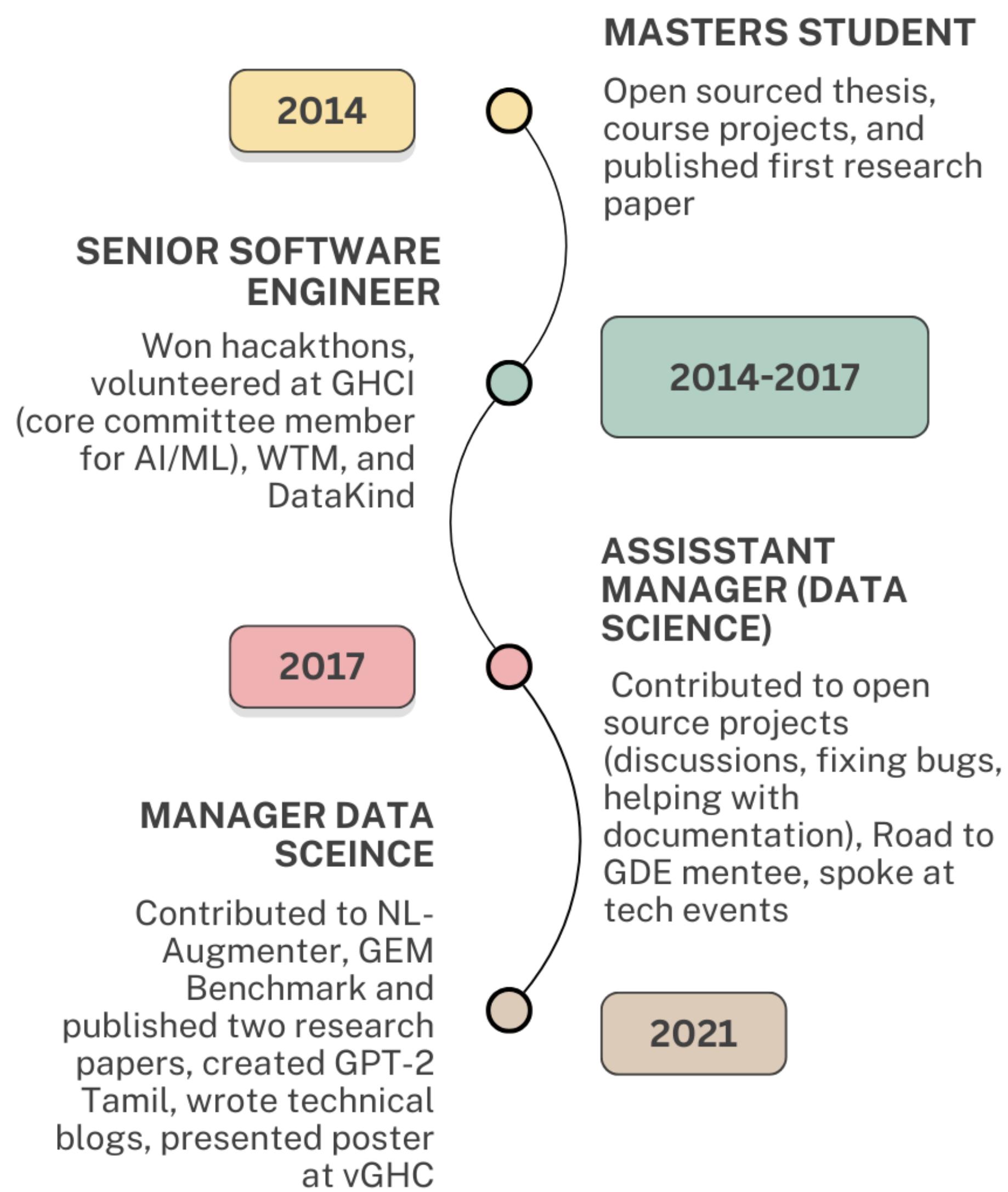


# **Contributing to open source**

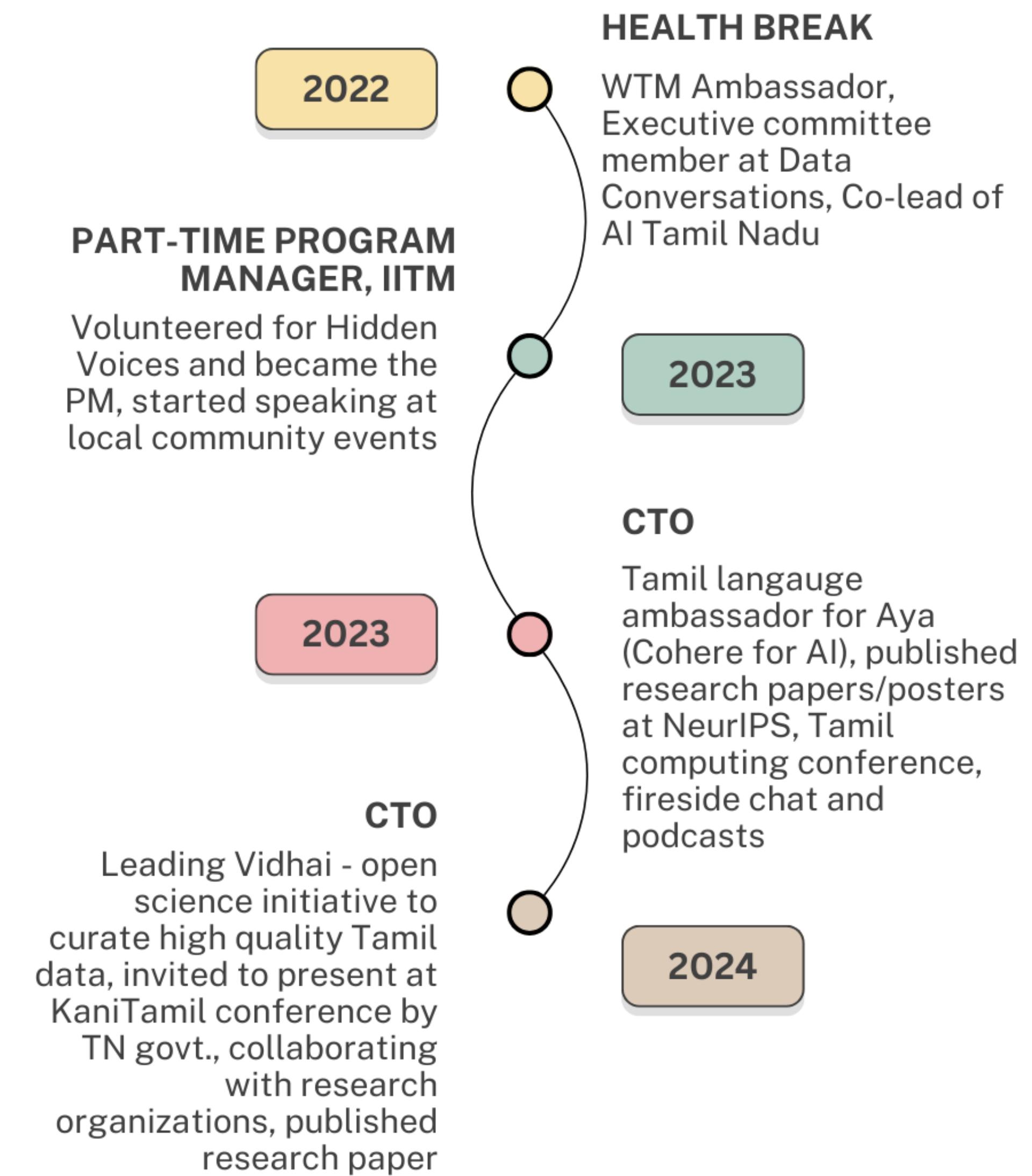
**Improving skills and increasing career opportunities the right way!**

