Open Source Funding

Your Doing It Wrong

Ohai


After a decade saying we need to support our digital infrastructure, are we any closer to doing so?

Part 1: Usage is the best determinant of criticality

Part 2: Funding for ‘critical infrastructure’ is misdirected

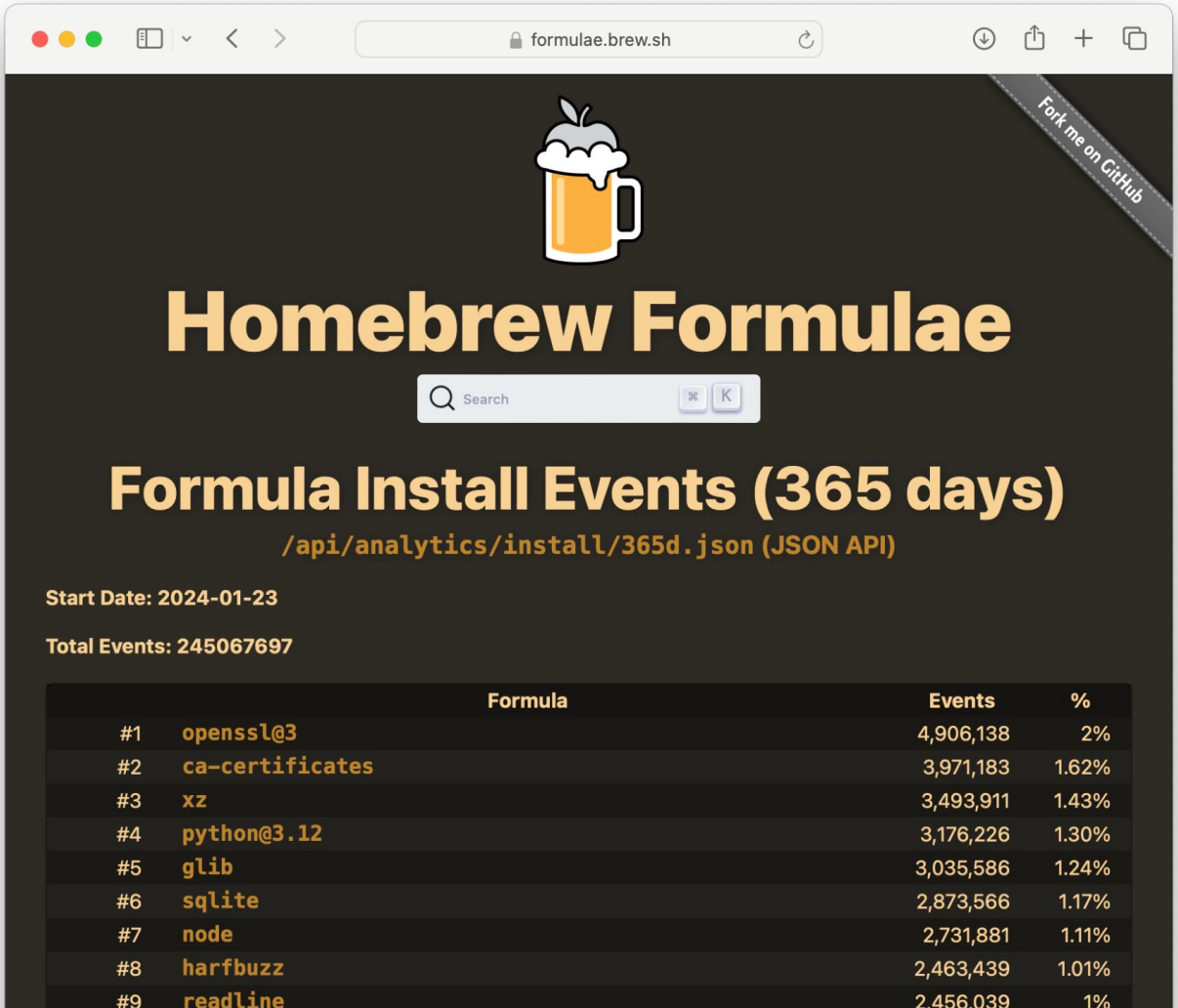
Identifying ‘critical infrastructure’


MARC MONDAVI'S
THE DIVINING ROD


WILLIAM FULD Talking Board Set
YES
OUIJA
NO
MYSTIFYING ORACLE
ABCDEFGHIJKLM
NOPQRS
VWXYZ
1
7890
GOOD BYE
KING of SWORDS

We can do better

Evidence-Based Methods

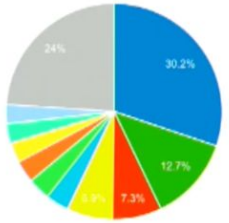

formulae.brew.sh
Fork me on GitHub
Homebrew Formulae
Search
⌘
K
Formula Install Events (365 days)
/api/analytics/install/365d.json (JSON API)
Start Date: 2024-01-23
Total Events: 245067697
Formula
Events
%
#1 openssl@3 4,906,138 2%
#2 ca-certificates 3,971,183 1.62%
#3 xz 3,493,911 1.43%
#4 python@3.12 3,176,226 1.30%
#5 glib 3,035,586 1.24%
#6 sqlite 2,873,566 1.17%
#7 node 2,731,881 1.11%
#8 harfbuzz 2,463,439 1.01%
#9 readline 2,456,039 1%

