

Open Project Management

from an “open” perspective

Instructor: Dr. Bradly Alicea

<http://bradly-alicea.weebly.com>



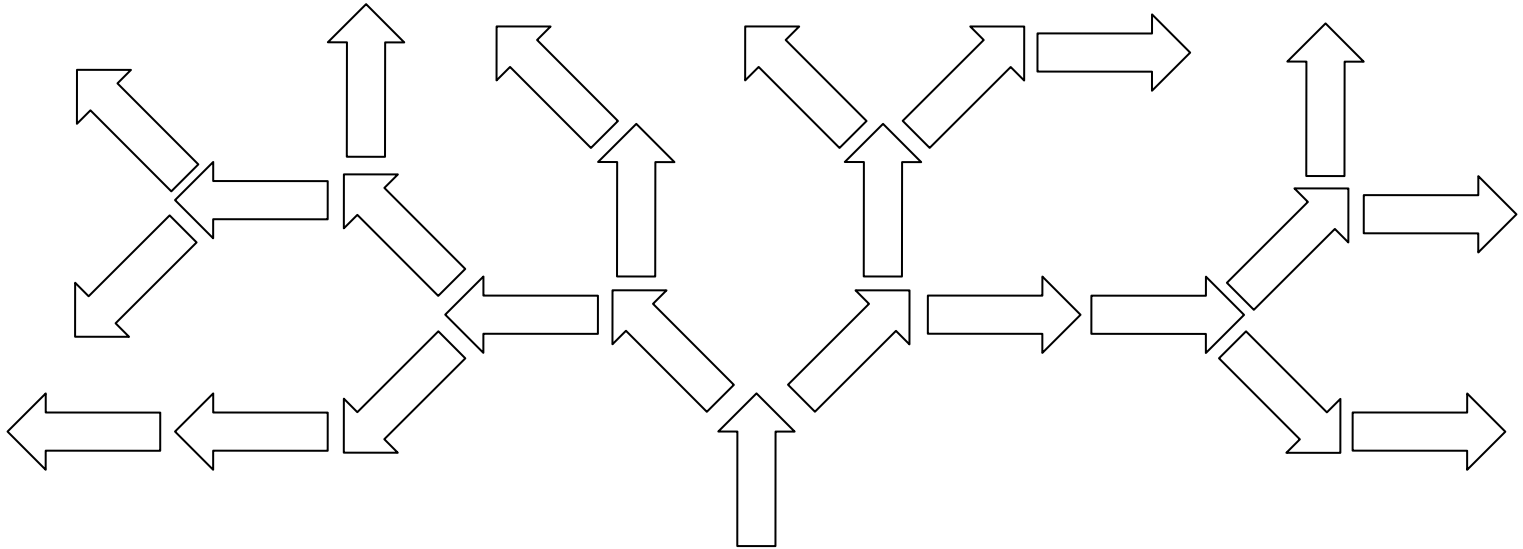
Lecture 4

All content



Open Project Management

Welcome Back!



Cloning a software repository

openworm / OpenWorm Public

Watch 158 Fork 166 Star 1.7k

<> Code Issues 41 Pull requests Discussions Actions Security Insights

master 23 branches 5 tags

Clone

Local Codespaces (New)

Clone

HTTPS SSH GitHub CLI

<https://github.com/openworm/OpenWorm.git>

Use Git or checkout with SVN using the web URL.

Open with GitHub Desktop

Download ZIP

Repository for the main Dockerfile with the Openworm software stack and project-wide issues

openworm.org

simulation biology openworm elegans

Readme MIT license 1.7k stars 158 watching 166 forks

Releases 5

0.9.2 (Latest) on Aug 15, 2022

+ 4 releases

Packages

No packages published

Contributors 18

ps://github.com/openworm/OpenWorm/archive/refs/heads/master.zip

OpenWorm main codebase

<https://github.com/openworm/OpenWorm>



Orthogonal Research & Education Lab

Overview Repositories 26 Projects 6 Packages Teams 11 People 22 Settings

Find a team...

