

The Business of Open Source @ Google



amanda casari (she/her)

Manager, Google OSPO

@amcasari

07FEB2025

Irish National Open Source Innovation Summit



Google Open Source



hello, i'm amanda casari 

Open Source + AI Lead, Google Open Source
Programs Office

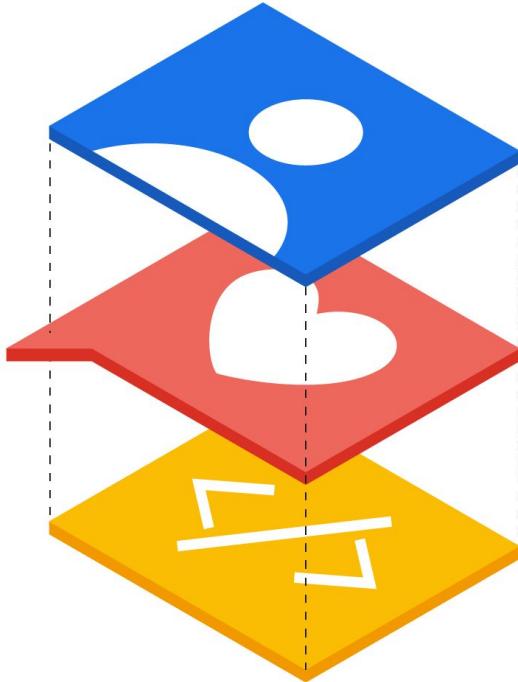
Manager, Open Source Contributor Experience Team

External Faculty @ Vermont Complex Systems Center

FOSS Fund Advisory Council, [Open Technology Fund](#)

Google representative on [NumFOCUS Advisory Board](#)

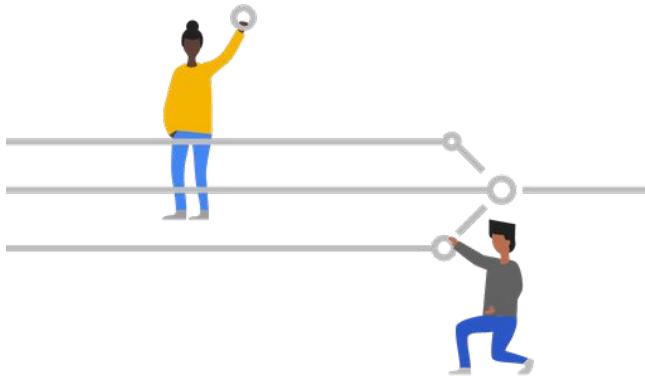
Co-author with Alice Zheng, “Feature Engineering for Machine Learning” [O'Reilly, 2018]



Google <3 Open Source

What is open source?

The [open source definition](#) standardized by the Open Source Initiative focuses *mostly* on how software is distributed and licensed.



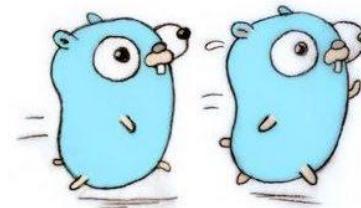
Android



Flutter



TensorFlow



What is openly available?

Openly available is how we broadly describe sharing technology under a wide spectrum of granting public permissions for creative works.



e.g. [Creative Commons licenses](#)

MapReduce: Simplified Data Processing on Large Clusters

Jeffrey Dean and Sanjay Ghemawat

jeff@google.com, sanjay@google.com

Google, Inc.

Abstract

MapReduce is a programming model and an associated implementation for processing and generating large data sets. Users specify a *map* function that processes a key/value pair to generate a set of intermediate key/value pairs, and a *reduce* function that merges all intermediate values associated with the same intermediate key. Many real world tasks are expressible in this model, as shown in the paper.

given day, etc. Most such computations are conceptually straightforward. However, the input data is usually large and the computations have to be distributed across hundreds or thousands of machines in order to finish in a reasonable amount of time. The issues of how to parallelize the computation, distribute the data, and handle failures conspire to obscure the original simple computation with large amounts of complex code to deal with these issues.