Andrew Nesbitt

Primary Dimension: **Event Category** Event Action Event Label

Plot Rows | Secondary dimension: npm version | Sort Type: Default | advanced

	Event Category	npm version	Total Events	Total Events	contribution to total: Total Events
			441,233 % of Total: 99.99% (441,290)	441,233 % of Total: 99.99% (441,290)	
1.	install	5.6.0	133,188	30.19%	
2.	install	3.10.10	55,926	12.67%	
3.	install	yarn-1.3.2	32,425	7.35%	
4.	install	5.5.1	30,392	6.89%	
5.	install	yarn-1.5.1	15,283	3.46%	
6.	install	5.7.1	14,894	3.38%	
7.	install	5.3.0	14,637	3.32%	
8.	install	3.10.3	13,351	3.03%	
9.	install	3.8.6	12,911	2.93%	
10.	install	4.6.1	12,419	2.81%	
11.	install	yarn-1.2.1	10,968	2.49%	
12.	install	3.10.8	10,630	2.41%	
13.	install	3.10.9	7,442	1.69%	
14.	install	4.2.0	5,886	1.33%	
15.	install	3.3.4	5,498	1.25%	
16.	install	4.1.2	4,611	1.05%	

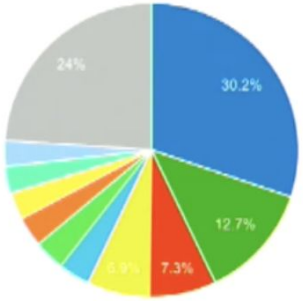


BATH Ruby 2018

Primary Dimension: **Event Category** Event Action Event Label

Plot Rows | Secondary dimension: npm version | Sort Type: Default | advanced

	Event Category	npm version	Total Events	Total Events	contribution to total: Total Events
			441,233 % of Total: 99.99% (441,290)	441,233 % of Total: 99.99% (441,290)	
1.	install	5.6.0	**133,188**	30.19%	
2.	install	3.10.10	**55,926**	12.67%	
3.	install	yarn-1.3.2	**32,425**	7.35%	
4.	install	5.5.1	**30,392**	6.89%	
5.	install	yarn-1.5.1	**15,283**	3.46%	
6.	install	5.7.1	**14,894**	3.38%	
7.	install	5.3.0	**14,637**	3.32%	
8.	install	3.10.3	**13,351**	3.03%	
9.	install	3.8.6	**12,911**	2.93%	
10.	install	4.6.1	**12,419**	2.81%	
11.	install	yarn-1.2.1	**10,968**	2.49%	
12.	install	3.10.8	**10,630**	2.41%	
13.	install	3.10.9	**7,442**	1.69%	
14.	install	4.2.0	**5,886**	1.33%	
15.	install	3.3.4	**5,498**	1.25%	
16.	install	4.1.2	**4,611**	1.05%	



about.scarf.sh

SCARF Platform Solutions Resources Pricing Company Join our community Log In SIGN UP

Scarf pioneered usage intelligence for open source

We believe that if open source is part of your business, open source should be a part of your success. Usage analytics, which has historically been taboo in open source, can actually add tremendous value for your project, community, and commercial business. We believe that transparently sharing usage data aligns with the value of “open” in Open Source, and should be embraced, rather than discouraged.

We are creating a world where users can contribute to any project, just by using it.

As open source community members ourselves, we know privacy is key – which is why Scarf never retains personally identifiable information (PII). Our commitment to responsible data practices are why we are used by some of the largest open source projects and foundations in the world.

“By leveraging analytics, our customers are able to demonstrate the impact of their open source project on their commercial business,"