New team

<input type="checkbox"/> Select all		Visibility ▾	Members ▾
<input type="checkbox"/>	Cognition Futures Cognition Futures is a new research team within Orthogonal Research and Educa...		9 members 0 teams
<input type="checkbox"/>	Education and Media		6 members 0 teams
<input type="checkbox"/>	Epistemological-Directories Front-end for repositories associated with the Epistemological Directories (E...		7 members 0 teams
<input type="checkbox"/>	Ethics & Society		8 members 0 teams
<input type="checkbox"/>	GSoC Google Summer of Code		5 members 0 teams
<input type="checkbox"/>	Innovation in Open Science		7 members 0 teams
<input type="checkbox"/>	MIT-CBMM-Summer OREL participants in MIT Center for Brains Minds and Machines Summer Course		2 members 0 teams
<input type="checkbox"/>	Neuromatch Neuromach Conferences & Summer Schools - Computational Neuroscience. members,...		10 members 0 teams
<input type="checkbox"/>	Representational-Brains-Phenotypes Projects in Meta-Brain Models and Empirical and Artificial Models of Neurobeh...		13 members 2 teams ▾

TO RECAP:

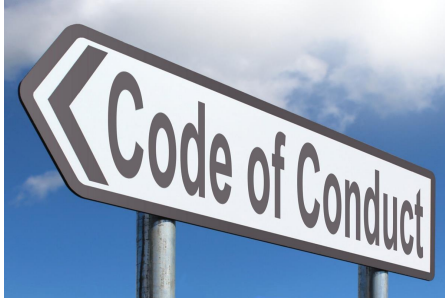
Open-source Community Governance

How do we manage multiple collaborators and their preferences?

Technologically-mediated: version-control, authority control, argument maps.

Governance: the creation of policies that enable collective wisdom and a means to mediate interactions through laws, norms, and power relations.

Code of Conduct

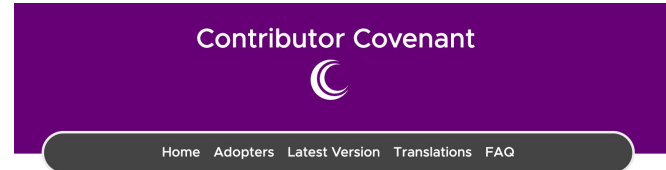


Shared values and norms of a collaborative community, will be unique to project type.

- centralized reference document, ethics benchmark.

Purposes:

- encouraging general compliance across organization.
- public declaration of current standards (versioned as needed).
- risk mitigation for misconduct.



A Code of Conduct for Open Source Communities

<https://www.contributor-covenant.org/>



ENTERPRISE

Richard Stallman's costly return to the Free Software Foundation

After the nonprofit renamed its founder to a leadership position despite years of complaints about his behavior, board members and contributors are heading for the exits.



Richard Stallman was reinstated as president of the Free Software Foundation. | Photo: Massimiliano Ferraro/Getty Images

<https://www.protocol.com/enterprise/richard-stallman-free-software-foundation>

SOFTWARE

Apache® Appropriation

It is not uncommon to learn about non-Indigenous entities appropriating Indigenous culture but none of them are as large, prestigious, nor well-known as The Apache® Software Foundation is in software circles.



ADAM RECVLOHE, HOLLY GRIMM, DESIREE KANE

6 JAN 2023 • 5 MIN READ

<https://blog.nativesintech.org/apache-appropriation/>

Defining a company policy to handle harassment in open source

We're sharing our guidance on Open Source harassment



Per Ploug
Open Source Community Manager

Posted on Feb 04, 2019

Tags: [Open Source](#), [InnerSource](#)

Open Source Participation

When you as a Zalando employee engage in open source communities as part of your work, you will interact with the wider open source communities outside Zalando - this is generally a good experience and collaborating with many different types of developers with different backgrounds is generally a positive input to your personal development.

However, there is also a small risk of encountering negative or even abusive behavior from community members when you act as an open source contributor or maintainer.