**Abinaya Mahendiran  
AI Days 2024 (AI For Society)  
March 31**

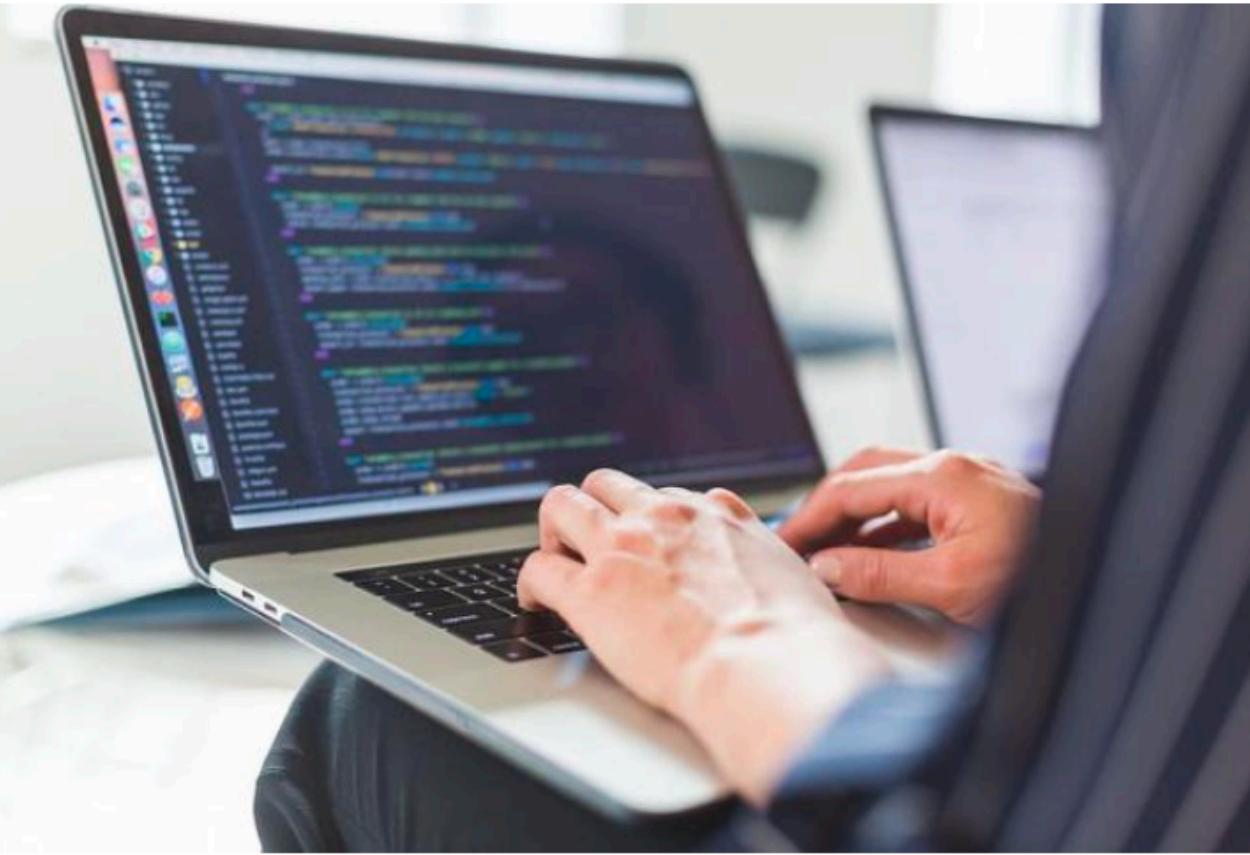
# CAREER JOURNEY



# CAREER JOURNEY

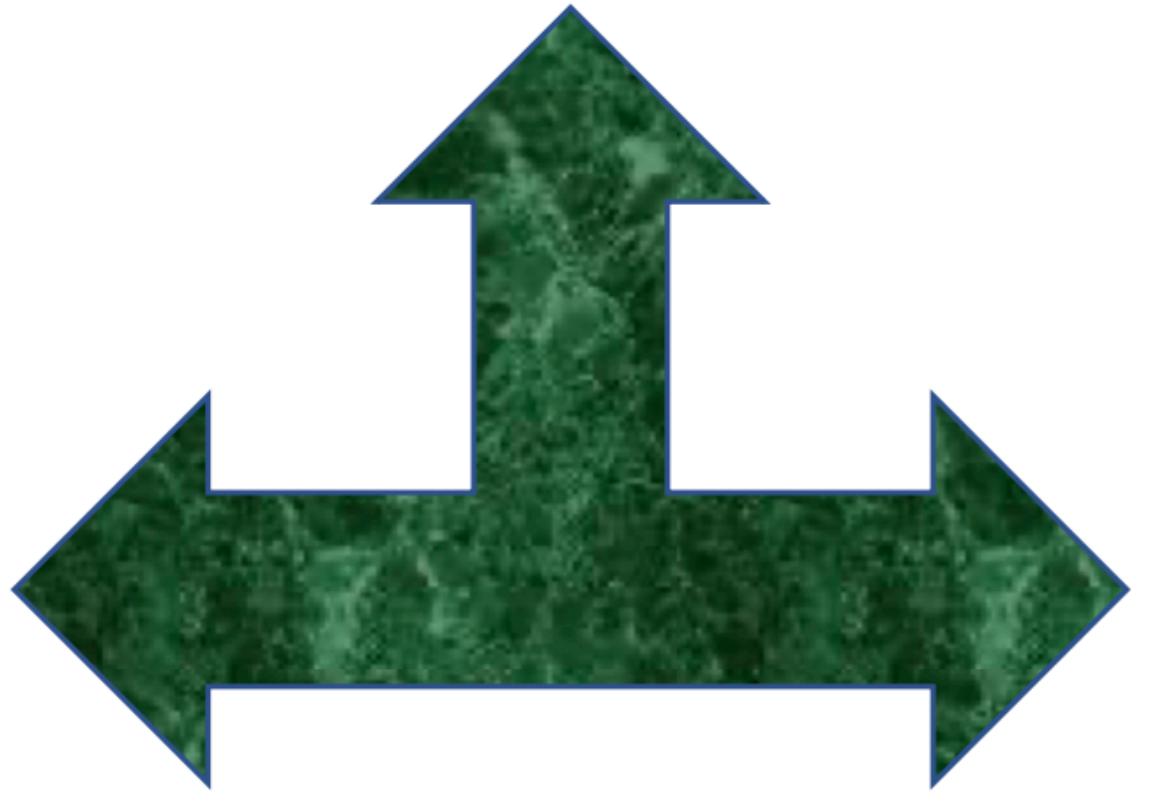


# The Why's?



Enhance your skills (technical/soft)

Improve the quality of the software



Meet like-minded people (network)