Avi Press, Founder and CEO, Scarf

Census III of Free and Open Source Software

DOWNLOAD REPORT

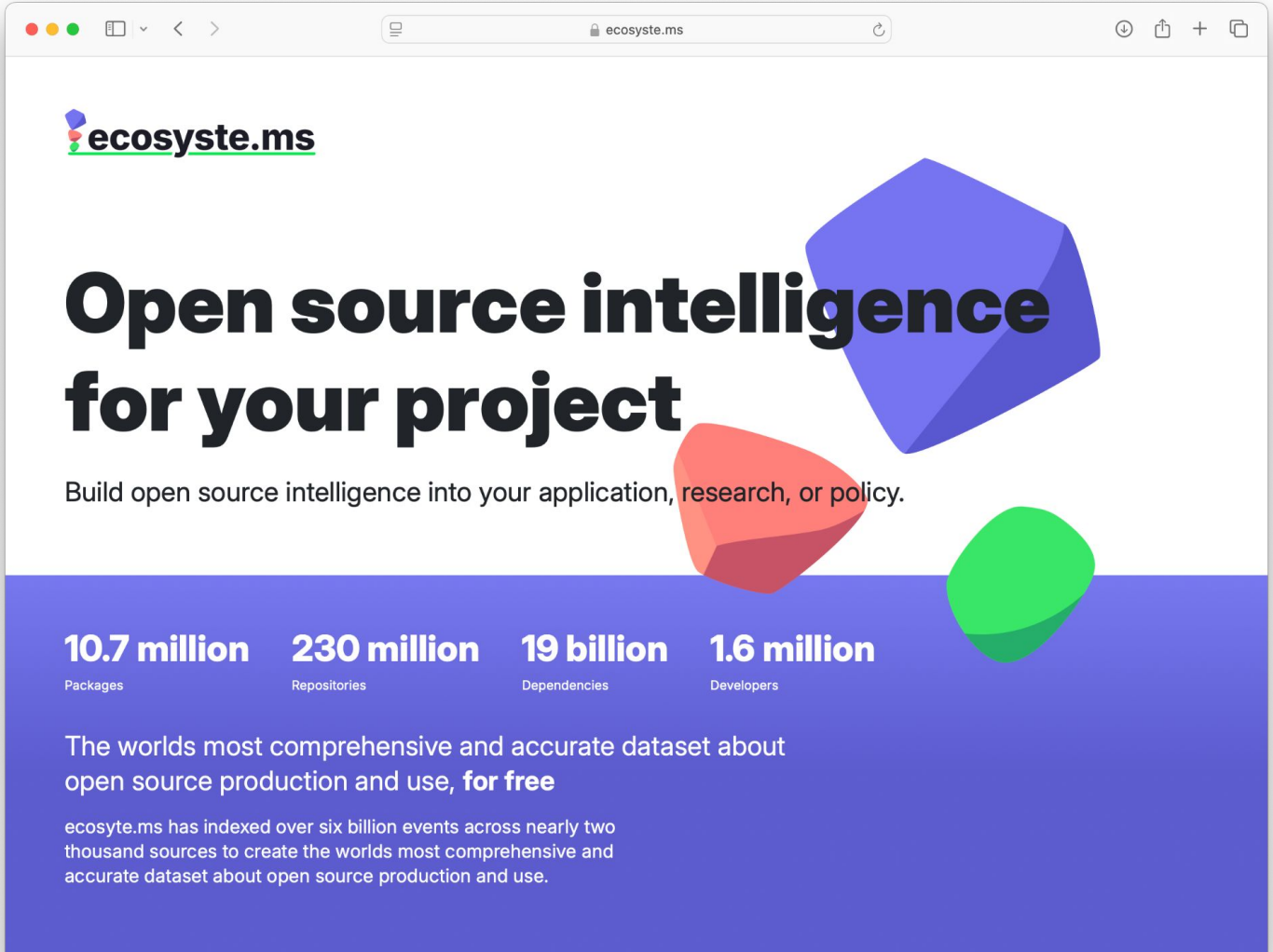
What are the most commonly used free and open source software (FOSS) packages? In this report, LF Research partnered with OpenSSF and the Laboratory for Innovation Science at Harvard to study the most common packages used at the application library level. With data from Software Composition Analysis partners FOSSA, Snyk, Sonatype, and Black Duck, the investigators captured detailed results of FOSS usage and key patterns of this usage that will help enhance the security of these packages.

DOWNLOAD REPORT

Authors

You Will Never Boil That Ocean

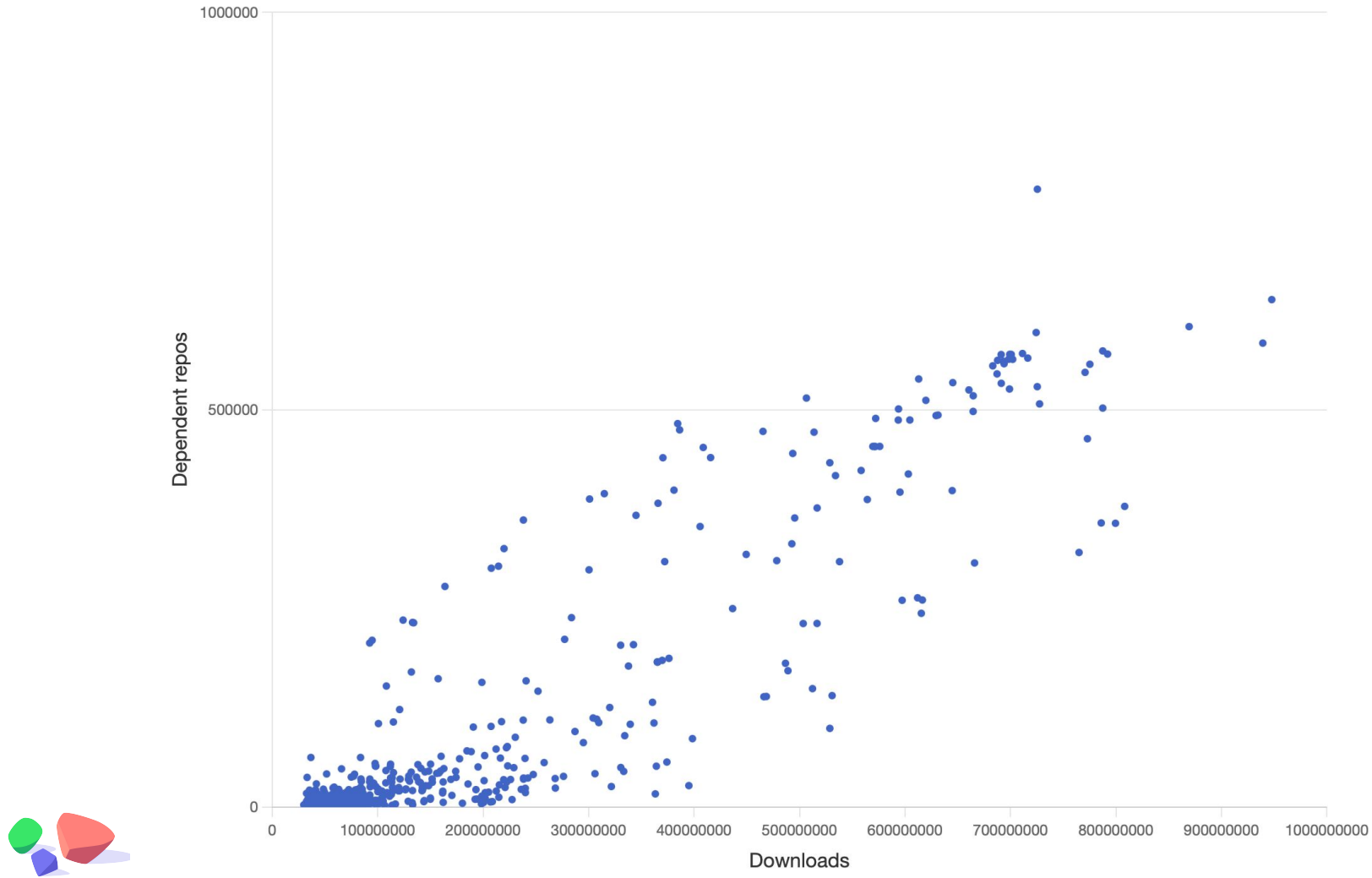
Can we save ourselves some time?