As an employer encouraging open source participation, we have decided to devise a [policy](#) for how we as a company can support our employees in case of harassment.

Statistics on harassment in open source

An extensive [survey by Github](#) in 2017 showed that nearly one out of five have experienced negative behavior personally and 50% have witnessed it between other people - fortunately outright harassment is much less likely with 14% witnessing it and 3% experiencing it personally.

<https://engineering.zalando.com/posts/2019/02/open-source-harassment-policy.html>

Conduct Conduit: Community Standards

What is allowable and what is not. Based on shared values, minimal requirements, standards for civility, privacy and security, and technology stack.

Facebook (<https://www.facebook.com/help/477434105621119>), Roblox (<https://en.help.roblox.com/hc/en-us/articles/203313410-Roblox-Community-Standards>), Wikipedia (https://en.wikipedia.org/wiki/Wikipedia:Expectations_and_norms_of_the_Wikipedia_community)

Facebook Content Management Controversies (https://en.wikipedia.org/wiki/Facebook_content_management_controversies)

Conduct Conduit: Community Standards

What is allowable and what is not. Based on shared values, minimal requirements, standards for civility, privacy and security, and technology stack.

Facebook (<https://www.facebook.com/help/477434105621119>), Roblox (<https://en.help.roblox.com/hc/en-us/articles/203313410-Roblox-Community-Standards>), Wikipedia ([https://en.wikipedia.org/wiki/Wikipedia:Expectations and norms of the Wikipedia community](https://en.wikipedia.org/wiki/Wikipedia:Expectations_and_norms_of_the_Wikipedia_community))

Facebook Content Management Controversies ([https://en.wikipedia.org/wiki/Facebook_content_management controversies](https://en.wikipedia.org/wiki/Facebook_content_management_controversies))

Open Standards:

- Freely available for adoption, implementation, and updates.
- Managed uniformly across organization, may converge with professional standards.
- Versioning managed via version-control, forking, and coordinated governance.

Contributing Guidelines: Wikimedia



Wikimedia.org

Launched September 2004

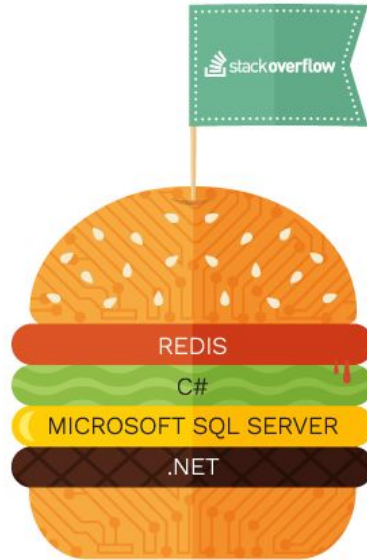
Goal: global movement to bring free educational content to the world.

Open-source community centered upon content creation.

Community: a group of people who register as members, participate, and exchange.

Reward and Motivation: status related to frequency of participation and attribution.

Further Standards: Technology Stacks



COURTESY (Infographic):

<https://www.thesoftwareguild.com/blog/build-your-own-technology-stack/>

The Impact of New Technologies on Open-source Projects

Why posting GPT and ChatGPT generated answers is not currently acceptable

This Help Center article provides insight and rationale on our policy regarding the usage of GPT and ChatGPT on Stack Overflow. While this is the position of Stack Overflow staff, it's meant to support the prior work done by moderators (namely, the [temporary policy](#) issued to ban contributions by ChatGPT).

Stack Overflow is a community built upon trust. The community trusts that users are submitting answers that reflect what they actually know to be accurate and that they and their peers have the knowledge and skill set to verify and validate those answers. The system relies on users to verify and validate contributions by other users with the tools we offer, including responsible use of upvotes and downvotes. Currently, contributions generated by GPT most often do not meet these standards and therefore are not contributing to a trustworthy environment. This trust is broken when users copy and paste information into answers without validating that the answer provided by GPT is correct, ensuring that the sources used in the answer are [properly cited](#) (a service GPT does not provide), and verifying that the answer provided by GPT clearly and concisely answers the question asked.