Attention Is All You Need

Ashish Vaswani*
Google Brain
avaswani@google.com

Noam Shazeer*
Google Brain
noam@google.com

Niki Parmar*
Google Research
nikip@google.com

Jakob Uszkoreit*
Google Research
usz@google.com

Llion Jones*
Google Research
llion@google.com

Aidan N. Gomez* †
University of Toronto
aidan@cs.toronto.edu

Lukasz Kaiser*
Google Brain
lukaszkaiser@google.com

Ilia Polosukhin* ‡
illia.polosukhin@gmail.com

Abstract

The dominant sequence transduction models are based on complex recurrent or convolutional neural networks that include an encoder and a decoder. The best performing models also connect the encoder and decoder through an attention mechanism. We propose a new simple network architecture, the Transformer,

[Code](#) [Issues 262](#) [Pull requests 29](#) [Actions](#) [Projects](#) [Security](#) [Insights](#)[main](#) ▾[1 Branch](#)[13 Tags](#)[Go to file](#)[Code](#) ▾**dliesecki** and **copybara-github**

Deprecate the AlphaFold 2 notebook



2ca41b1 · last week

156 Commits



afdb

Release code for v2.3.0

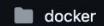
3 years ago



alphafold

Do not use jnp.clip arguments which are not supported in ...

2 weeks ago



docker

Update Conda to 24.11.1 to fix Docker build failure and sm...

2 months ago



docs

Release code for v2.3.0

3 years ago



imgs

Initial release of AlphaFold.

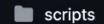
4 years ago



notebooks

Deprecate the AlphaFold 2 notebook

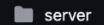
last week



scripts

Update wwPDB paths from ftp.wwpdb.org to files.wwpdb....

2 years ago



server

Update JSON format docs with new template on/off, cuto...

2 weeks ago



.dockerrcignore

Update .dockerrcignore

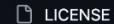
3 years ago



CONTRIBUTING.md

Initial release of AlphaFold.

4 years ago



LICENSE

Initial release of AlphaFold.

4 years ago



README.md

Deprecate the AlphaFold 2 notebook

last week



requirements.txt

Pin setuptools on a version smaller than 72.0.0.

6 months ago

About

Open source code for AlphaFold 2.[Readme](#)[Apache-2.0 license](#)[Activity](#)[Custom properties](#)[13.1k stars](#)[228 watching](#)[2.3k forks](#)[Report repository](#)

Releases 13

[Alphafold v2.3.2](#) Latest

on Apr 5, 2023

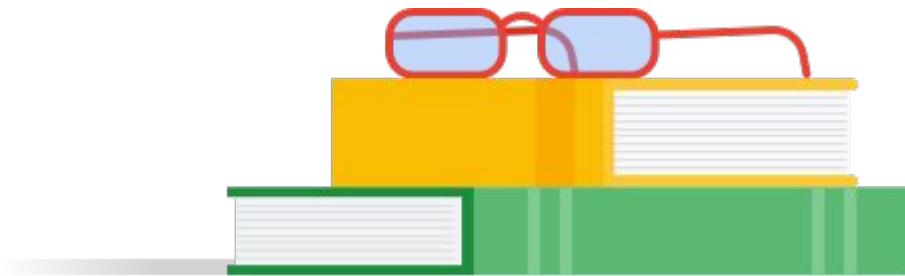
[+ 12 releases](#)