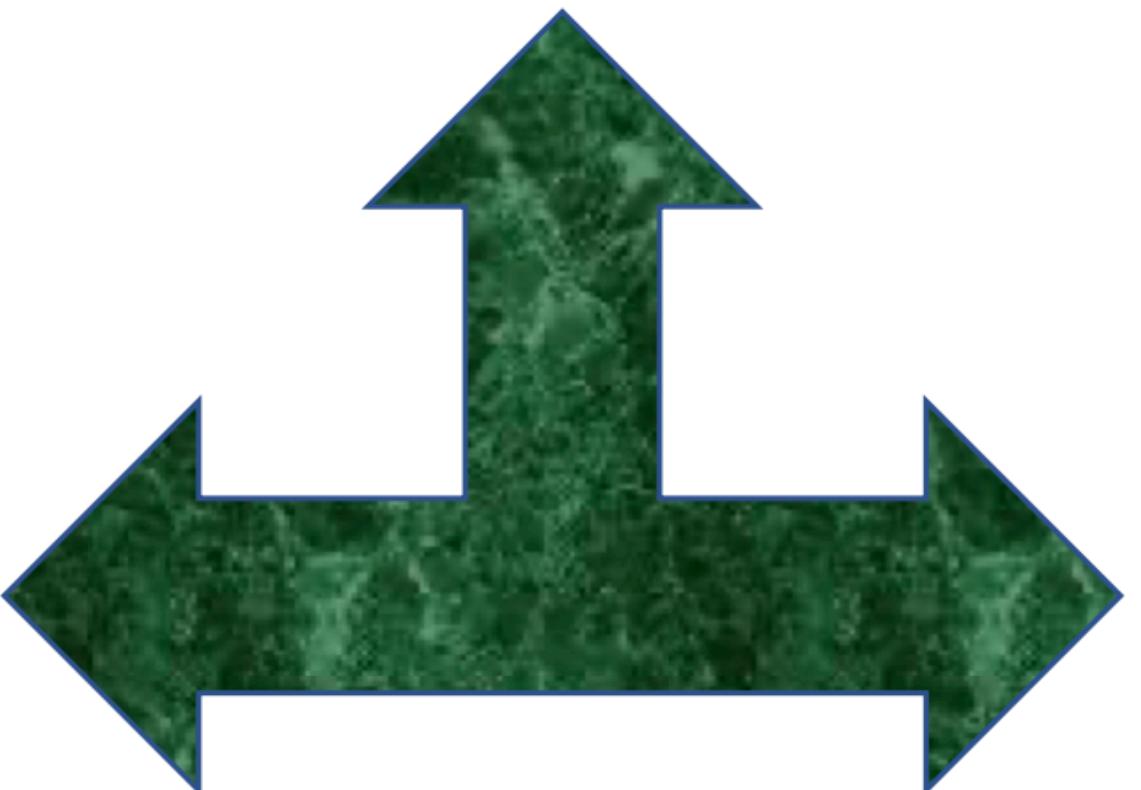


# The Why's?



Find (become) a mentor or sponsor

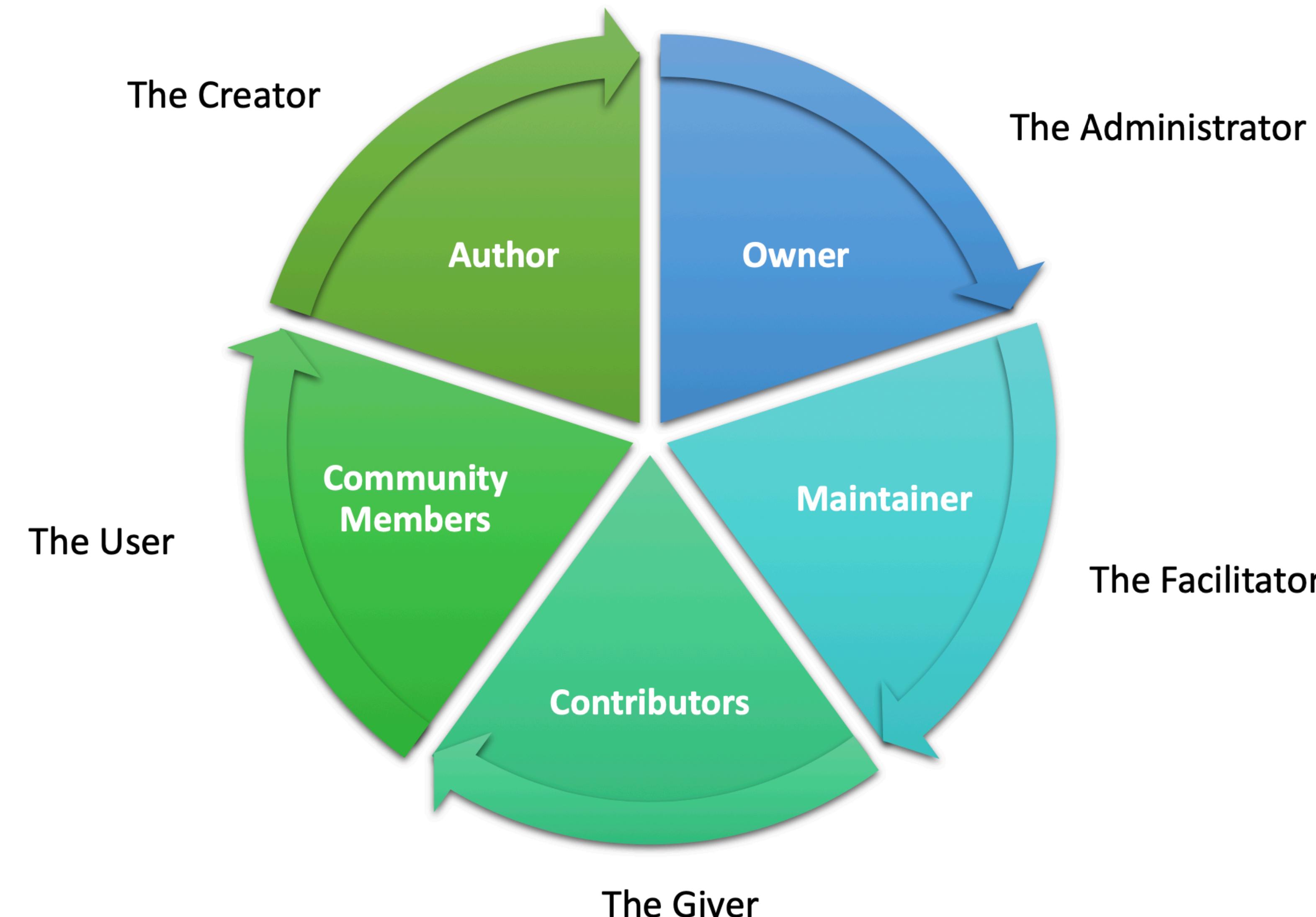
Build a brand / career



Learn people skills



# Anatomy of Projects : The People



# Anatomy of Projects: The Artefacts

License

README

Contributing  
Guide

Code of  
Conduct

Other  
Documentation

# Anatomy of Projects: The Artefacts

keras-team / keras Public

Code Issues 284 Pull requests 116 Actions Projects 1 Wiki Security Insights

master 166 branches 75 tags Go to file Add file Code

 nkovela1 and tensorflow-gardener Expands Keras internal testing coverage... e7c4d89 12 hours ago 8,111 commits

.devcontainer	Use the official vscode Python image as the base image for devcontain...	4 months ago
.github	Move stale management to github action	last month
.vscode	Use the official vscode Python image as the base image for devcontain...	4 months ago
keras	Expands Keras internal testing coverage for the new v3 saving format ...	12 hours ago
shell	Make OSS scripts to be executable.	last year
third_party	Internal change	2 years ago
.bazelrc	Internal Code Change	last month
.bazelversion	Pin the bazel version used by keras OSS build.	4 months ago
.gitignore	gitignore .DS_Store	last year
BUILD	Fix circular import error between tensorflow_io and keras.	last month
CONTRIBUTING.md	Add <code>-oss_excluded</code> to TF build/test tag filters	last month
ISSUE_TEMPLATE.md	Add issue template for keras-team/keras.	2 years ago
LICENSE	Change the <code>LICENSE</code> file to be a verbatim copy of the complete Apach...	last year
README.md	Update the Keras README	2 months ago
WORKSPACE	feat: Update protobuf version to match latest TF protobuf update	3 months ago