The objective nature of the content on Stack Overflow means that if any part of an answer is wrong, then the answer is objectively wrong. In order for Stack Overflow to maintain a strong standard as a reliable source for correct and verified information, such answers must be edited or replaced. However, because GPT is good enough to convince users of the site that the answer holds merit, signals the community typically use to determine the legitimacy of their peers' contributions frequently fail to detect severe issues with GPT-generated answers. As a result, information that is objectively wrong makes its way onto the site. In its current state, GPT risks breaking readers' trust that our site provides answers written by subject-matter experts.

“Researchers should ask themselves how the transparency and trust-worthiness that the process of generating knowledge relies on can be maintained if they or their colleagues use software that works in a fundamentally opaque manner”.

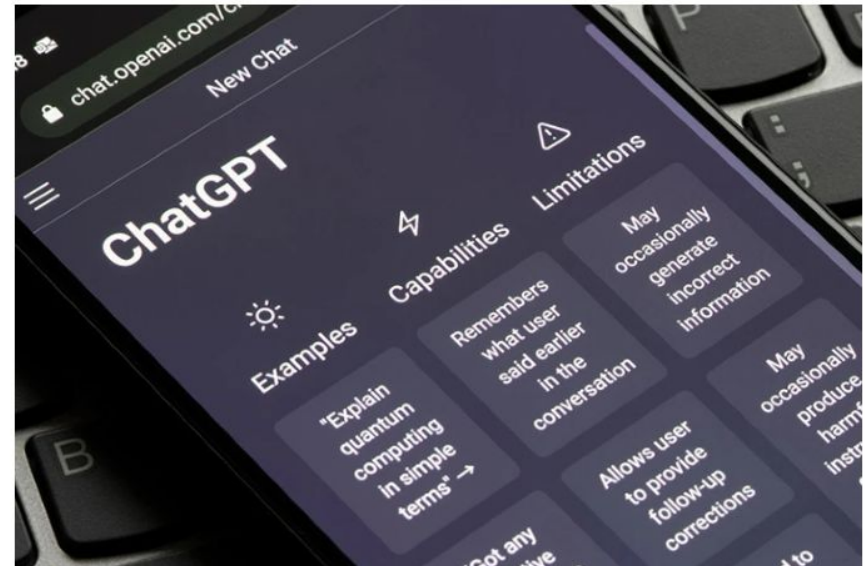
EDITORIAL | 24 January 2023

Tools such as ChatGPT threaten transparent science; here are our ground rules for their use

As researchers dive into the brave new world of advanced AI chatbots, publishers need to acknowledge their legitimate uses and lay down clear guidelines to avoid abuse.



Nature, 613, 612 (2023).



Stable Diffusion model trained on LAION open-source datasets, billions of images with English captions.



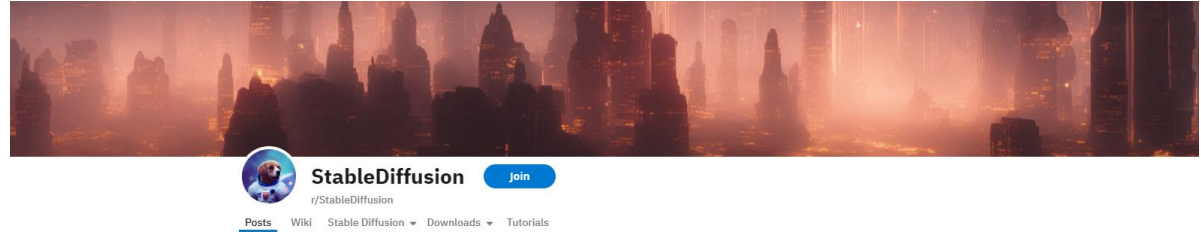
<https://waxy.org/2022/08/exploring-12-million-of-the-images-used-to-train-stable-diffusions-image-generator/>

User-generated content (photo and art) platforms as source for the data: Pinterest, wordpress.com, Smugmug, Blogspot, Flickr, DeviantArt, Wikimedia, and Tumblr.