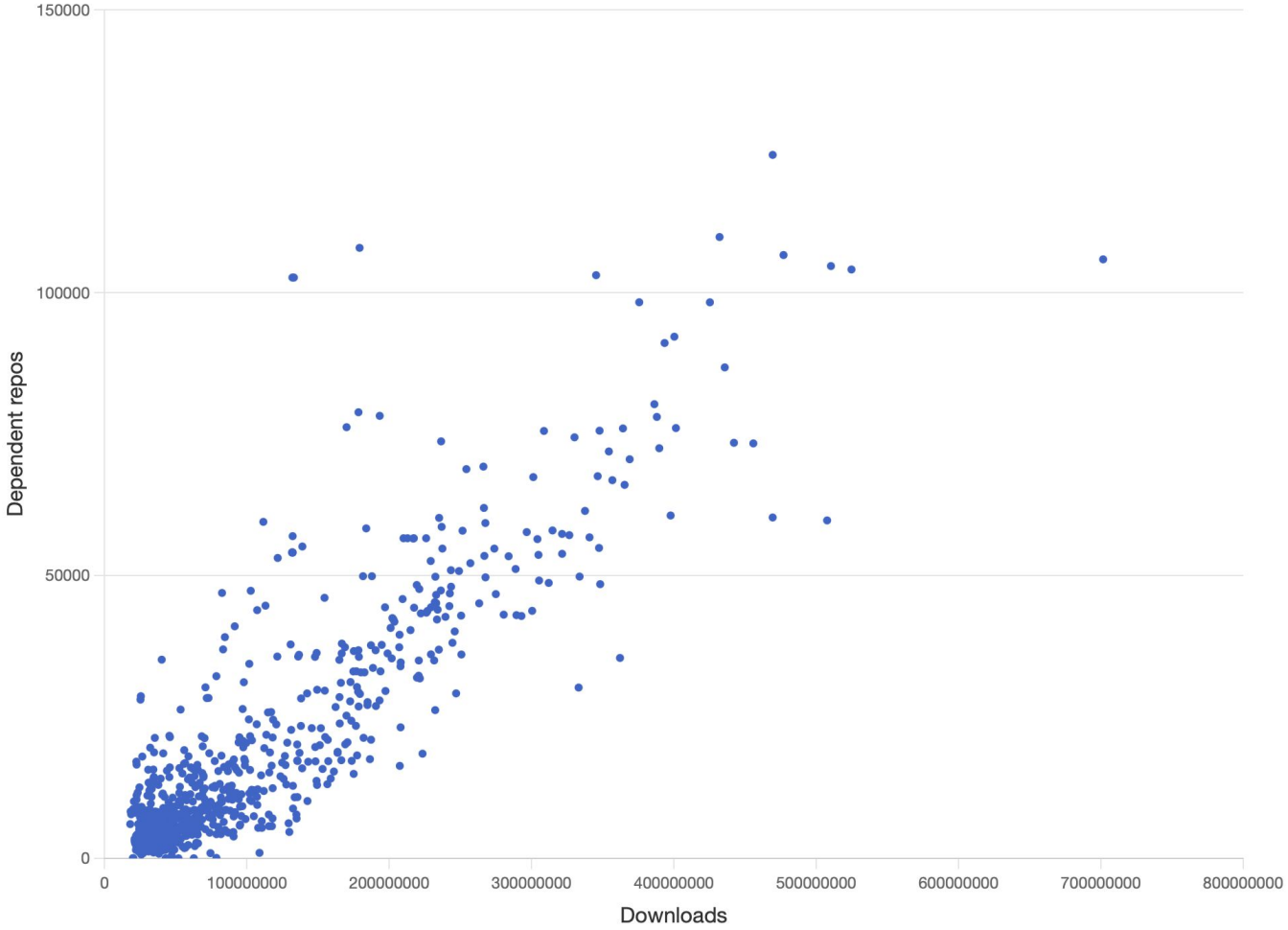

ecosyste.ms
ecosyste.ms
Open source intelligence for your project
Build open source intelligence into your application, research, or policy.
10.7 million
Packages
230 million
Repositories
19 billion
Dependencies
1.6 million
Developers
The worlds most comprehensive and accurate dataset about open source production and use, for free
ecosyte.ms has indexed over six billion events across nearly two thousand sources to create the worlds most comprehensive and accurate dataset about open source production and use.