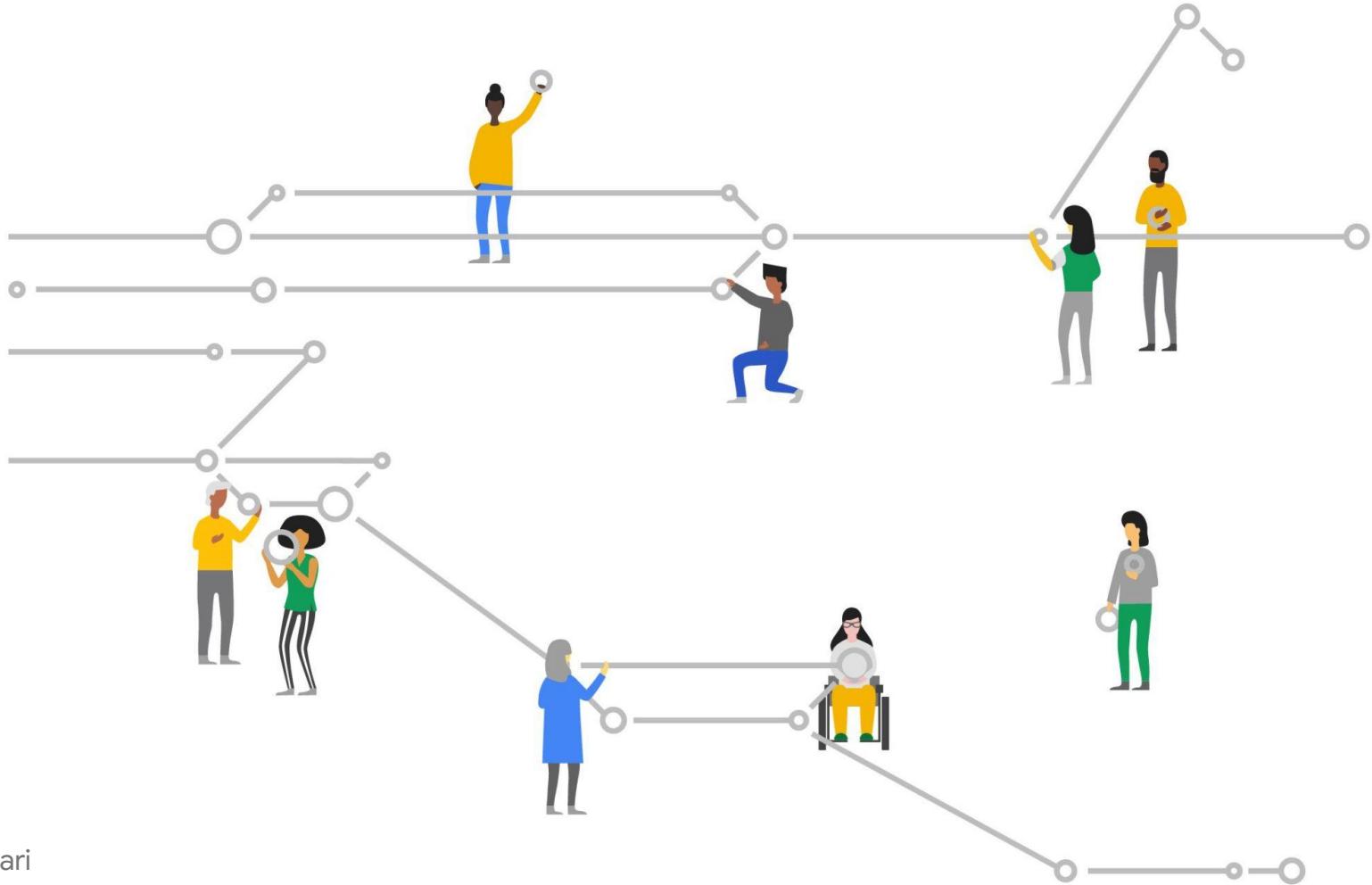
Contributors 18



"Open" is about more than just licenses + code.

The "open" in open source is also a collective (*but not always standardized*) framework that shapes how software is created, released, shared, and distributed — **as well as** the community and how it forms around the projects.





Our mission is to organize the
world's **information** and make it
universally accessible and
useful.

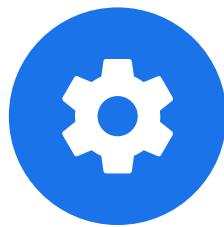


Google Open Source

Bringing all the value of open source to Google and
all the resources of Google to open source.

opensource.google

Open source touches every part of Google.



Programming
Languages



Software
development



Global
networking



Scientific
Research



The Business Impact of Open Source

**When open source becomes
essential for your organization,
you need new tools.**

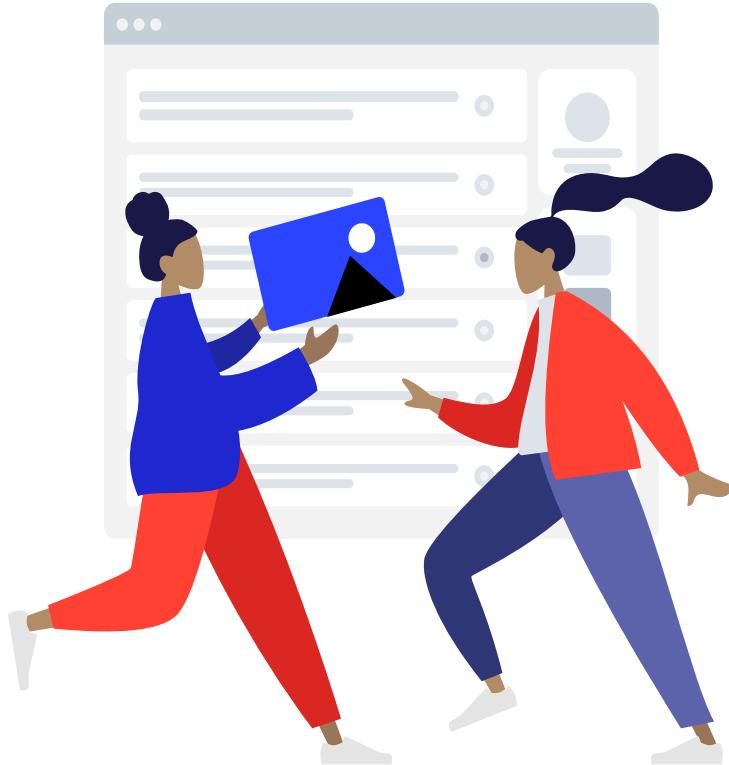


Organizations need common maps to navigate open source.