- images sampled and encoded into a latent space, then available for general public to use under open-source license (CreativeML Open RAIL-M).

Stable Diffusion Community

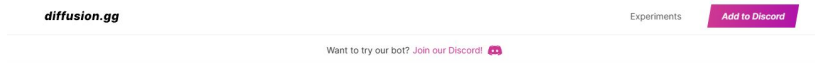
Funded by Stability.ai, multiple forks that can be further trained on customized datasets.



List of forks:

<https://github.com/CompVis/stable-diffusion/network/members>

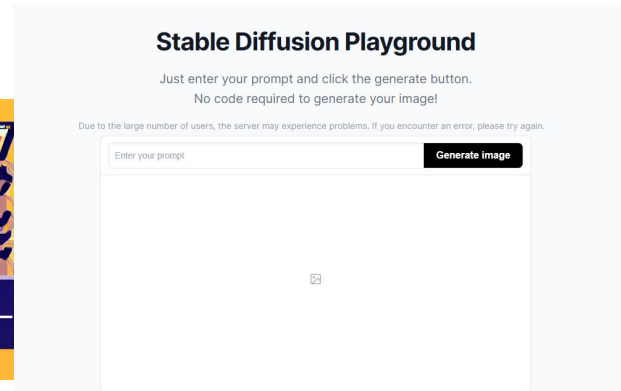
Github repos, Reddit forum, Discord server, a FAQ, and online demos



Create beautiful Cartoons from text

Use /draw in Discord to create unique and original images from natural language.

[Connect to Discord](#)



Ethics of Developing Open-source Generative Models



Jae Duk Seo

Jan 15 · 6 min read · Member-only · Listen



AI Art Lawsuit Chaos: Stable-Diffusion & Mid-Journey

Stable Diffusion, Mid-journey, Deviant Art on Trial: Lawsuit Details Inside



[Image from Unsplash](#)

Lawsuit challenging Stable Diffusion filed against DeviantArt & others



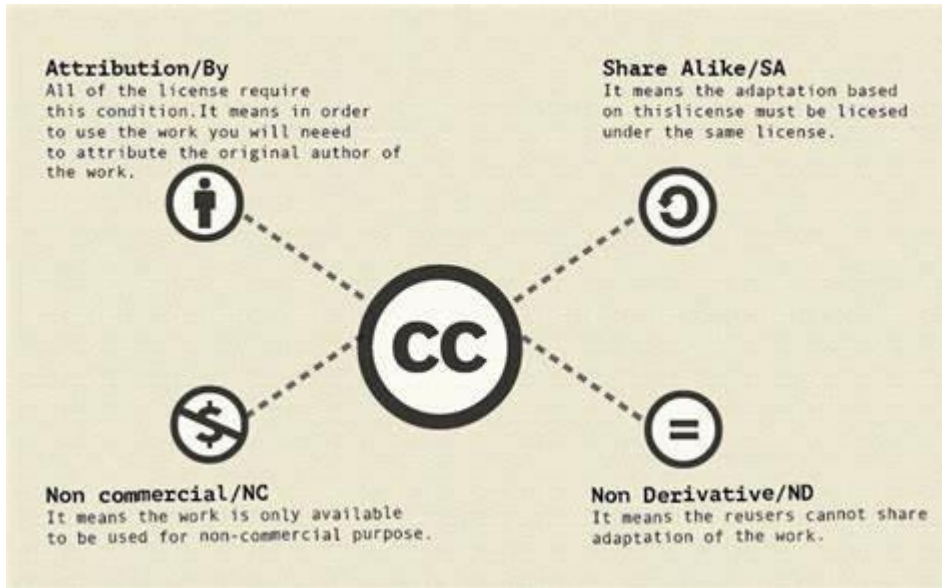
January 13, 2023

How do you select a license for your content?