20 Billion Data Points

Can we prove it?

TLDR: Yes


1000000
500000
0
Dependent repos
0
100000000
200000000
300000000
400000000
500000000
600000000
700000000
800000000
900000000
1000000000
Downloads


150000
100000
50000
0
Dependent repos
0
100000000
200000000
300000000
400000000
500000000
600000000
700000000
800000000
Downloads

Pearson Coefficient

Correlation	Measure
Very strong correlation	0.8-1.0
Strong correlation	0.6-0.79
Moderate correlation	0.4-0.59
Weak correlation	0.2-0.39

Correlation

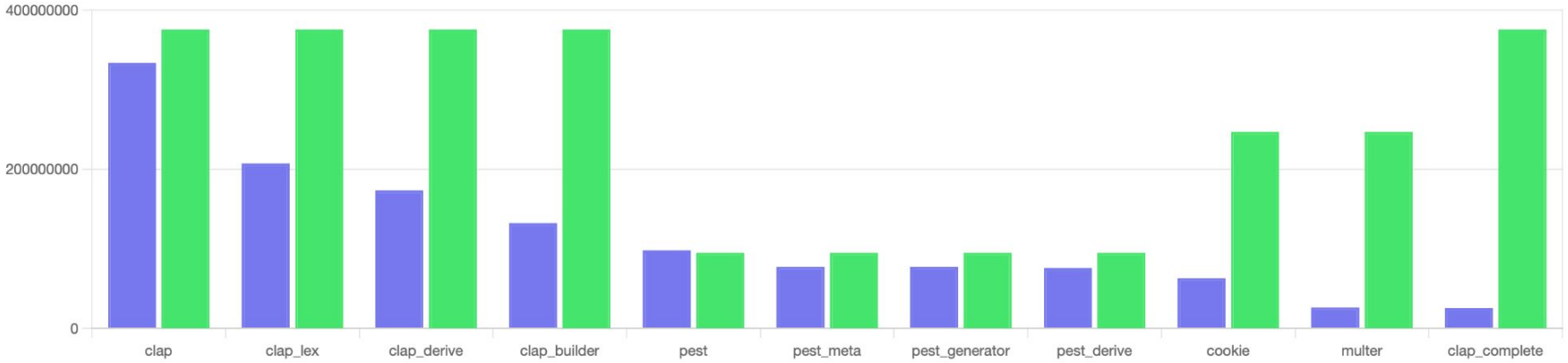
Registry	Metric	Correlation
Crates.io	Dependent repos	0.92
Packagist	Dependent repos	0.91
pypi.org	Dependent repos	0.66
npmjs	Dependent repos	0.59

Formula?

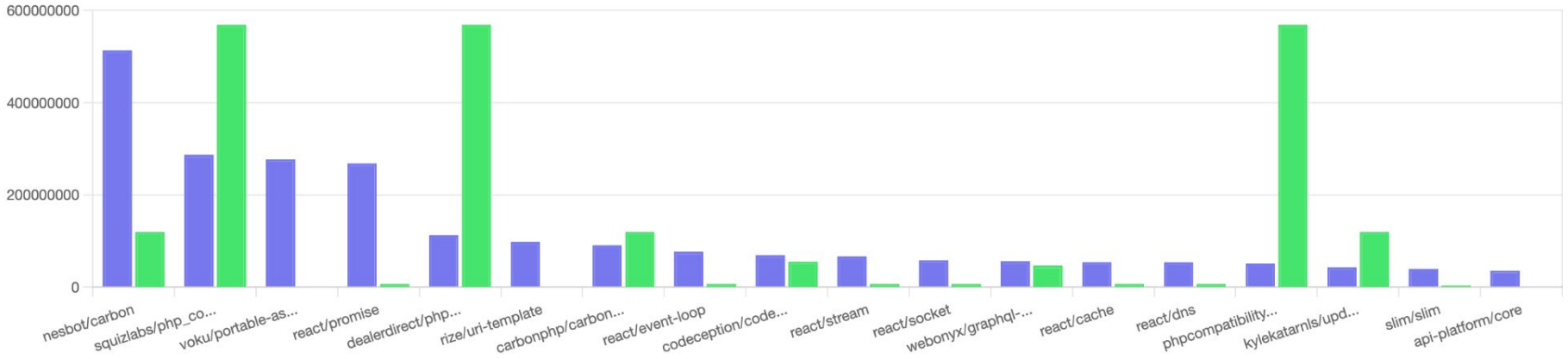
Part 2: Funding for ‘critical infrastructure’ is misdirected

