開源軟體介紹
– 以OPENCV為例

葉信和／Hsin-Ho Yeh
Founder／CEO／Software Engineer @ 信誠金融科技
hsinho.yeh@footprint-ai.com

信誠金融科技股份有限公司
XINCHEN FINTECH CO., LTD.

2023/10

Tintin
Machine Learning Platform for Everyone

Download Slides

http://bit.ly/3QmqsTw

https://github.com/FootprintAI/talks/tree/main/slides

Intro & timeline

- footprint-ai.com (信誠金融科技) is committed to providing

  Software-as-a-Service for serving AI/ML applications, specializing in:

  - Cloud-native architecture, MLOps, Green Tech, Internet-scale Data

    Analysis

- Practical experience in taking an idea (zero) to many products (one)

  - Singapore-based insurtech startup iglooinsure (18M USD/2020 Series A,

    46M USD/2023 Series B).

- Business partner with IBM Taiwan (IBM.tw)

- Joined Nvidia Inception program in 2021.

- Moved into FinTechSpaces in 2023.

About me

- 2020 - Present at 信誠金融科技
  - Tintin: a machine learning platform for everyone
    - https://get-tintin.footprint-ai.com
- 2016 - 2020 at IglooInsure (16M+ in series A+ 2020)
  - Provide digital insurance for e-conomic world
  - Funded in KUL, Headquartered in Singapore
  - First employee/ Engineering Lead / Regional Head/ Chief Engineer
- 2013 - 2016 at Studio Engineering @ hTC
  - Principal Engineer on Cloud Infrastructure Team
- 2009 - 2012 at IIS @ Academia Sinica
  - Computer vision, pattern recognition, and data mining
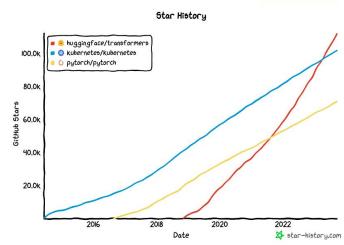- CS@CCU, CS@NCKU alumni

Agenda

- What is Open Source?
- OpenCV Introduction
- How to contribute to Open Source?
- OpenCV Case Study
- FAQ

Huggingface

- An open source data science and machine learning platform, aka github for AI.
- Notable with its open source transformer library, which simplifies the process of working with pre-trained models.
- Huggingface Hub [1]
  - A central hub for community to share and explore machine learning.

Star History

year	2021		2023
Fundraising status	40M (B)		235M(D)

https://en.wikipedia.org/wiki/Hugging_Face
https://github.com/huggingface/transformers
https://collabnix.com/what-is-hugging-face-and-why-it-is-damn-popular

What is open source competitive advantage?

- Open Source (OS) company has gone to public: Hashicorp, JFrog, Elastic, MongoDB and Gitlab
- Key advantages from open source
  - Crowdsourcing product development
    - anyone on the internet can contribute.
  - low-cost sales channel: anyone can download / use the software for free.
    - Trust to the provider is increasing during daily based usage.
    - Acquire a base of free users and offer them with upselling.

https://www.forbes.com/sites/forbesfinancecouncil/2022/06/09/open-source-as-a-competitive-advantage

What is Open Source Software?

"Open-source software is a type of computer software … the copyright holder grants users the rights to **use**, **study**, **change**, and **distribute** the software to anyone and for any purpose." [1]

[1] https://en.wikipedia.org/wiki/Open-source_software

What is Open Source Software? (1/4)

	Open Source Software	Proprietary Software
Ownership	Everyone	A group of limited people
Usability	Difficult use for unskilled users Ex: Desktop Linux	Feature neat to end-user Ex: Microsoft Windows
Security	Better (? Ex: PHP source code is injected backdoors by hackers	Less
Stability	Less	Better
Opacity	Transparent	Closed

Ref: http://www.optimusinfo.com/downloads/white-paper/open-source-vs-proprietary-software-pros-and-cons.pdf

What is Open Source Software? (2/4)

Misunderstanding on open source software

- Open source software is **Free** in terms of Freedom, not Monetary.
- Open source developer is **NOT** a free labor [1]
    - Sponsorship by companies
    - Add values(a.k.a layers) on the existing project
    - Professional consult
- Open source means **Pool** quality
    - Good quality when the project gained enough eyeballs
    - On average the density of code defect is smaller than proprietary software [2]

[1] https://softwareengineering.stackexchange.com/questions/100685/making-money-with-open-source-as-a-developer
[2] https://blogs.worldbank.org/opendata/quality-open-source-software-how-many-eyes-are-enough

What is Open Source Software? (3/4)

Where open source communities lives?