About Deep Learning for humans  
keras.io/  
python data-science machine-learning  
deep-learning tensorflow  
neural-networks

Readme Apache-2.0 license 58.2k stars 1.9k watching 19.3k forks

Report repository

Releases 47  
Keras Release 2.12.0 Latest on Mar 22 + 46 releases

Packages No packages published

# Anatomy of Projects: The Tools

- **Issue Tracker – Issues, bugs, features, enhancements**

The screenshot shows a dark-themed issue tracker interface with the following details:

- Filters:** is:issue is:open
- Labels:** 30
- Milestones:** 0
- New issue** button
- Open Issues Summary:** 284 Open, 11,415 Closed
- Issues List:** A list of 7 open issues, each with a green circular icon, a title, a description, and a timestamp.

Author	Label	Projects	Milestones	Assignee	Sort

**Issues:**

- steps\_per\_execution for validation data** `type:feature`  
#18070 opened 1 hour ago by iussef
- Spurious results with `model.compile(jit_compile=True)` and explicit XLA compilation of the model**  
#18062 opened yesterday by sayakpaul
- Conv1DTranspose "groups" argument** `keras-team-review-pending` `type:feature`  
#18051 opened 4 days ago by WaterBoiledPizza
- Keras should offer a Callback to Pick the Best model weights** `type:feature`  
#18046 opened last week by tim-janik
- print in dataset\_utils instead of io\_utils.print\_msg** `type:support`  
#18041 opened last week by pedrobrs
- data cardinality error** `type:Bug`  
#18013 opened last week by deepquantum88

# Anatomy of Projects: The Tools

- **Pull Requests – Tested code for an issue and reviews**

The screenshot shows a list of pull requests on a GitHub interface. Each PR is represented by a card with the following details:

