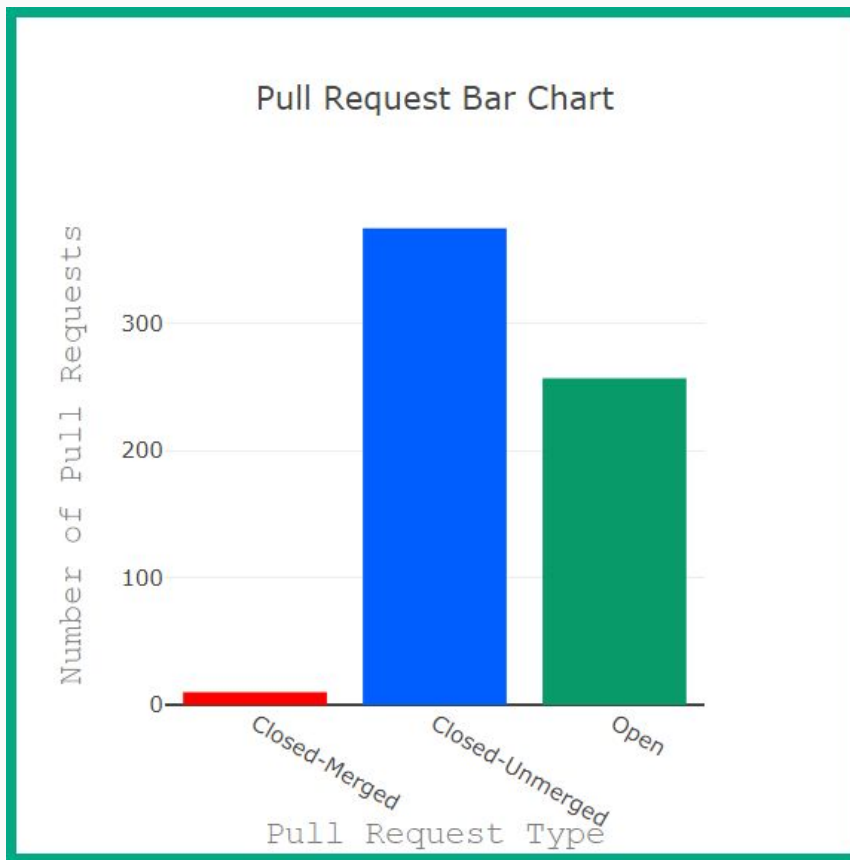


Pull Request Frequency by Month

Number of Pull Requests

Date

Legend: Closed (blue), Closed-Unmerged (orange), Open (green)



Open Source Newcomer Issues:

- Inability to Gauge Activity ✓



torvalds/linux
Linux kernel source tree

Repo information

Item	Content
Date Created ip	Sept 4, 2011, 3:40 a.m.
Date Last Updated ip	April 8, 2018, 9:08 a.m.
Date Last Commit ip	April 8, 2018, 9:10 a.m.
Clone URL ip	https://git.kernel.org/torvalds
Number of Stars ip	73408
Programming Language ip	C
Has a Read Me ip	False
Development License ip	Other
Number of Open Issues ip	285
Number of Pulls ip	2808
Number of Followers ip	8058