- Github
  - Free for any public repositories
  - Github Sponsor

Become a sponsor to **OpenCV**

OpenCV
opencv

OpenCV is the largest Computer Vision library in the world. It's used worldwide: in the world's biggest companies, underwater, and in space.

OpenCV was started in 1998 and played the key role in the development of Computer Vision and Artificial Intelligence. It is released under Apache 2.0 license and is open-source and free for everyone.

In the sponsorship program, our main focus is direct support of OpenCV core developers. This means more new functionality, better support, and faster code review for the pull requests from our huge community.

Support OpenCV today to help us develop Computer Vision industry worldwide.

Current sponsors 3

Past sponsors 5
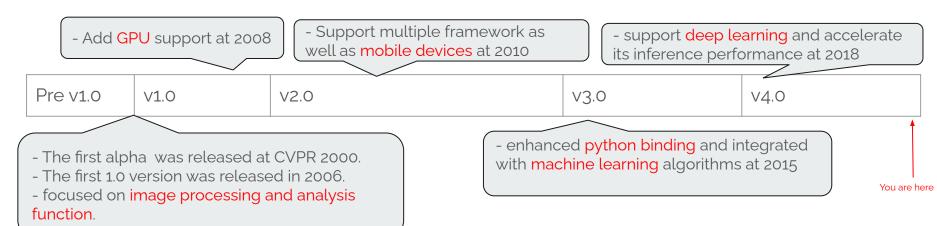
What is Open Source Software? (4/4)

Key files to understand a github project
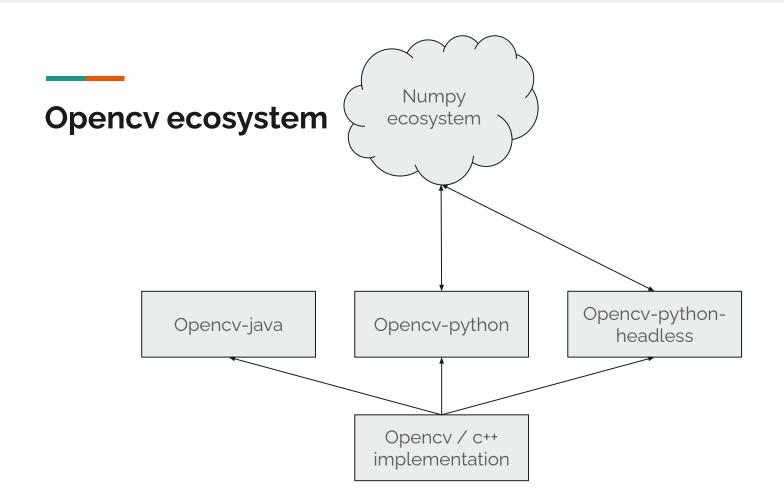
- README.md
  - How to build and use the project
- LICENSE
  - Terms and Conditions
- CONTRIBUTING.md
  - Rules on contributing

Ref: https://www.slideshare.net/atwong/how-to-contribute-to-open-source-59292026

OpenCV introduction

Opencv history

OpenCV (Open Source Computer Vision Library) is an open-source computer vision and machine learning software library.

- Add GPU support at 2008

- Support multiple framework as well as mobile devices at 2010

- support deep learning and accelerate its inference performance at 2018

Pre v1.0	v1.0	v2.0	v3.0	v4.0

- The first alpha was released at CVPR 2000.
- The first 1.0 version was released in 2006.
- focused on image processing and analysis function.