Does everyone in your organization:

- know what licenses they can use to create new products?
- know if you have policies on releasing what they have created?
- feel comfortable with what software they can use to create something new?
- have someone to ask "open" questions to?





Welcome, OSPOs ✨

@amcasari

 Google Open Source

OSPOs come in many shapes and sizes.

They may focus on:

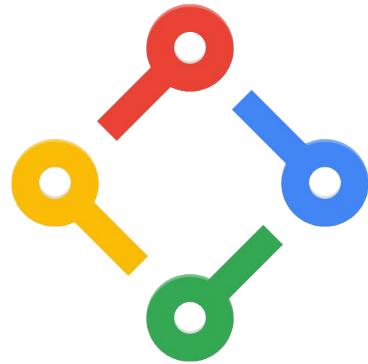
- navigating open source licenses
- funding open source projects your organization relies on
- focusing on internal compliance
- developing and documenting organizational best practices
- synthesizing decentralized efforts into scalable programs
- tracking organizational investments in open source ecosystems
- <and many, many more!>



OS[P]Os come in many shapes and sizes.

- Programs
- Projects
- Policy
- Policies
- Progress
- Processes
- Products
- Problems
- Pi<e>!





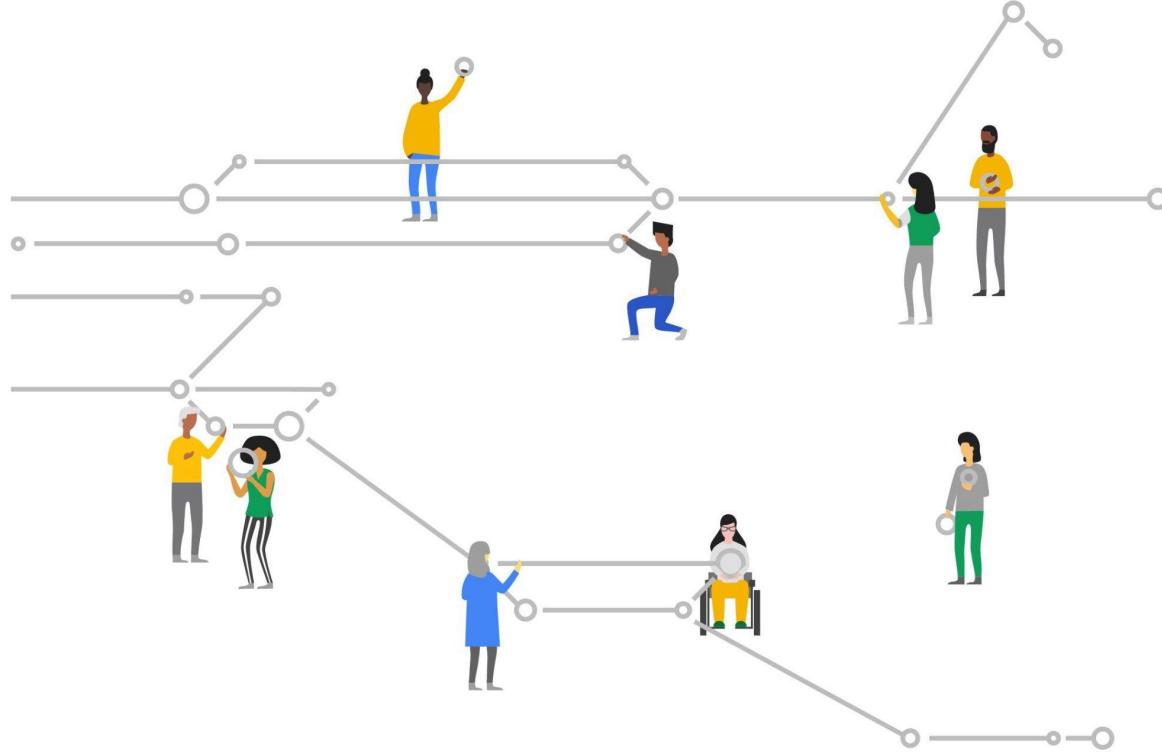
Google Open Source

Bringing all the value of open source to Google and
all the resources of Google to open source.

opensource.google

How do we know which "all" to focus on and for whom?