Dearth of Data

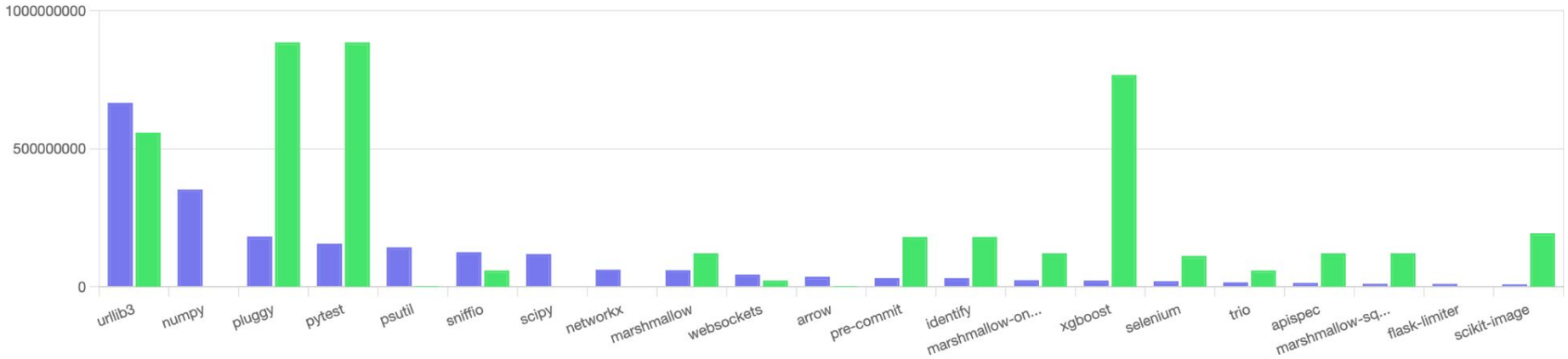
Ecosystem: cargo



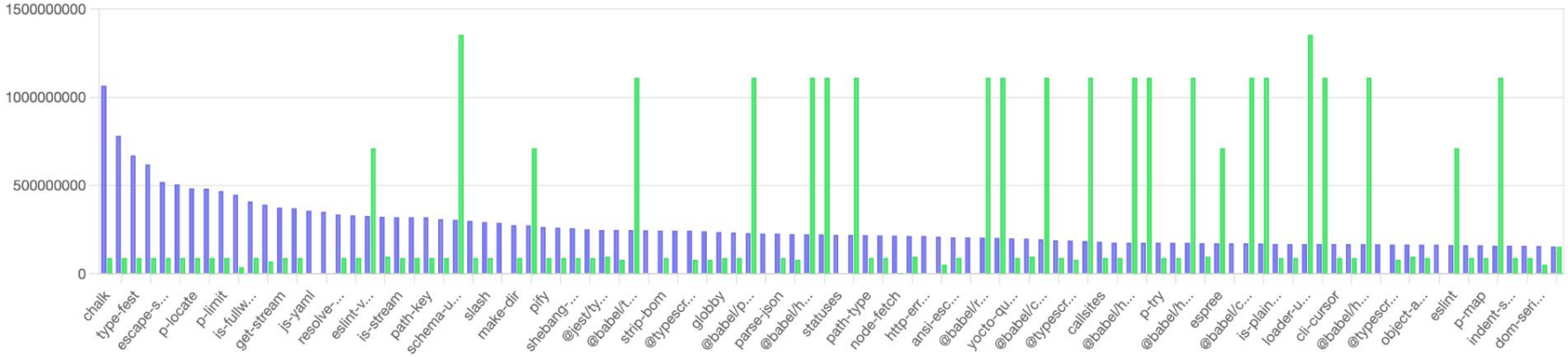
Ecosystem: packagist



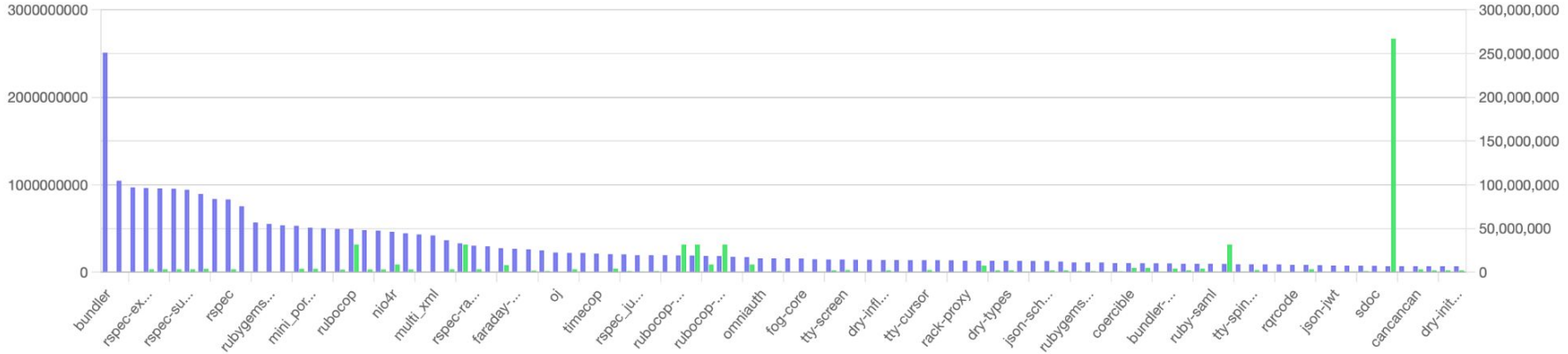
Ecosystem: pypi



Ecosystem: npm



Ecosystem: ruby



Inconclusive

After a decade saying we need to support our digital infrastructure, are we any closer to doing so?

Nope

What we need is...