- PR #18036: Update indentation level - activations.py** (Status: ✗)  
#18060 opened 2 days ago by copybara-service bot
- PR #18058: Remove github user sushreebarsa from assignees.** (Status: ✗) ready to pull size:XS  
#18058 opened 2 days ago by shmishra99 • Approved
- PR #18054: use io\_utils.print\_msg in dataset\_utils** (Status: ✗) size:XS  
#18054 opened 4 days ago by pedrobrs • Changes requested
- PR #18052: Add ignore nan metric loss** (Status: ✓) size:M  
#18052 opened 4 days ago by sagi-ezri
- PR #17979: [keras/testing\_infra] Standardise docstring usage of "Default to"** (Status: ✗)  
#18047 opened 5 days ago by copybara-service bot
- PR #18045: Update example losses to bce- metrics/confusion\_metrics.py** (Status: ✓) size:S  
#18045 opened last week by Frightera
- PR #18043: Update indentation level - losses.py** (Status: ✗) ready to pull size:XL  
#18043 opened last week by Frightera
- PR #18042: Improve error message for input data to fit.** (Status: ✓) size:S  
#18042 opened last week by tomrtk • Approved

# Anatomy of Projects: The Tools

- **Mailing Lists** – Group conversations, updates and important information
- **Discussion Forums** – Get help, clear doubts, and feature requests
- **Synchronous chat channel** – Casual conversations, collaboration and quick idea exchanges

# The How's: Finding Projects

- Pick a project that you already use!
- Search for tags
- Check if the project proactively accepts contributions
- See if the community is active and welcoming

**good-first-issues**

**help-wanted**

**beginner-friendly**



# The How's: Finding Projects



- [GitHub Explore](#)
- [First Timers Only](#)
- [CodeTriage](#)
- [24 Pull Requests](#)
- [Up For Grabs](#)
- [Contributor-ninja](#)
- [First Contributions](#)
- [Hacktoberfest](#)



# The How's: Contributing to the project

- Get familiar with codebase
- Start small
- Follow the project guidelines
- Make meaningful contributions
- Ask for help and be patient
- Communicate effectively
- Do your homework
- Keep your communication public
- Respect community decisions
- Spread positivity

# The Ways!

» Code examples / Natural Language Processing

## Natural Language Processing

Text classification

- ★ Text classification from scratch
- Review Classification using Active Learning
- Text Classification using FNet
- Large-scale multi-label text classification
- Text classification with Transformer
- Text classification with Switch Transformer
- Text classification using Decision Forests and pretrained embeddings
- Using pre-trained word embeddings
- Bidirectional LSTM on IMDB

## Tutorials / Blogs

Python Enhancement Proposals | Python » PEP Index » PEP 8

### Contents

- Introduction
- A Foolish Consistency is the Hobgoblin of Little Minds
- Code Lay-out
  - Indentation
  - Tabs or Spaces?
  - Maximum Line Length
  - Should a Line Break Before or After a Binary Operator?
  - Blank Lines
  - Source File Encoding
  - Imports
  - Module Level Dunder Names
  - String Quotes
  - Whitespace in Expressions and Statements
  - Pet Peeves
  - Other Recommendations

### PEP 8 – Style Guide for Python Code

**Author:** Guido van Rossum <guido at python.org>, Barry Warsaw <barry at python.org>, Nick Coghlan <ncoghlan at gmail.com>

**Status:** Active  
**Type:** Process  
**Created:** 05-Jul-2001  
**Post-History:** 05-Jul-2001, 01-Aug-2013

▶ Table of Contents

#### Introduction

This document gives coding conventions for the Python code comprising the standard library in the main Python distribution. Please see the companion informational PEP describing [style guidelines for the C code in the C implementation of Python](#).

This document and [PEP 257](#) (Docstring Conventions) were adapted from Guido's original Python Style Guide essay, with some additions from Barry's style guide [2].