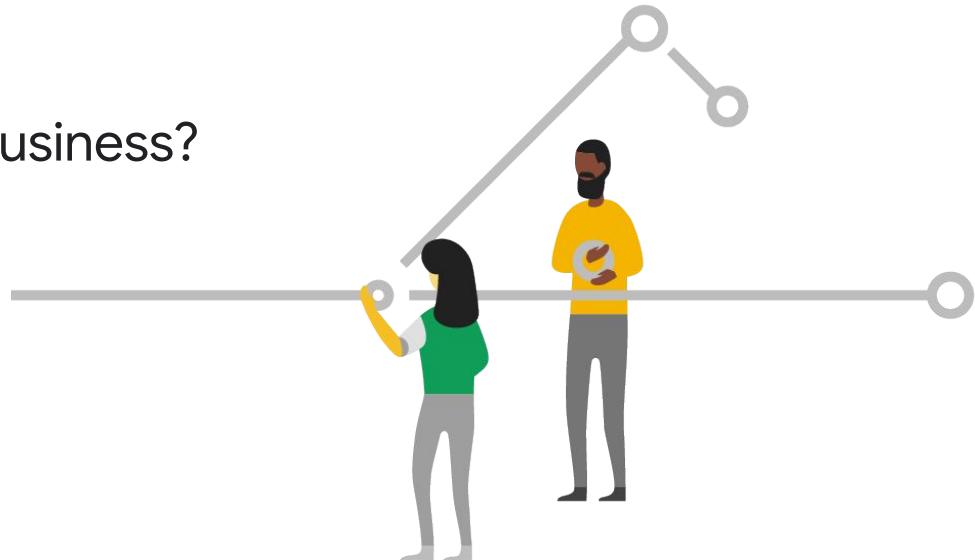




With great open source priorities, there must also come great expectation management.

Questions we ask about our open source priorities

1. Why are we doing this?
2. Who are we doing this for?
3. Who are the stakeholders?
4. What is the impact on our business?



Questions we ask about our open source programs

1. Why are we doing this?
2. Who are we doing this for?
3. Who are the stakeholders?
4. What is the impact on our business?



@amcasari

Focus on useful metrics.

Consider the processes, outcomes, and feedbacks.



More on this: bit.ly/foss-backstage-prm

@amcasari

Google Open Source

What defines success for your stakeholders?

| Stakeholder (Team/Role) | Why is this stakeholder important to you? | How does this stakeholder define success? | How often does this stakeholder measure our progress? | Signals to explore for KPIs (metrics of success) |
|-------------------------|---|---|---|--|
| | | | | |

- For each stakeholder, map out what they care about to identify additional metrics which will help you explain your vision of success
- How a stakeholder defines success may be listed in their —
 - Organizational/team mission or vision
 - Business goals (such as OKRs or quarterly goals)
 - Shared roadmaps
 - Shared dashboards (what do they track and why?)

@amcasari

Google Open Source

The Business Impact of Open Source

@amcasari



With great open source priorities, there must also come great investment management.

@amcasari

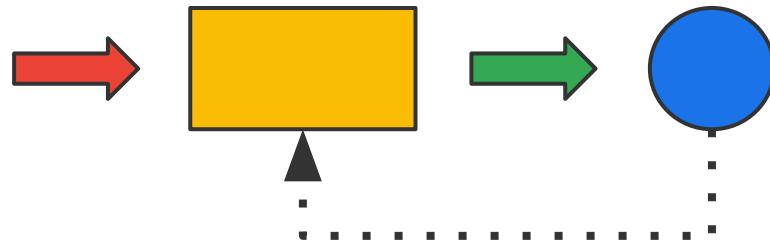
 Google Open Source

Avoid estimating ROI for your open source efforts.

- **Return on Investment is a very specific economic model**
 - Created as a way to relate *profits* to *capital invested*
 - Generally used when economic modeling is simplified to monetary equivalents
- **Many investments in open source do not readily simplify to standardized monetary equivalents**
 - *Code reviews for strategic open source projects*
 - *Board memberships for standards bodies*
 - *Employees volunteering as mentors + sponsors in their communities*

Focus on the business impact of your programs.