The Value of Open Source Software

Harvard Business School Strategy Unit Working Paper No. 24-038

42 Pages • Posted: 16 Jan 2024

Manuel Hoffmann
Harvard Business School

Frank Nagle
Harvard Business School

Yanuo Zhou
University of Toronto

Date Written: January 1, 2024

Abstract

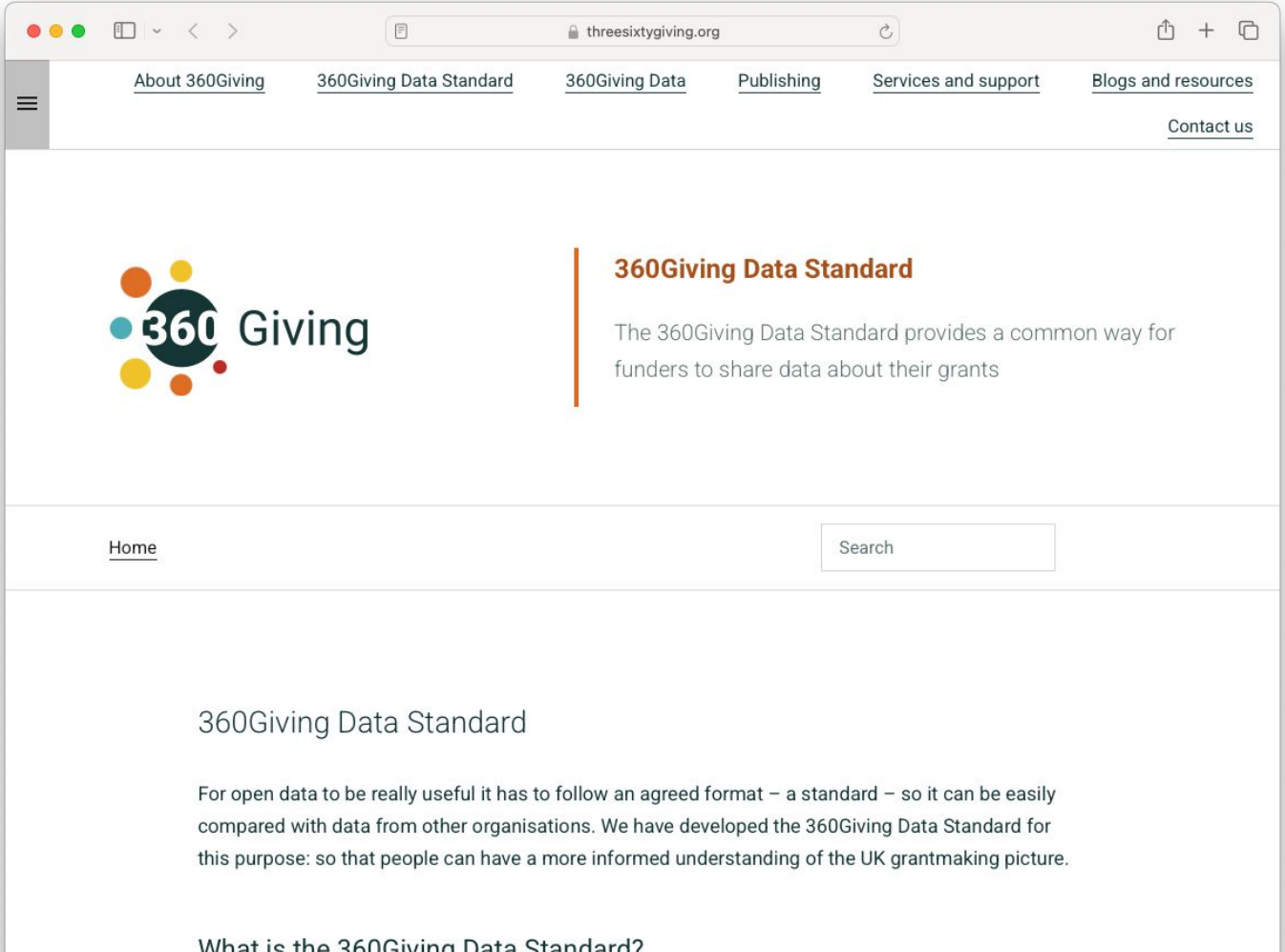
The value of a non-pecuniary (free) product is inherently difficult to assess. A pervasive example is open source software (OSS), a global public good that plays a vital role in the economy and is foundational for most technology we use today. However, it is difficult to measure the value of OSS due to its non-pecuniary nature and lack of centralized usage tracking. Therefore, OSS remains largely unaccounted for in economic measures. Although prior studies have estimated the supply side costs to recreate this software, a lack of data