Here is how pull requests are distributed within torvalds/linux



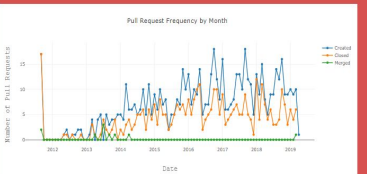
Pull Request Bar Chart

Here are how many people are contributing to torvalds/linux per month

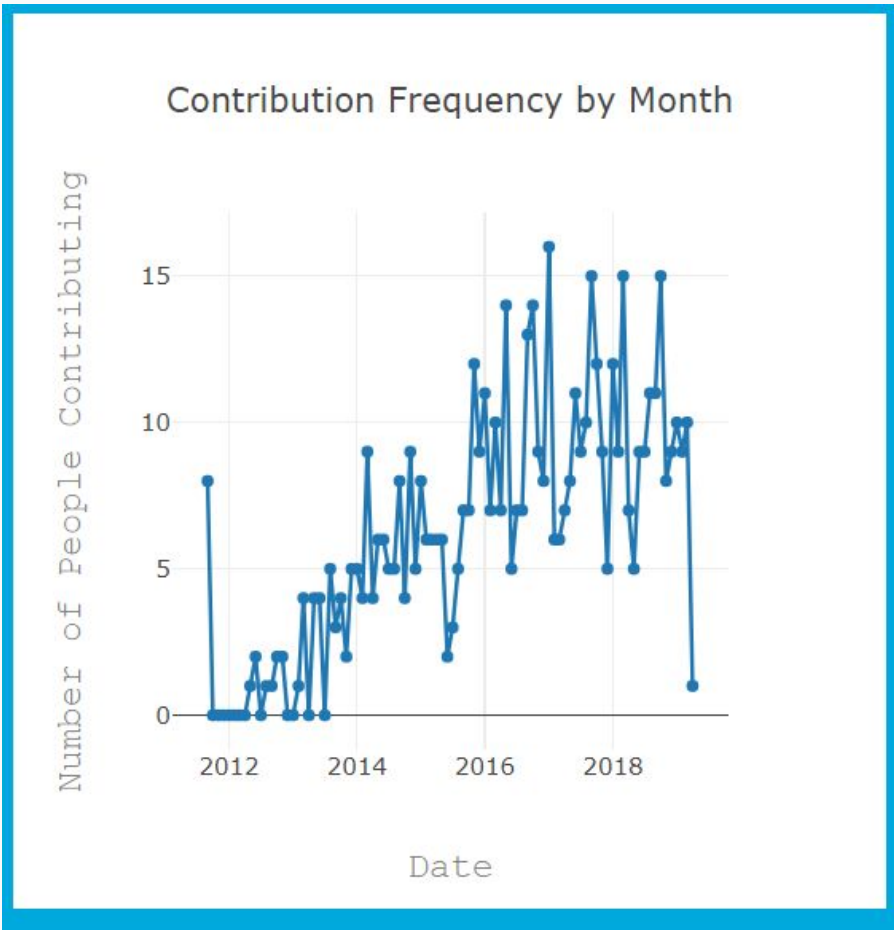


Contribution Frequency by Month

Here are how pull requests are being decided upon in torvalds/linux per month



Pull Request Frequency by Month



Open Source Newcomer Issues:

- Unaccepting of New Contributions






torvalds/linux
Linux kernel source tree

Repo information

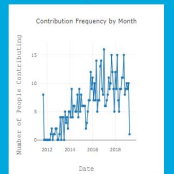
Item	Content
Date Created ip	Sept 4, 2011, 3:40 p.m.
Date Last Updated ip	April 8, 2019, 9:58 p.m.
Date Last Commit ip	April 8, 2019, 9:19 p.m.
Clone URL ip	https://git.kernel.org/torvalds
Number of Stars ip	73408
Programming Language ip	C
Has a Readme Page ip	False
Development License ip	Other
Number of Open Issues ip	235
Number of Pulls ip	2808
Number of Followers ip	8858