We've shifted to an input-output model of "Investment : Impact"



- Identify + track your investments
 - Project funding
 - Event sponsorships
 - Employee contributor programs
- Identify + track these impacts
 - How do you + your stakeholders track success?

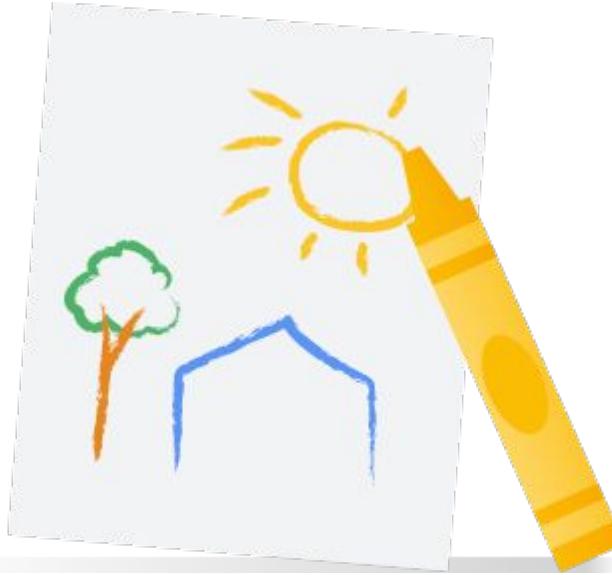
You can demonstrate business impact with useful metrics and KPIs.





Metrics are measurements.

Measurements are recorded observations — in and of themselves, there is no judgement we should make.



Observations are NOT unbiased.

- What you choose to observe, the lens you use to observe it, what you choose to leave out — all of these decisions bias the outcomes, regardless of your intention.
- You **must** record design decisions and **be clear** as to your motives for observing and measuring.



Think of a scientific experiment.

- When you have a question, you want observations to understand what is happening.
- Collecting data from your observations paints a picture/tells a story/gives insight into what is happening.

Focus on useful metrics.

Consider the processes, outcomes, and feedbacks.



Identify
objectives



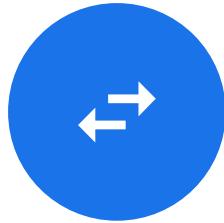
Define
observable
aspects



Establish a
baseline



Define your
desired trend



Determine
signal
phenomena

More on this: bit.ly/foss-backstage-pram

Google Open Source Blog

The latest news from Google on open source releases, major projects, events, and student outreach programs.

2023 Open Source Contributions: A Year in Review

Tuesday, August 13, 2024



At Alphabet, open source remains a critical component of our business and internal systems. We depend on thousands of upstream projects and communities to run our infrastructure, products, and services. Within the Open Source Programs Office (OSPO), we continue to focus on investing in the sustainability of open source communities and expanding access to open source opportunities for contributors around the world. As participants in this global ecosystem, our goal with this report is to provide transparency and to report our work within and around open source communities.

Use case: Measuring Alphabet's contributions to open source

- Author + project lead: Sophia Vargas
- Annual analysis examining Google's contributions to open source
 - Centered on Google Open Source's mission - "resources of Google for open source."
 - Explicitly outlines data samples + evolving methodology
 - Quantitative + qualitative methods
- More at opensource.googleblog.com

KPIs are a specific kind of metric – they measure the pulse points that correlate to a process, program, or product being considered to be "**successful**".



You need to define what success is, or looks like, **for you and all of your stakeholders** to create useful KPIs.



What defines success for you?

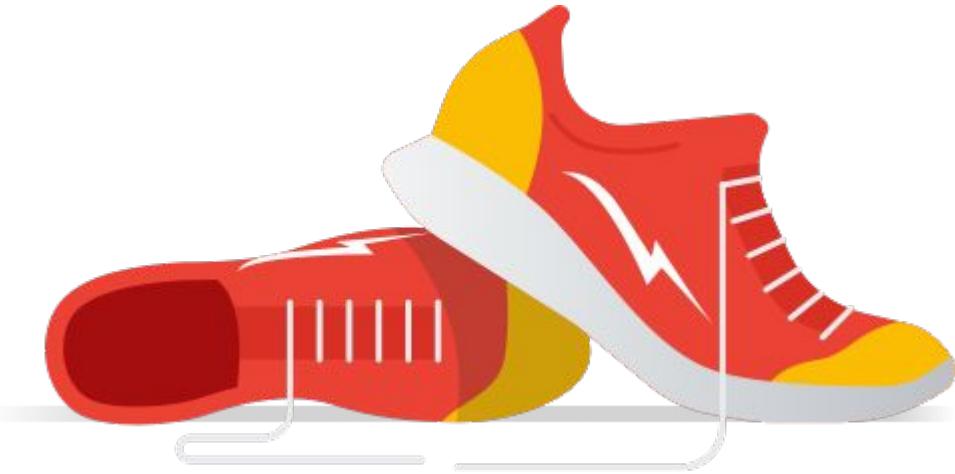
- How do you introduce people to your work?
- Have you written a mission statement?
- How do you talk about your team to other people in your organization?
- How do other people introduce your team's work?
- What do you point back to when you want to show The Why?