## Events near Coimbatore, IN

Events Groups

## Meetups / Technical Events



Serverless Apache pulsar with Astra Streaming

Pulsar User group • Bangalore, IN

New Group

12 attendees



SAT, MAY 13 · 5:00 PM IST

CBEStartups Monthly Meetup

Coimbatore Start-up Meetup • Coimbatore, IN

22 attendees

## Documentation

# The Ways!

```
import os

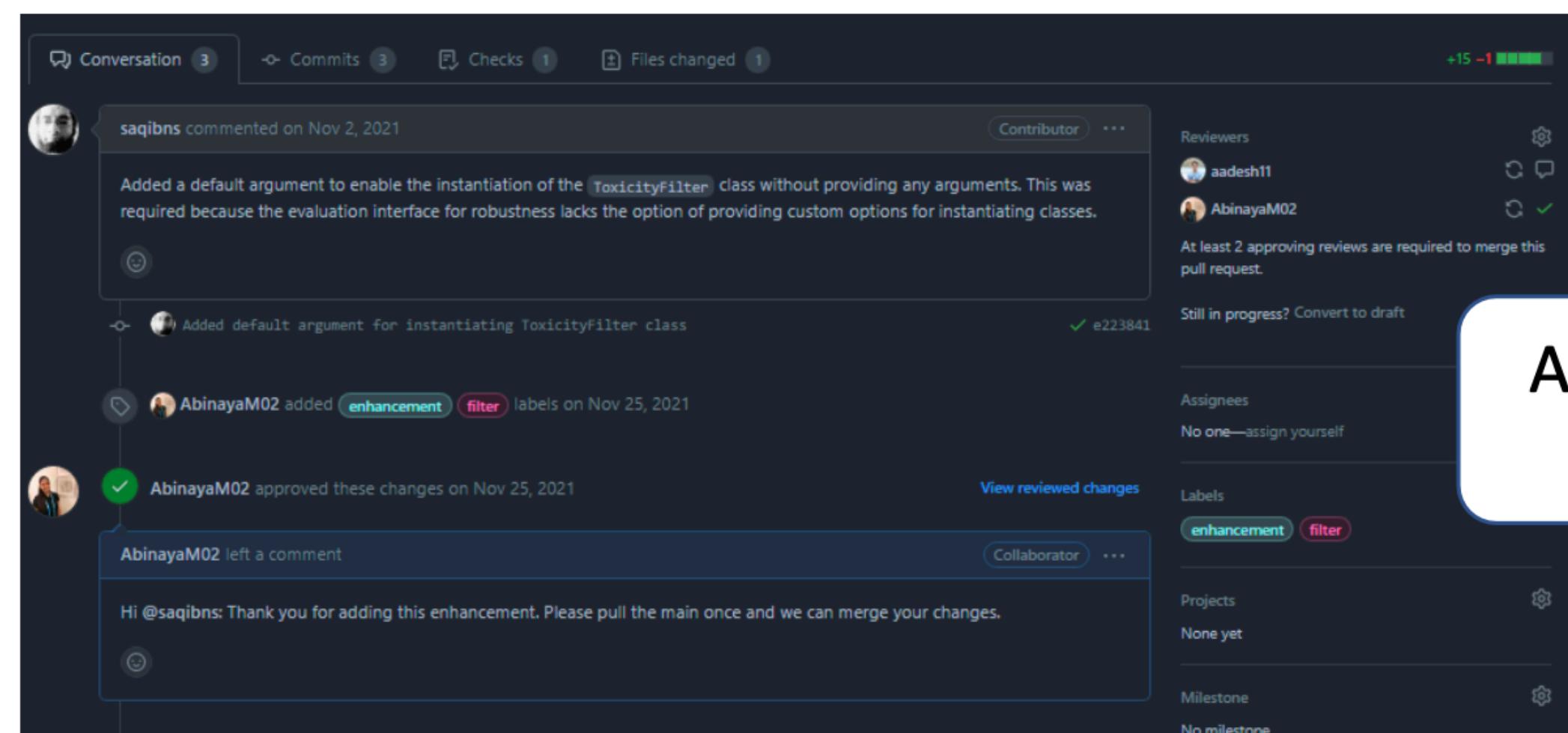
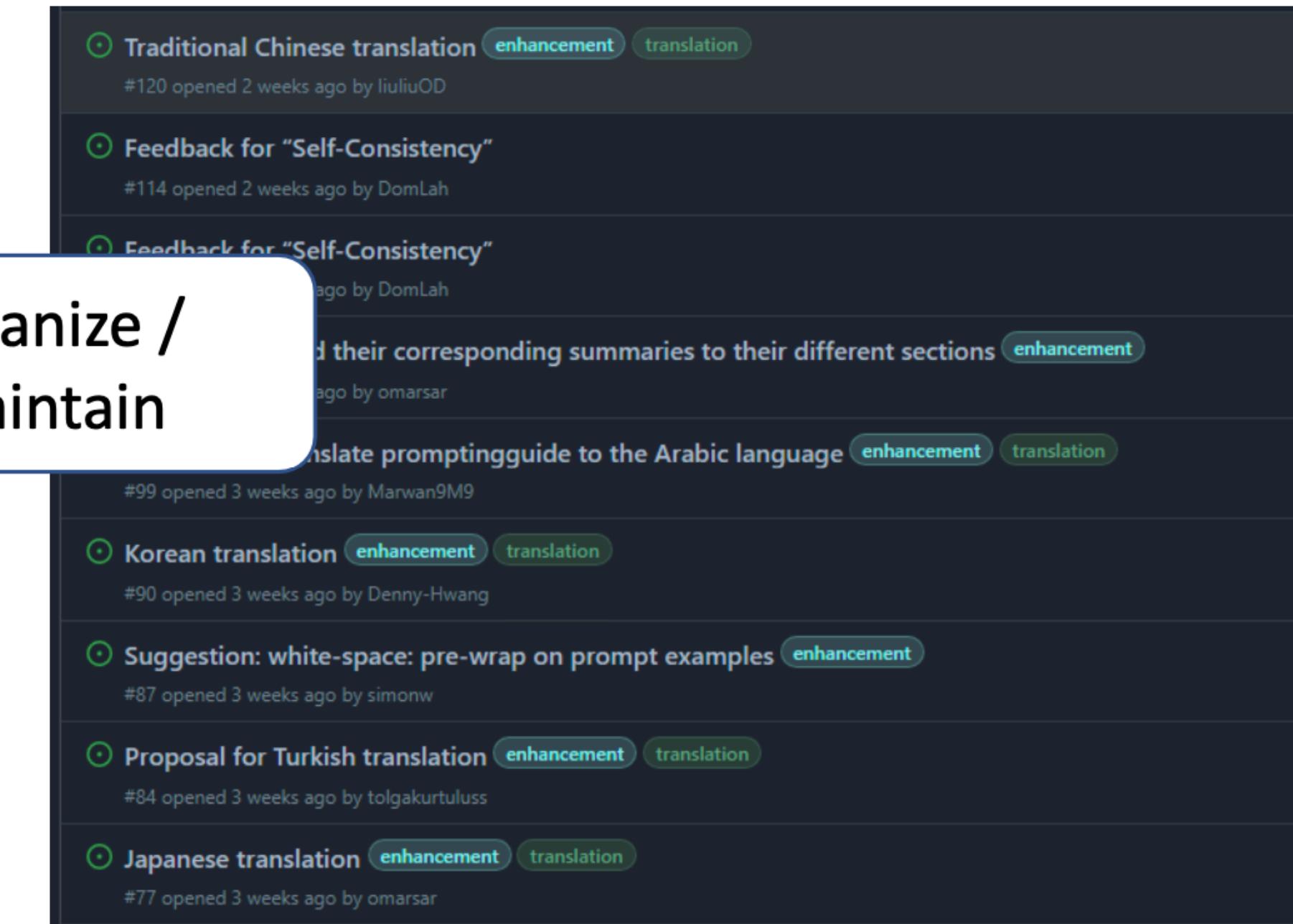
num_skipped = 0
for folder_name in ("Cat", "Dog"):
    folder_path = os.path.join("PetImages", folder_name)
    for fname in os.listdir(folder_path):
        fpath = os.path.join(folder_path, fname)
        try:
            fobj = open(fpath, "rb")
            is_jfif = tf.compat.as_bytes("JFIF") in fobj.peek(10)
        finally:
            fobj.close()

        if not is_jfif:
            num_skipped += 1
            # Delete corrupted image
            os.remove(fpath)

print("Deleted %d images" % num_skipped)
```