Here is how pull requests are distributed within torvalds/linux



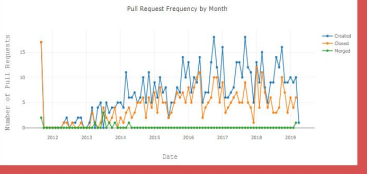
Pull Request Bar Chart

Here is how many people are contributing to torvalds/linux per month

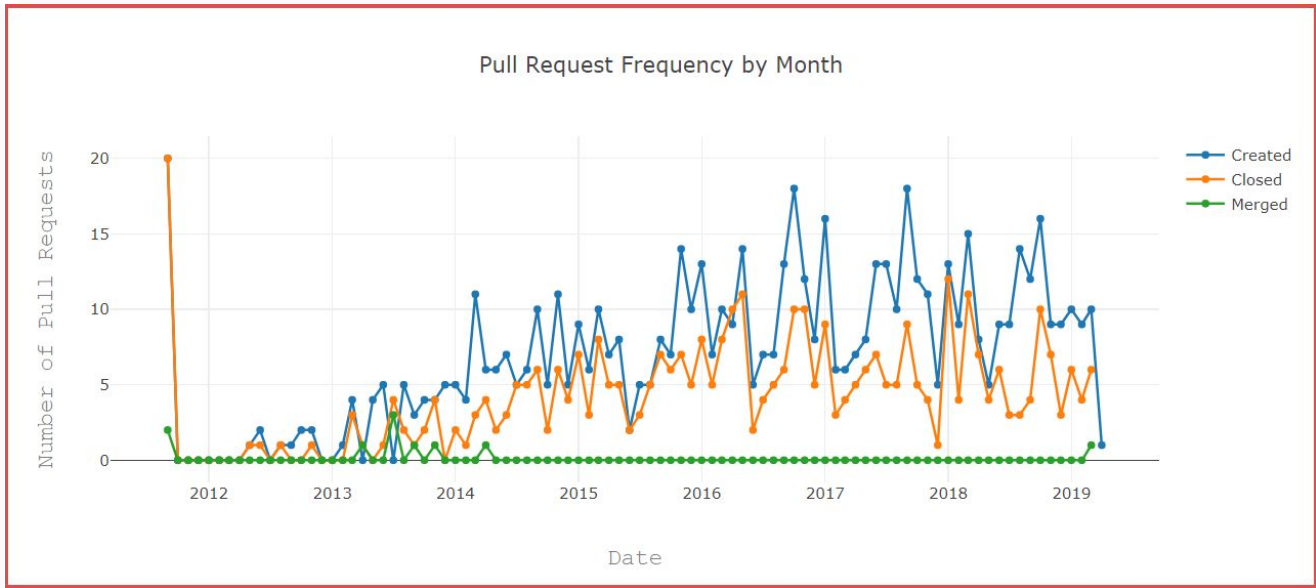


Contribution Frequency by Month

Here is how pull requests are being decided upon in torvalds/linux per month



Pull Request Frequency by Month



Testing

- **Unit Testing**
 - 100 Unit Tests
 - 7 Test Suites
- **Integration Testing**
 - Combining unit tests as well as front-end testing.
- **User Testing**
 - Recruiting students and faculty to perform tasks using our app.



Risks & Obstacles

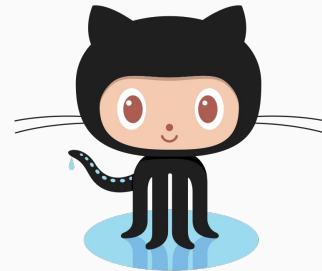
Risks

- We provide facts, not opinions
- Repository name changes
- Microsoft changing API
- Errors during data collection