What defines success for your stakeholders?

| Stakeholder (Team/Role) | Why is this stakeholder important to you? | How does this stakeholder define success? | How often does this stakeholder measure our progress? | Signals to explore for KPIs (metrics of success) |
|----------------------------|--|---|--|--|
| | | | | |

- For each stakeholder, map out what they care about to identify additional metrics which will help you explain your vision of success
- How a stakeholder defines success may be listed in their —
 - Organizational/team mission or vision
 - Business goals (such as OKRs or quarterly goals)
 - Shared roadmaps
 - Shared dashboards (what do they track and why?)

*How often and when to measure KPIs depends on how quickly things change that you are measuring, **not how often someone wants a report on them.***



Considerations for **how + when**

to measure KPIs include:

- Aggregation buckets
 - *how to group or compare metrics*
- Collection cycles
 - *when to observe and how often*
- Reporting cycles
 - *when to make public your observations*
- Privacy + security risks





Use case: Defining KPIs for open source events

- Project Team: Amanda Casari, Kassandra Dhillon, Teresa Terasaki, María Cruz
- Collaborative effort to improve success metrics for all Google Open Source events
 - Multiple programs support different kinds of events
 - Identified overlapping stakeholders + success metrics
 - Clarified individual program missions to define program-specific KPIs



Normalize Open + Open Source + AI



@amcasari

Created using Gemini ✨

 Google Open Source

I'm an engineer.

I'm an engineer and a pragmatist.

As much as one can be,
I try to be a voice of reason.

What is an "AI"?

genAI

generative
pretrained
transformers

Neural
Networks

Deep Learning

Machine Learning

AI

AI - What is it good for?

Artificial intelligence – when machines perceive their environment, use learning and intelligence to take actions that maximize the chance of achieving a defined goal.

genAI advances problem solving in –

- image generation
- text summarization
- text generation
- extending context windows for semantic search
- question and answer design patterns (chatbots)

What is an "AI System"?

THE DATA SCIENCE HIERARCHY OF NEEDS

LEARN/OPTIMIZE

AGGREGATE/LABEL

EXPLORE/TRANSFORM

MOVE/STORE

COLLECT

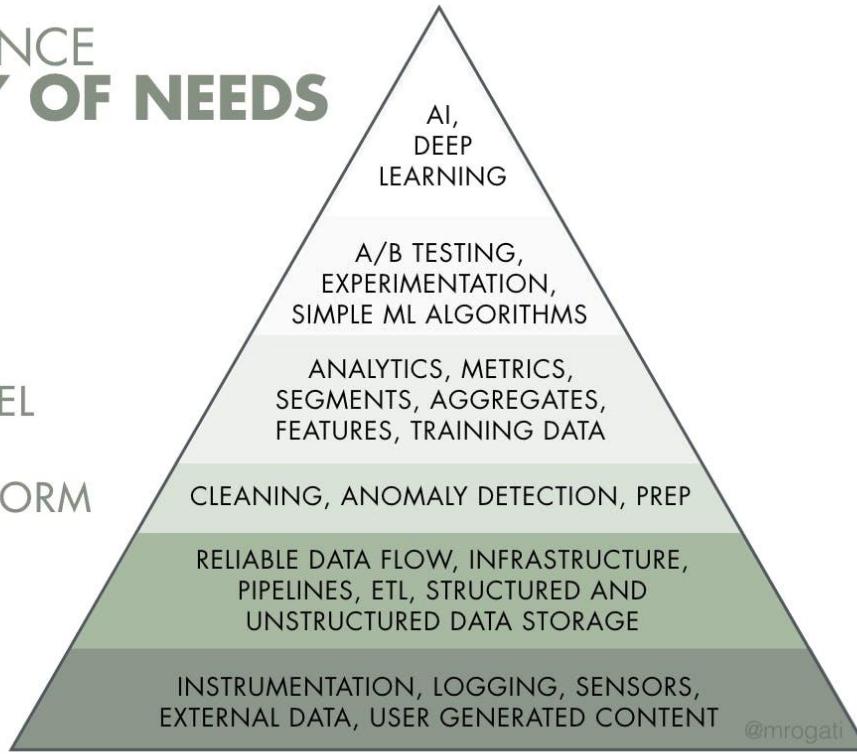


illustration + concept: Monica Rogati

From [The AI Hierarchy of Needs](#) (2017)