- enhanced python binding and integrated with machine learning algorithms at 2015

You are here

https://computervisionproject.com/opencv/intro/opencvbackground.html

14

Opencv ecosystem

Numpy ecosystem

Opencv-java

Opencv-python

Opencv-python-headless

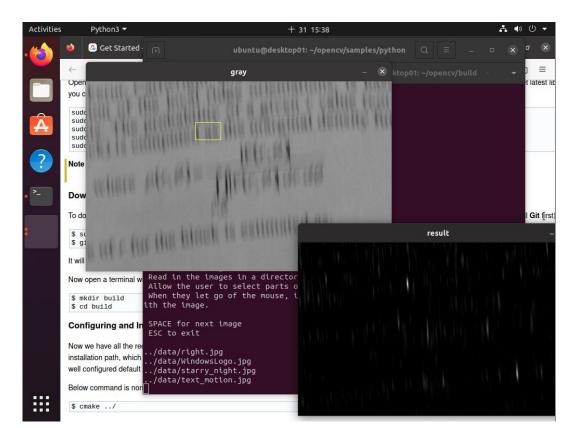Opencv / c++ implementation

Opencv installation

```
// pre-build binaries
https://github.com/opencv/opencv/releases/tag/4.8.1

// python pypi (python package index)
pip3 install opencv-python
```
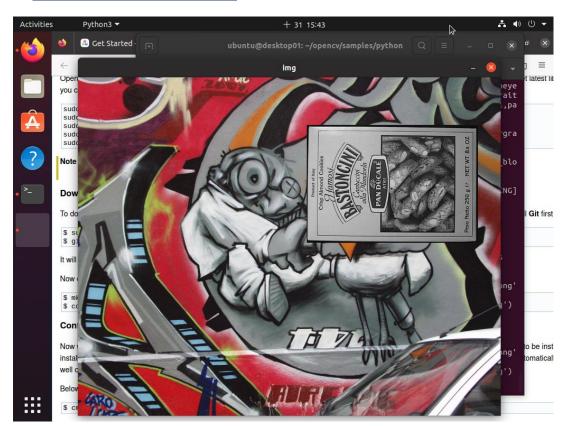
Opencv: installation from source code (advanced)

- Build with cpp
  - Linux: https://docs.opencv.org/master/d7/d9f/tutorial_linux_install.html
  - Windows: https://docs.opencv.org/master/d3/d52/tutorial_windows_install.html
- Build with python
  - Linux: https://docs.opencv.org/4.x/d2/de6/tutorial_py_setup_in_ubuntu.html
  - Windows: https://docs.opencv.org/4.x/d5/de5/tutorial_py_setup_in_windows.html
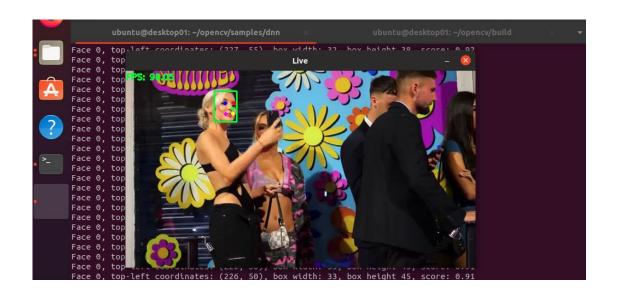- Took ~~several minutes~~ half hours to build code

Example I mouse_and_match

Example II scene_render
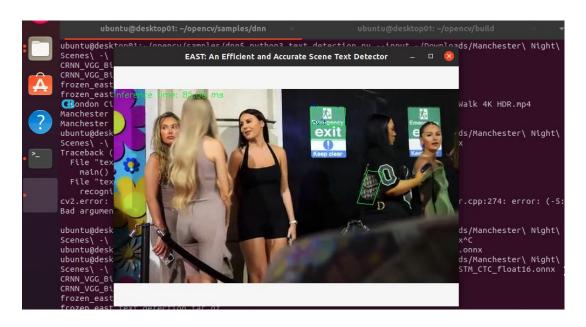
Example III [face detector](#)