Fair Use and Ethics of the Remix Economy

Licensing, Creative Commons

Founded in 2001 by Lessig, Eldred, and Abelson.























































- purpose: expansion of reasonable, flexible copyright (copyleft).
- As the stakes grew for copying and remixing via the internet, need became apparent.

First license: December 2002.

Two billion works under these licenses, including the whole of Wikipedia.

Licensing, Creative Commons

CREATIVE COMMONS LICENSES		 COPY & PUBLISH	 ATTRIBUTION REQUIRED	 COMMERCIAL USE	 MODIFY & ADAPT	 CHANGE LICENSE
	PUBLIC DOMAIN					
	CC BY					
	CC BY-SA					
	CC BY-ND					
	CC BY-NC					
	CC BY-NC-SA					
	CC BY-NC-ND					

				
You can redistribute (copy, publish, display, communicate, etc.)	You have to attribute the original work	You can use the work commercially	You can modify and adapt the original work	You can choose license type for your adaptations of the work.

The Creative Commons licences explained.



You can redistribute (copy, publish, display, communicate, etc.)



You can use the work commercially



You have to attribute the original work



You can modify and adapt the original work



You can choose license type for your adaptations of the work.

Licensing, Open-source Software



“permissive software licenses”, “acknowledgment of the original source of the code in all advertising materials”

<https://pitt.libguides.com/openlicensing/opensoftware>

Licensing, Open-source Software



Can “use, copy, modify, merge publish, distribute, sublicense and sell copies of the software”

<https://pitt.libguides.com/openlicensing/opensoftware>

Licensing, Open-source Software



"Work based on another work" which triggers the licenses copyleft obligations.

Licensing, Open-source Software



“users may distribute, modify, and distribute modified versions. Permits derivative works to be distributed under a different, more restrictive, license (not copyleft).”



<https://pitt.libguides.com/openlicensing/opensoftware>

Project Management Strategies for Open Data

Making things open and shareable but also
private and secure when appropriate

Open Data



Wikipedia:

“Openly accessible, exploitable, editable and shared by anyone for any purpose. Open data is licensed under an open license”

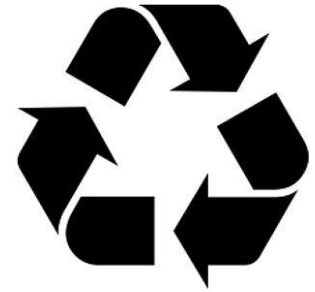
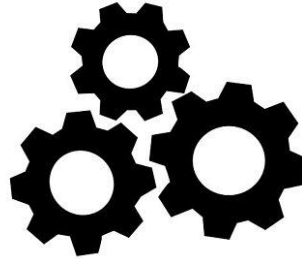
Open Data allows for data to be easily shared and reused, but requires governance, issues of ownership, and adherence to community standards.

F
Findable

A
Accessible

I
Interoperable

R
Reusable



Formats: csv, json, XML.

Standards: FAIR, Production, Archival, Privacy.

Production and Reuse: vocabularies, licenses, metadata and identification, quality assurance.

Does your project organization need to invest in stakeholder relations?

- relevant communities (e.g. municipalities, medical doctors, public advocates).
- recruit from areas of domain knowledge?

Does your project organization need to invest in stakeholder relations?

- relevant communities (e.g. municipalities, medical doctors, public advocates).
- recruit from areas of domain knowledge?

Does your project require a vested interest in privacy standards?

- PPI (protected personal information), standards for anonymity and opting-out.
- adherence to data privacy laws (GDPR in Europe; in US, state by state).

Does your project organization need to invest in stakeholder relations?

- relevant communities (e.g. municipalities, medical doctors, public advocates).
- recruit from areas of domain knowledge?

Does your project require a vested interest in privacy standards?

- PPI (protected personal information), standards for anonymity and opting-out.
- adherence to data privacy laws (GDPR in Europe; in US, state by state).

Does your project need to invest in a formal data strategy?

- data analysis and handling pipeline (different from project planning pipeline!)
- metadata management (annotation, ontologies).