MODELS / WEIGHTS

/ AI TECHNOLOGY \

SOFTWARE

/ EXPLORE, PROTOTYPING \

DATA

From *The AI Hierarchy of Needs* (2017)



illustration + concept: Monica Rogati

Open Source + AI for Business

Models/Weights

Open models and weights are a starting point - what you pick up determines where you go next (including what business decisions you can make with what you build from these tools)

Software

Free and open source software have good existing business integration and operational patterns.

Data

Your Legal team will be your best friend - before you make decisions on what to do next, make sure the licenses, terms, and conditions of the data you are using allow for that.

Open Source + AI for Your Business

Build

You have the resources and the people to start pilot initiatives yourself.

Buy

You need to make tradeoffs between your bottom line and what you can justify maintaining over time.

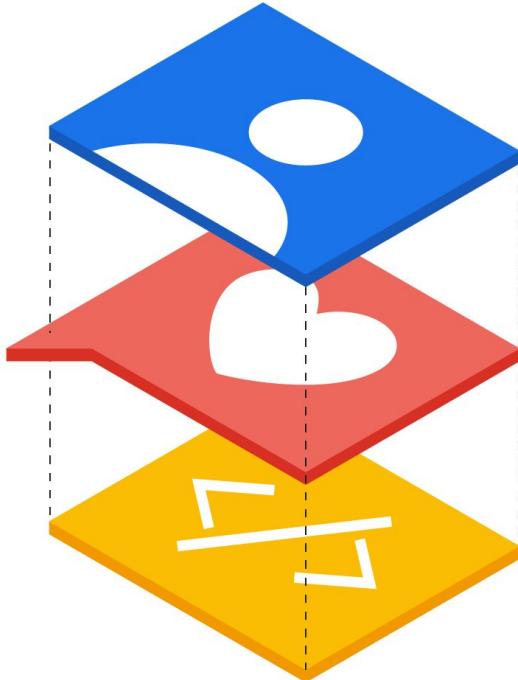
Hire

It's in your best or core business interest to have your own capacity, but you cannot get there yet from here.

If you only remember FOUR things...

- Open and open source is the license and is more than the software.
- Managing open source in organizations requires good expectation management practices.
- AI systems are not new - we can look to existing systems and business models to make reasonable decisions.
- Talk to your LawyerCat. They are nice people.





We are all sailing these open
source rafts, together 🐠

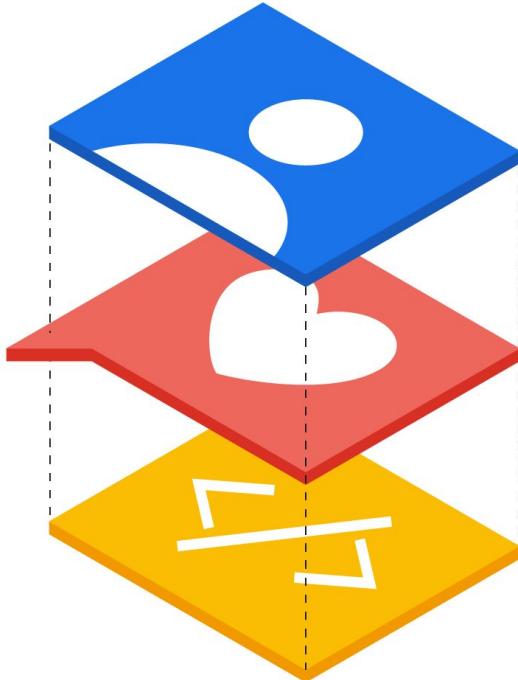
Thank you!



amanda casari (she/her)
Manager, Google OSP0
@amcasari



Google Open Source



**Live demo of scrolling
through a project to pick
out the terms?**

**YOUR PARTNER DOESN'T LIVE
HERE.**

**CLEAN UP YOUR SKIPPED
SLIDES - DELETE ALL BELOW
AFTER FINAL EDITS MADE**