Video source: https://www.youtube.com/watch?v=ufPn6RaDxbg

Example III text detector

Video source: https://www.youtube.com/watch?v=ufPn6RaDxbg
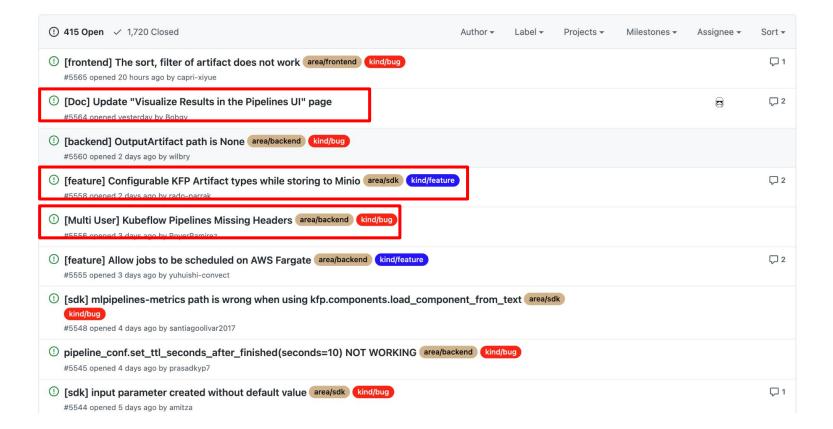
How to contribute Open Source?

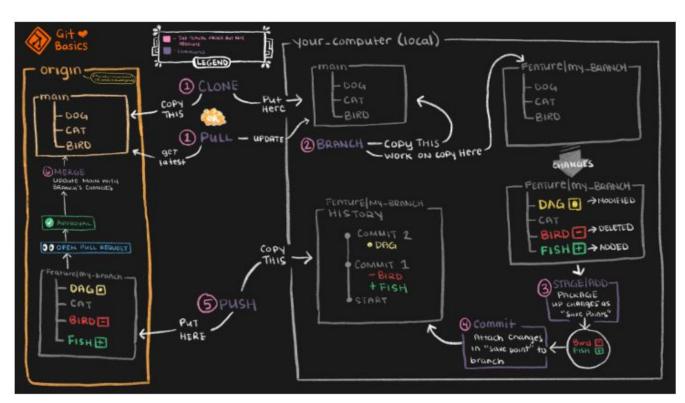# How to contribute to open source?
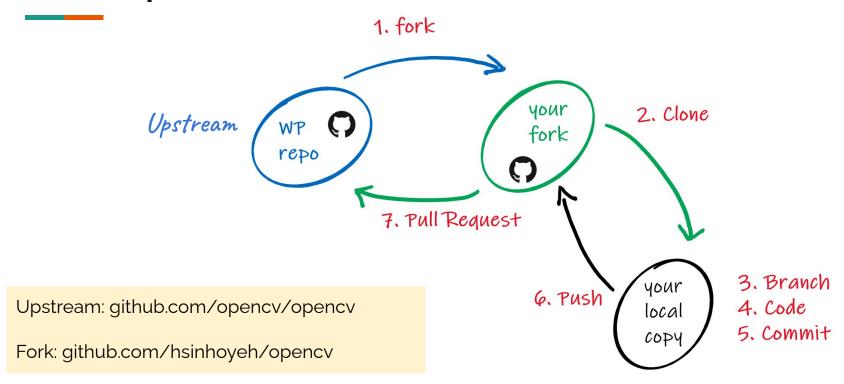
The two most popular way to contribute