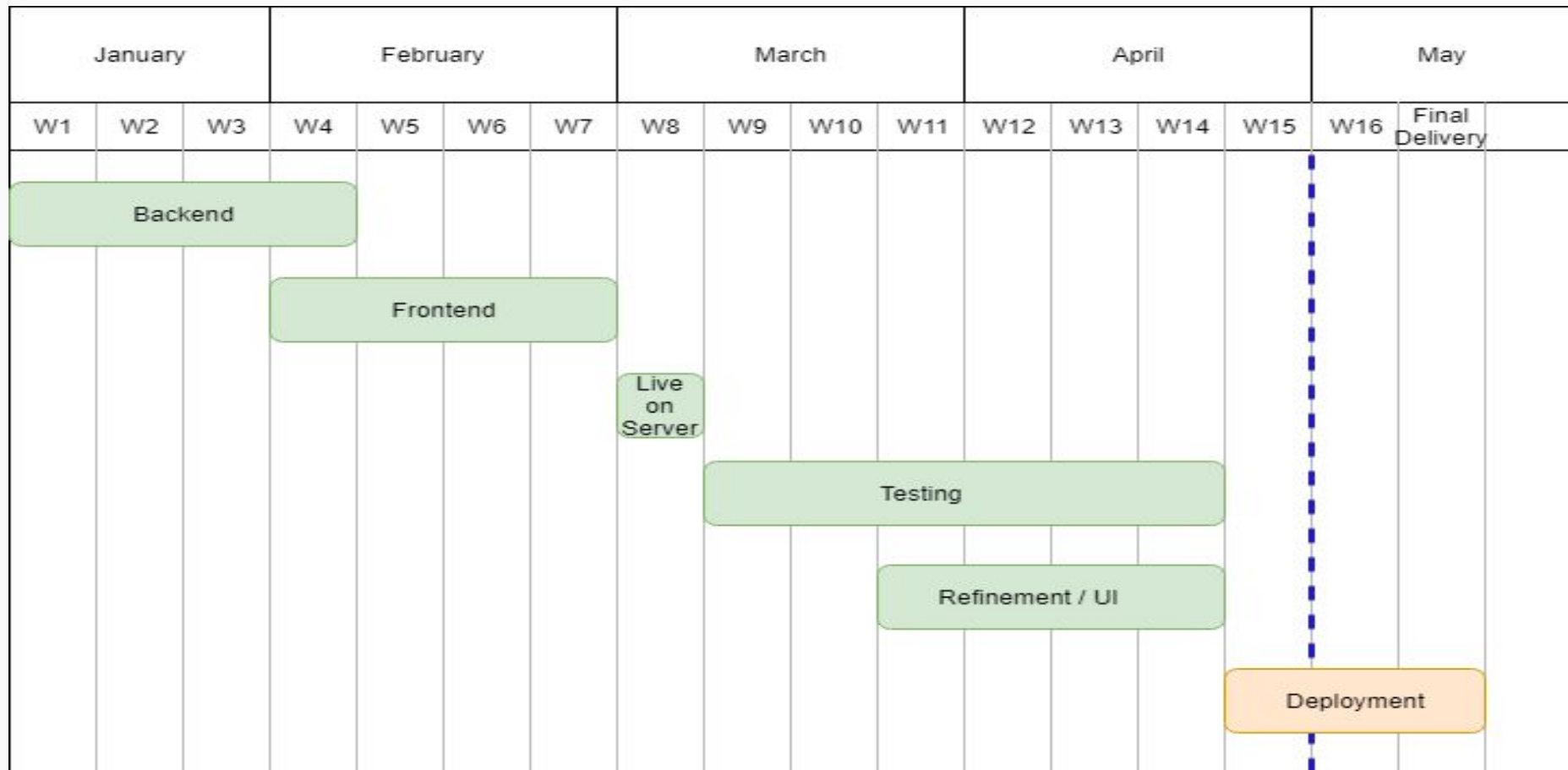


Obstacles

- Speed of data delivery to the user.
 - Increased mining time, decreased display time.
- Mining Data Asynchronously
 - Scheduling and automation

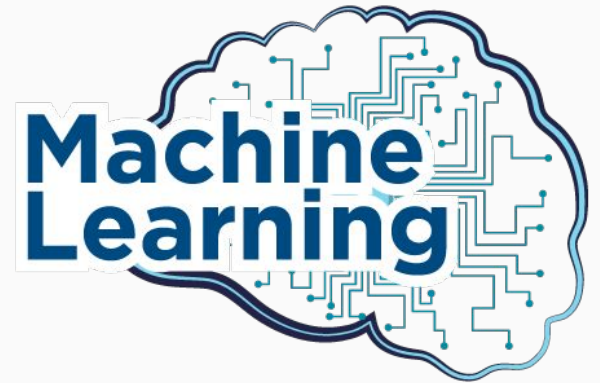


August / September			September				October				November				December			
W1	W2	W3	W4	W5	W6	W7	W8	W9	W10	W11	W12	W13	W14	W15	W16			
			Requirements Acquisition															
									Software Design									
													Testing Technologies					



Future Work

- Sentiment analysis of pull request comments
- Repository activity projection using machine learning
- Ability to track repositories being mined
- Ability to “favorite” repositories of interest



Conclusion

- OSS importance
- Problems
 - Finding a First Project
 - Information Overload
 - Inactivity/No Feedback
 - Unaccepting of New Contributions
- Solution: Redesign OSS interface
- Impact: Make that 10% difference.



Git OSS-um: A Tool for Newcomers to Open Source Software



Team Members

- Gary Baker (Team Lead)
 - gpb27@nau.edu
- Van Steinbrenner
- Stephen White

- Contact Us: gitossum@gmail.com
- Visit Us: tinyurl.com/gitossum
- See the Product: gitossum.com
- Poster in Fieldhouse: 2-4 pm