Code / Code  
Reviews

Organize /  
Maintain

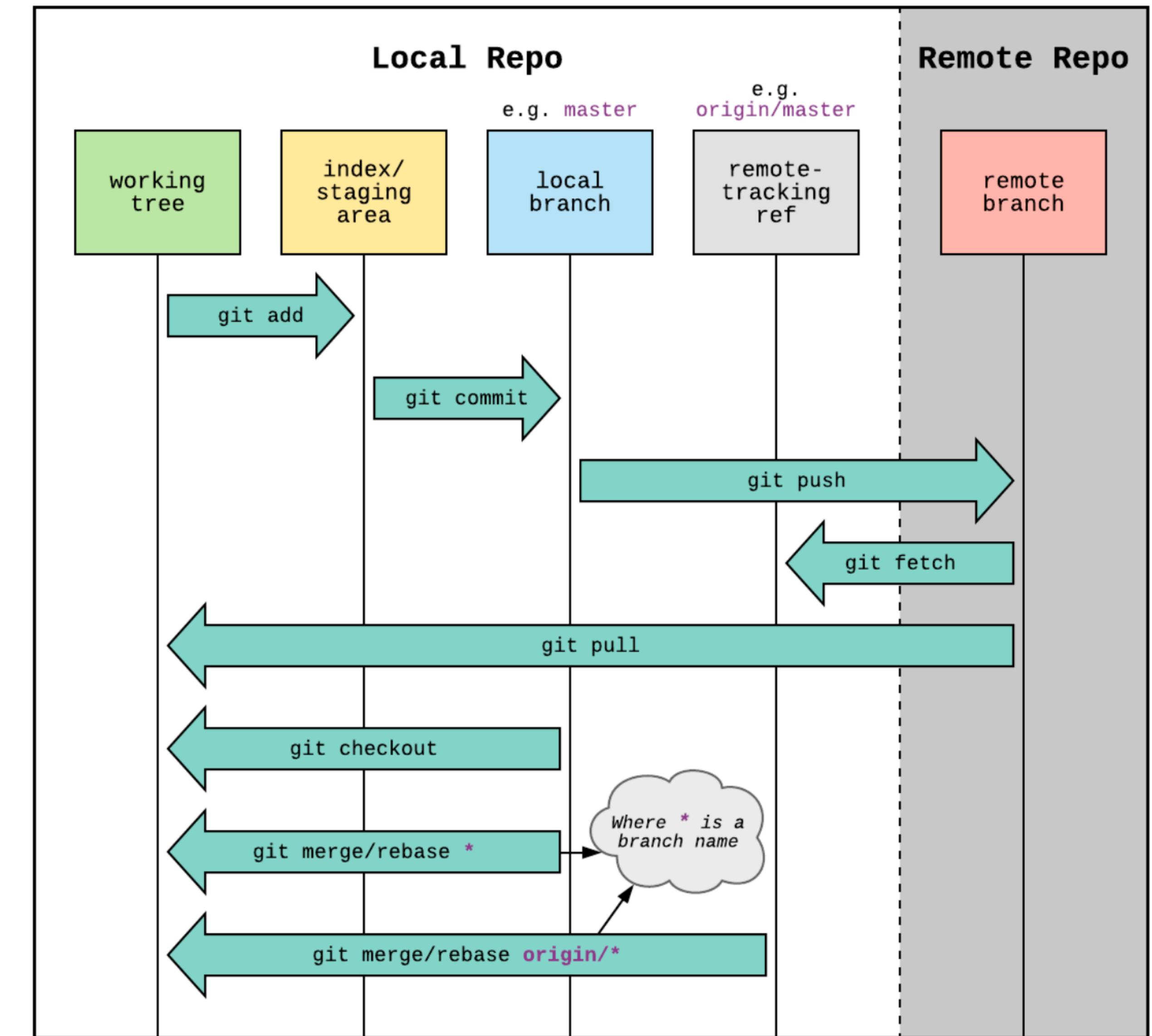
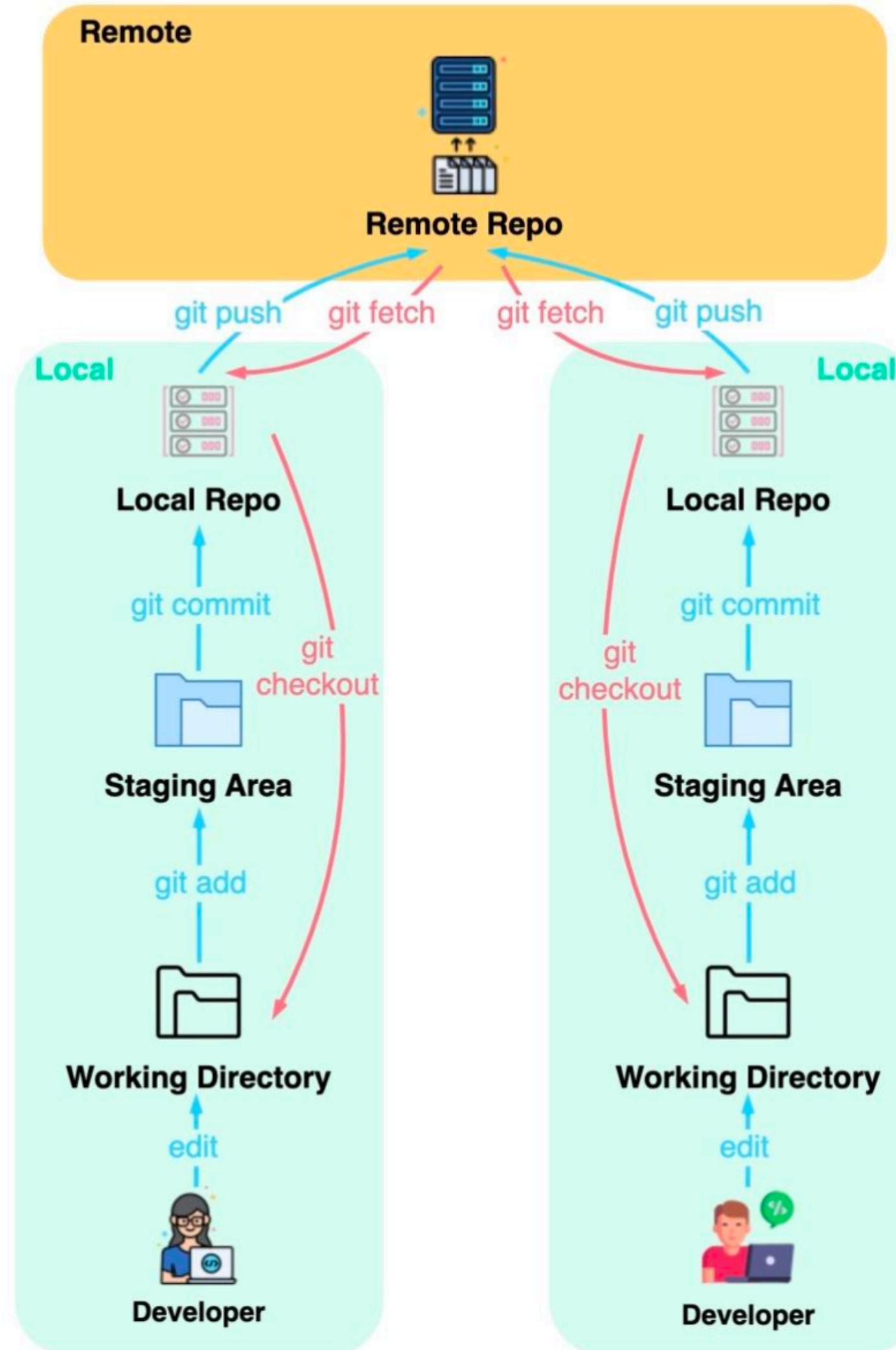