1. Raising an issue where you think the project could be improved
   a. Report a problem/bug (kind/bug)
   b. Suggest new features (kind/feature)
   c. Gap in documentation (doc)
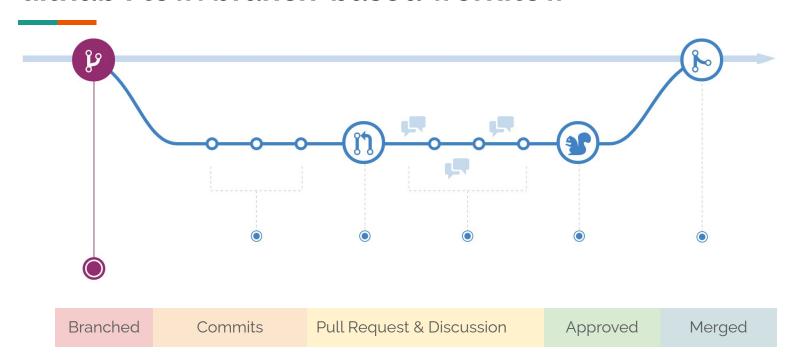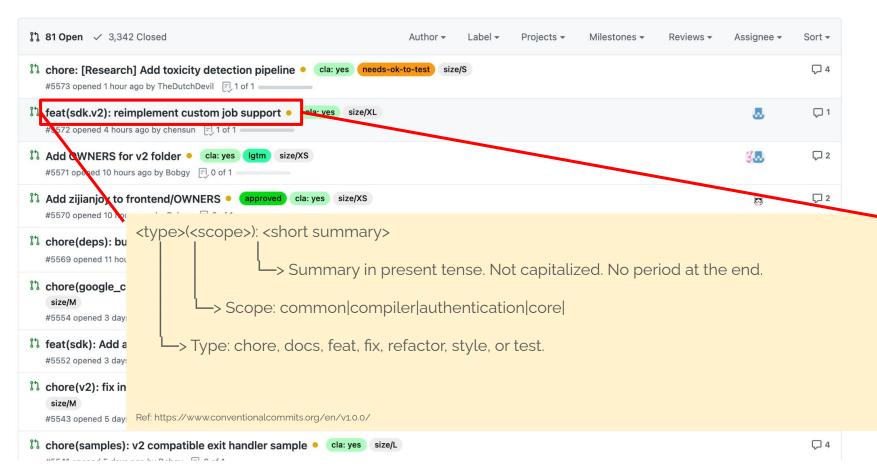2. Contribute to code
   a. Pull-Request in Github

⚠ **415 Open**   ✓ 1,720 Closed   Author ⌄   Label ⌄   Projects ⌄   Milestones ⌄   Assignee ⌄   Sort ⌄

⚠ **[frontend] The sort, filter of artifact does not work** area/frontend kind/bug   💬 1
#5565 opened 20 hours ago by capri-xiyue

⚠ **[Doc] Update "Visualize Results in the Pipelines UI" page**   😶   💬 2
#5564 opened yesterday by Bobgy

⚠ **[backend] OutputArtifact path is None** area/backend kind/bug
#5560 opened 2 days ago by wilbry

⚠ **[feature] Configurable KFP Artifact types while storing to Minio** area/sdk kind/feature   💬 2
#5558 opened 2 days ago by rado-parrak

⚠ **[Multi User] Kubeflow Pipelines Missing Headers** area/backend kind/bug
#5556 opened 3 days ago by RoverRamirez

⚠ **[feature] Allow jobs to be scheduled on AWS Fargate** area/backend kind/feature   💬 2
#5555 opened 3 days ago by yuhuishi-convect

⚠ **[sdk] mlpipelines-metrics path is wrong when using kfp.components.load_component_from_text** area/sdk
kind/bug
#5548 opened 4 days ago by santiagoolivar2017

⚠ **pipeline_conf.set_ttl_seconds_after_finished(seconds=10) NOT WORKING** area/backend kind/bug
#5545 opened 4 days ago by prasadkyp7

⚠ **[sdk] input parameter created without default value** area/sdk kind/bug   💬 1
#5544 opened 5 days ago by amitza