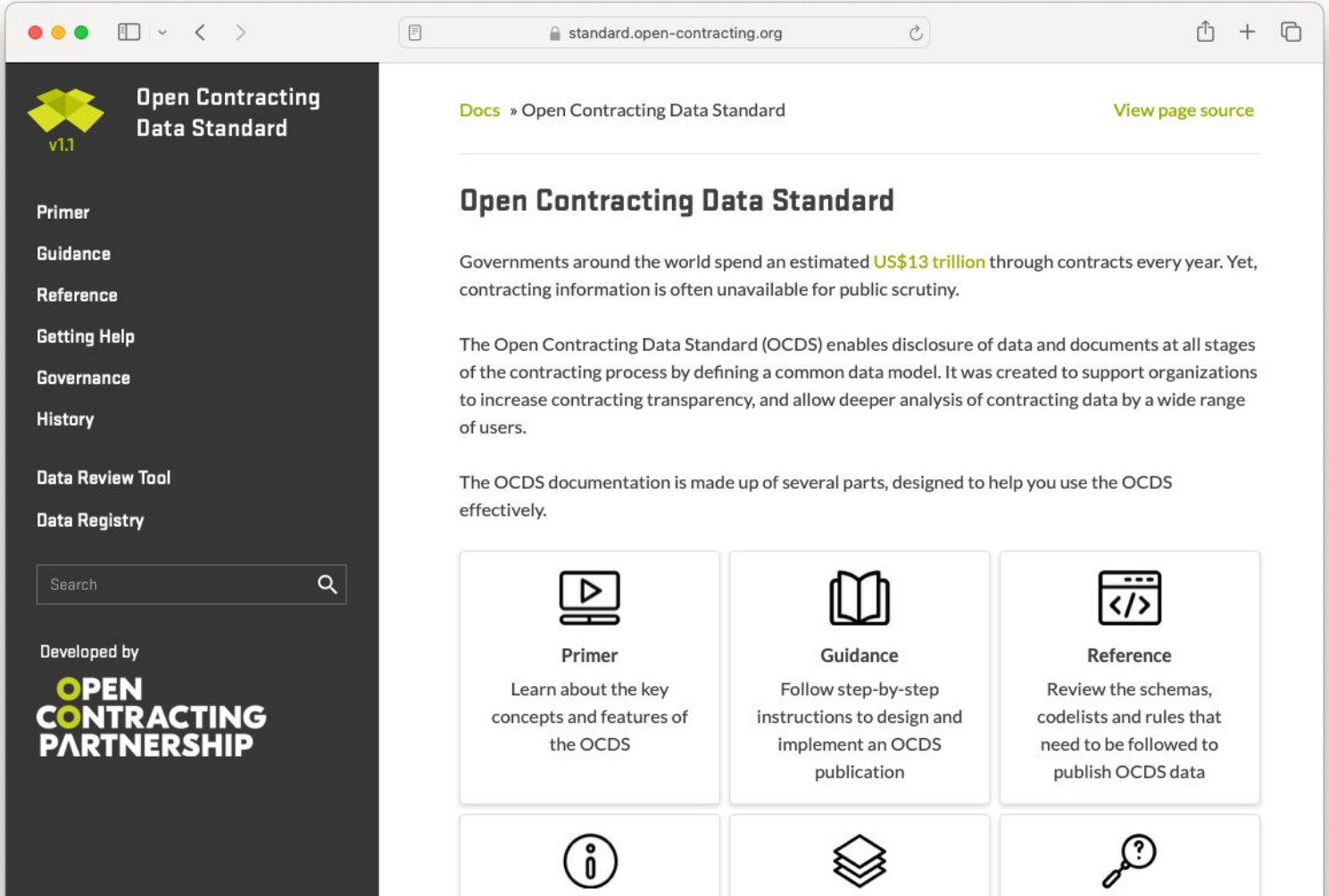
The state of open infrastructure grant funding

Introduction

At IOI, our work to increase investment in open infrastructure (OI) relies on a deep understanding of how infrastructure is funded and by whom. Since 2020, we've studied, analysed, and published our findings about the funding landscape for OI for research and scholarship. We recognize that grant funding is just one revenue stream that supports open infrastructure, but as our research has shown, financial contributions (which include grants) are the primary source of revenue for many OIs.[1] Here, we take our deepest dive yet into the available data in order to better understand the amount, impact, and limitations of grant funding to OIs. To the limited extent possible, we also try to position this analysis within an overall picture of the financial operations of open infrastructures.

The dataset we have assembled includes 514 awards made by 23 distinct funders to 36 open infrastructures (OIs), totaling US$415,845,753, and dating back to 2000. Of these awards, we estimate 149 of them (totaling $174,491,754) represent direct support to OIs, and the remainder support activities which depend on that infrastructure in some measure. This work greatly expands IOI's 2022 analysis of 137 funder-reported grant awards (for a total of $124,972,660) made to 28 key infrastructures and data from 22 funding organizations (Dunks, 2022) including ten members of the Open Research Funders Group. In updating our


threesixtygiving.org
About 360Giving
360Giving Data Standard
360Giving Data
Publishing
Services and support
Blogs and resources
Contact us
360 Giving
360Giving Data Standard
The 360Giving Data Standard provides a common way for funders to share data about their grants
Home
Search
360Giving Data Standard
For open data to be really useful it has to follow an agreed format – a standard – so it can be easily compared with data from other organisations. We have developed the 360Giving Data Standard for this purpose: so that people can have a more informed understanding of the UK grantmaking picture.
What is the 360Giving Data Standard?


standard.open-contracting.org
Open Contracting Data Standard
v1.1
Primer
Guidance
Reference
Getting Help
Governance
History
Data Review Tool
Data Registry
Search
Developed by
OPEN CONTRACTING PARTNERSHIP
Docs » Open Contracting Data Standard
View page source
Open Contracting Data Standard
Governments around the world spend an estimated US$13 trillion through contracts every year. Yet, contracting information is often unavailable for public scrutiny.
The Open Contracting Data Standard (OCDS) enables disclosure of data and documents at all stages of the contracting process by defining a common data model. It was created to support organizations to increase contracting transparency, and allow deeper analysis of contracting data by a wide range of users.
The OCDS documentation is made up of several parts, designed to help you use the OCDS effectively.
Primer
Learn about the key concepts and features of the OCDS
Guidance
Follow step-by-step instructions to design and implement an OCDS publication
Reference
Review the schemas, codelists and rules that need to be followed to publish OCDS data

What if we had the data?

More to come...

Questions?

🌐 https://ecosyste.ms

✉️ hello@ecosyste.ms

🐙 github.com/ecosyste-ms