Answer / Clear  
Doubts

# Git, Please!

How does Git Work?

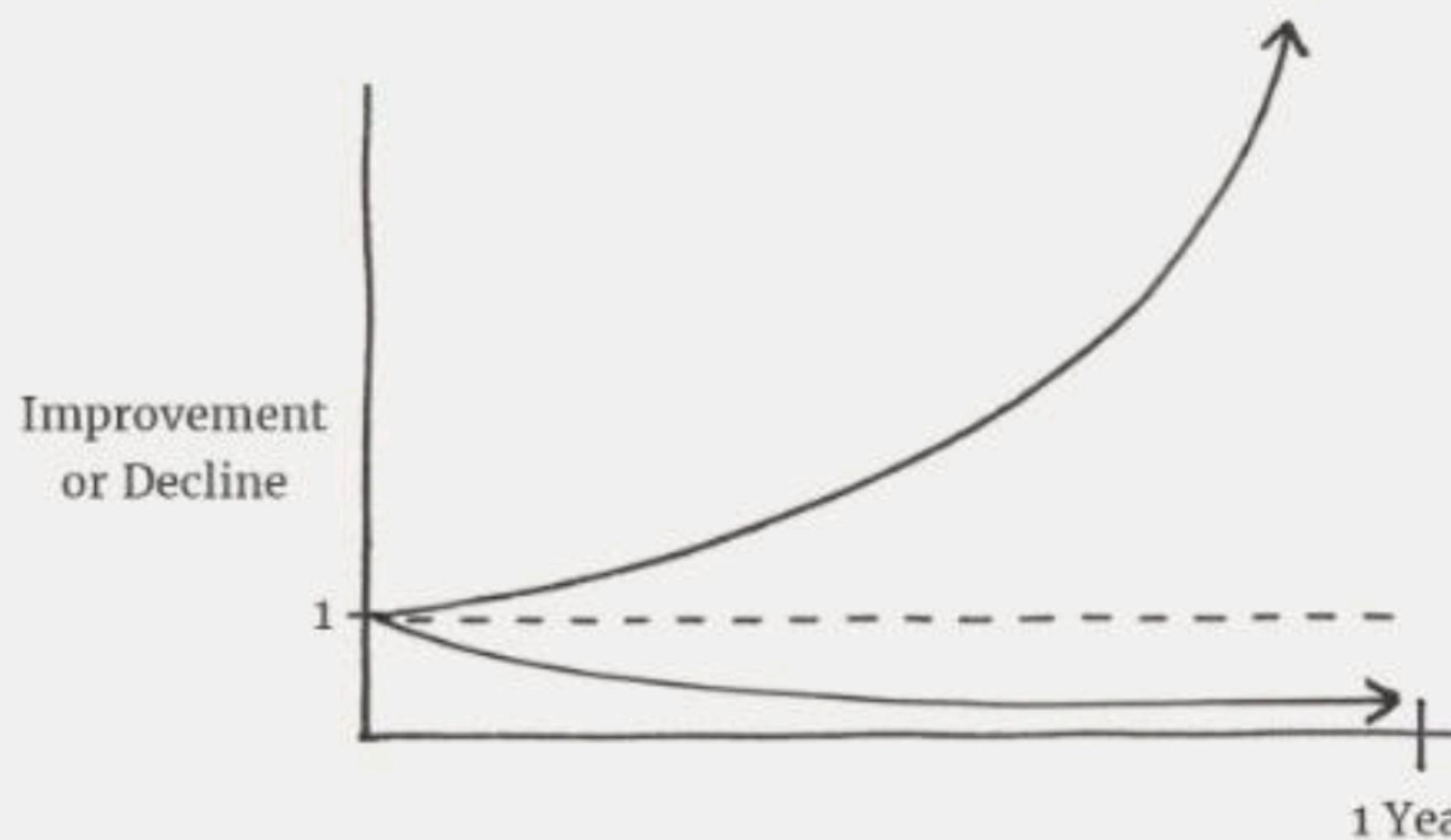
 blog.bytebytego.com



# Consistency is the key!



1% better every day  $1.01^{365} = 31.18$   
1% worse every day  $0.99^{365} = 0.03$



Credit: JamesClear.com

# Resources for developing coding skills

**Teach yourself Programming in Ten Years :**  
<https://norvig.com/21-days.html>

- TopCoder: <https://www.topcoder.com/>
- Leetcode: <https://leetcode.com/>
- CodeChef: <https://www.codechef.com/>
- HackerRank: <https://www.hackerrank.com/>
- Hackerearth: <https://hackerearth.com/>
- Kaggle: <https://www.kaggle.com/>
- Git: <https://git-scm.com/docs/gittutorial>
- How to contribute to open source: <https://opensource.guide/how-to-contribute/>



Have any questions?!





## Abinaya Mahendiran

- CTO, Nunnari Labs,
- Program Manager, IITM
- M.Tech IT, International Institute of Information and Technology Bangalore
- Volunteer at AI Tamil Nadu, WTM, Data Conversations, GHCI, WAI, Women Who Code
- Interests: Building NLP/NLU/NLG/MLOps/Gen AI systems, Open source, Applied Research

 <https://abinayam02.github.io>

 <https://www.linkedin.com/in/abinayamahendiran/>

 @freakynut

 <https://medium.com/@abinayamahendiran>

 [https://topmate.io/abinaya\\_mahendiran](https://topmate.io/abinaya_mahendiran)