Git Basic

Pull-Request flow in Github

1. fork

Upstream

WP repo

your fork

2. Clone

7. Pull Request

6. Push

your local copy

3. Branch
4. Code
5. Commit
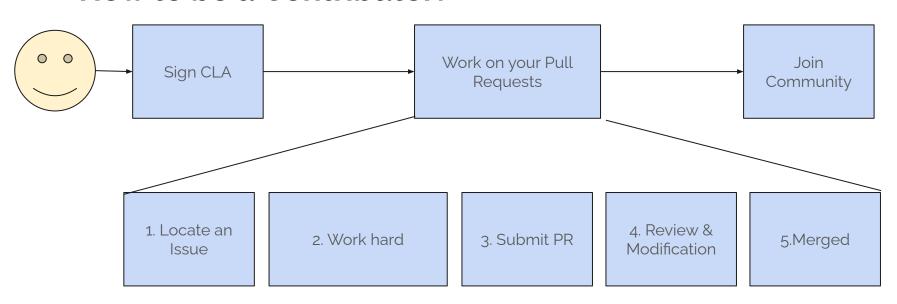
Upstream: github.com/opencv/opencv

Fork: github.com/hsinhoyeh/opencv

Ref: https://developer.wordpress.org/block-editor/contributors/code/git-workflow/
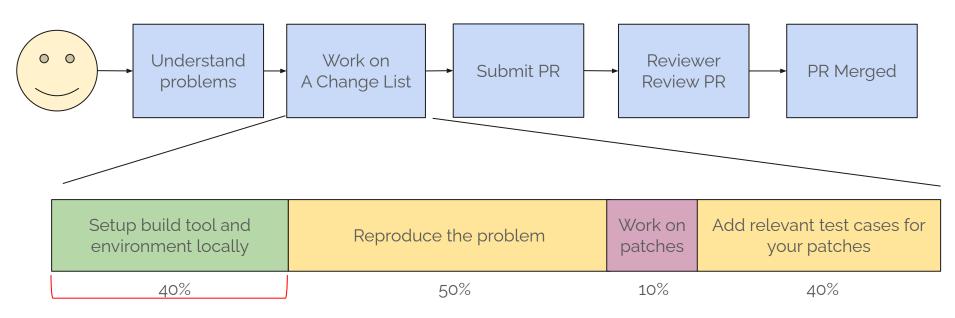
26

Github Flow: branch-based workflow

| Branched | Commits | Pull Request & Discussion | Approved | Merged |

Ref: https://guides.github.com/introduction/flow/

81 Open ✓ **3,342 Closed**   Author ▾   Label ▾   Projects ▾   Milestones ▾   Reviews ▾   Assignee ▾   Sort ▾

chore: [Research] Add toxicity detection pipeline ● `cla: yes` `needs-ok-to-test` `size/S`   💬 4
#5573 opened 1 hour ago by TheDutchDevil 1 of 1

feat(sdk.v2): reimplement custom job support ● `cla: yes` `size/XL`   💬 1
#5572 opened 4 hours ago by chensun 1 of 1

Add OWNERS for v2 folder ● `cla: yes` `lgtm` `size/XS`   💬 2
#5571 opened 10 hours ago by Bobgy 0 of 1

Add zijianjoy to frontend/OWNERS ● `approved` `cla: yes` `size/XS`   💬 2
#5570 opened 10 ho

chore(deps): bu
#5569 opened 11 hou

chore(google_c
`size/M`
#5554 opened 3 days

feat(sdk): Add a
#5552 opened 3 days

chore(v2): fix in
`size/M`
#5543 opened 5 days

chore(samples): v2 compatible exit handler sample ● `cla: yes` `size/L`   💬 4

```
<type>(<scope>): <short summary>
          │        │         │
          │        │         └─> Summary in present tense. Not capitalized. No period at the end.
          │        │
          │        └─> Scope: common|compiler|authentication|core|
          │
          └─> Type: chore, docs, feat, fix, refactor, style, or test.
```

Ref: https://www.conventionalcommits.org/en/v1.0.0/

How to be a Contributor?

Sign CLA	Work on your Pull Requests	Join Community

1. Locate an Issue	2. Work hard	3. Submit PR	4. Review & Modification	5.Merged

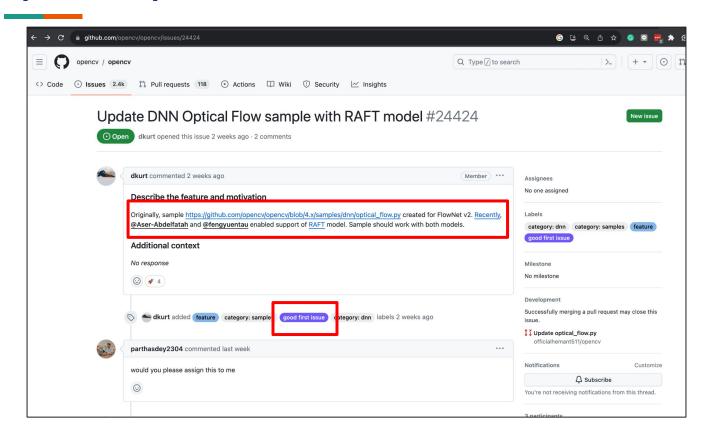Ref: https://www.kubeflow.org/docs/about/contributing/

29

Find your first issue

1. Existing issues
   a. Issues labelled with `good first issue`: opencv/opencv
   b. Issues labelled with `help wanted`: opencv/opencv
2. Organic issues
   a. Create an issue to describe issues you encountered and also provide a solution (a.k.a Pull-request) for it

Why is the initial contribution significant?

OpenCV Case Study

opencv/opencv#24424
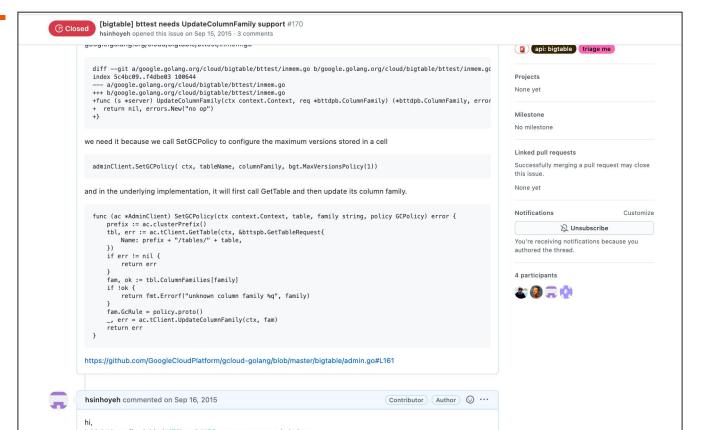
opencv/opencv#24451

[opencv/opencv#24446](https://github.com/opencv/opencv/pull/24446)

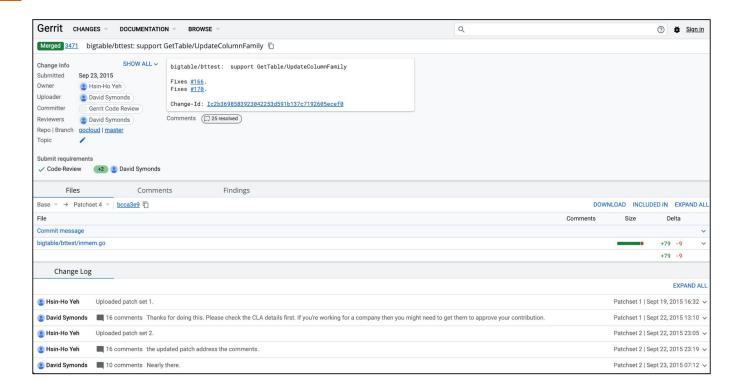# Other Case Study

nats-io/nats.go#22 Case Study

[googleapis/google-cloud-go#170](#) Case Study (1/2)

[googleapis/google-cloud-go#170](#) Case Study (2/2)

[istio/istio#46327](https://github.com/istio/istio/issues/46327) Case Study
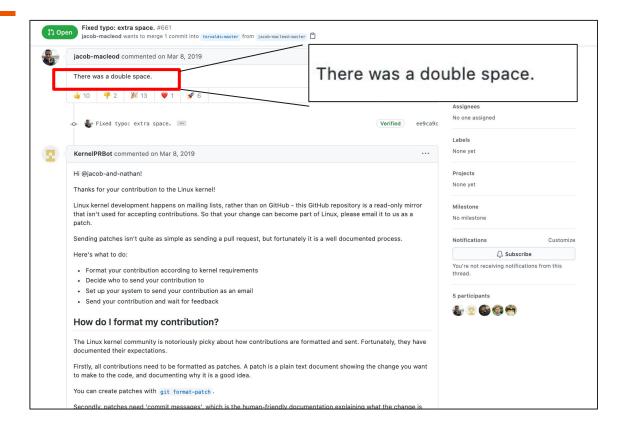

天下武功，唯快不破

[torvalds/linux#661](#) Case Study - Not all CLs are valuable

Conclusion

- Open source is a public artifacts that help you grow a reputation and a career.
- The work to contribute to Open Source could take days or weeks, but the rewards could keep years.
- Let's make changes to the software world now, even steps are small :)

Thank You

Questions?

Contact Us:
partnership@footprint-ai.com