Mapping the Open Source Ecosystem for Climate Science and Sustainable Technology

Richard Littauer[1], Tobias Augspurger[2], Alex Stinson[3], Pierre Vogler-Finck[4], Andrew Nesbitt[5], Chris Harris[6]

[1] OpenSustain.tech; Te Herenga Waka Victoria University of Wellington; SustainOSS; CURIOSS. Te Whanganui-a-Tara Wellington, Aotearoa New Zealand. Email: richard.littauer@ecs.vuw.ac.nz https://orcid.org/0000-0001-5428-7535
[2] OpenSustain.tech; Open Energy Transition. Aachen, Germany. Email: tobias.augspurger@openenergytransition.org. https://orcid.org/0000-0003-1744-1112
[3] OpenSustain.tech; Data Uruguay. Progreso, Uruguay. 0009-0000-9718-8801
[4] OpenSustain.tech; PierreVF UG. Freiburg, Germany. https://orcid.org/0000-0002-3893-8956
[5] OpenSustain.tech; Ecosyste.ms. United Kingdom. https://orcid.org/0009-0007-2710-1118
[6] OpenSustain.tech; Ground Truth Data. Vancouver, Canada. https://orcid.org/0009-0008-2868-6338

October 2025

The climate crisis poses a severe threat to the natural systems that support modern civilization, disrupting essential cycles that provide freshwater, fertile soils, and stable weather patterns. These disruptions are projected to lead to widespread biodiversity loss and to upset local and global economies. To ensure that the scientific basis of these projections is transparent and credible, researchers globally are increasingly making climate data and models openly available following best practices from the open science, software, and data communities. This openness supports informed decision-making and helps safeguard sustainable development from being compromised by short-term political or economic agendas.

Despite this progress in open science, the broader application of open source software and open data in climate and sustainability-related technologies remains limited. National governments, international organizations, academia, industry, and civil society have all played roles in both contributing to the crisis and proposing solutions. However, fragmented, proprietary approaches persist. Open source offers a powerful alternative: lowering costs, enhancing verifiability, and enabling collaboration across disciplines and sectors.

In this paper, we introduce OpenSustain.tech, the most comprehensive dataset of over 2,500 open source projects directly addressing the climate crisis. We detail the transparent methodology used to curate this collection, including human expert review across multiple fields. We further analyze the network of transitive dependencies among these projects, extending previous work in mapping the climate-focused open source ecosystem.

Finally, we discuss the strategic importance of open source in advancing climate solutions, including its potential economic and societal value. By building shared digital infrastructure, we argue that open source can play a foundational role in climate mitigation, adaptation, and sustainable development.

Table of Contents

Introduction

The climate crisis poses a severe threat to the natural systems that support modern civilization, disrupting essential cycles that provide freshwater, fertile soils, biodiverse ecosystems, and stable weather patterns. These disruptions result in human deaths, widespread biodiversity loss and large scale disruptions of society at all levels, politically, economically and demographically[1]. The disruptions are projected to get worse. As key international figures like the Intergovernmental Panel on Climate Change (IPCC) and the UN Secretary General point out, all possible parts of society need to work together to address the urgent climate crisis.

Open work can help answer this call. Open work, or open organizing,[2] is a new term encompassing different fields:[3] open science, open data, open models, open access, open knowledge and open source, among others. There is progress in each of these subfields. For instance, to ensure that the scientific basis of climate models and other projections are transparent and credible, researchers globally are increasingly making climate data and models openly available.[4] This openness supports informed decision-making and may help safeguard sustainable development from being compromised by short-term political or economic agendas.

Open access publishing, the restriction-free ability to download and read scientific publications, has gained some traction in recent years. However, as a study in Carbon Brief evaluated, even in the IPCC Sixth Assessment report, still, only half of its top-cited journals are "open access by default". This openness is also not universal, as much of the content

---

[1] IPBES (2019): Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. S. Díaz, J. Settele, E. S. Brondízio, H. T. Ngo, M. Guèze, J. Agard, A. Arneth, P. Balvanera, K. A. Brauman, S. H. M. Butchart, K. M. A. Chan, L. A. Garibaldi, K. Ichii, J. Liu, S. M. Subramanian, G. F. Midgley, P. Miloslavich, Z. Molnár, D. Obura, A. Pfaff, S. Polasky, A. Purvis, J. Razzaque, B. Reyers, R. Roy Chowdhury, Y. J. Shin, I. J. Visseren-Hamakers, K. J. Willis, and C. N. Zayas (eds.). IPBES secretariat, Bonn, Germany. 56 pages. https://doi.org/10.5281/zenodo.3553579

[2] Meluso, John and Chambers, Cassandra R. and Littauer, Richard and Llamas, Nerea and Long Lingo, Elizabeth and Mhangami, Marlene and Pitt, Beck and Splitter, Violetta and Wang, Huajin, Opening Up: Interdisciplinary Guidance for Managing Open Ecosystems (May 8, 2024). Available at SSRN: https://ssrn.com/abstract=4821969 or http://dx.doi.org/10.2139/ssrn.4821969

[3] https://openworkdefinition.com/

[4] Stockhause, Martina, David Huard, Alaa Al Khourdajie, et al. 2024. 'Implementing FAIR Data Principles in the IPCC Seventh Assessment Cycle: Lessons Learned and Future Prospects'. *PLOS Climate* 3 (12): e0000533. https://doi.org/10.1371/journal.pclm.0000533.

comes from authors in the Global North, and the cost structure of open access publishing is prohibitive for scholars working in the Global South.[5]

Open source software has seen marked growth in the climate sector in the last few years. Open source is normally defined as the public release of software under licenses which allow for easy use, modification, and dissemination. Many open source licenses are only considered as being "open source" if they match the Open Source Definition, provided by the non-profit Open Source Initiative.[6] These licenses include the MIT license, the GPL, Apache license, and other popular licenses, now used globally. The MIT license, for example, was created in 1988,[7] and according to a 2020 report, it was used by over a million software projects.[8] GitHub notes that it was used by around a third of all licensed repositories on their platform.[9] Another way to understand the spread of open source is to consider the impact of open source on society. One working paper from Harvard Business School noted that the supply side of OSS was worth 4 billion USD, while the demand was an order of magnitude more, closer to 9 trillion USD.[10]

The impact on the environmental sector is less directly clear. The software, scientific literature, and data needed for sustainable development are not widely available, especially in contexts where people and polluting industries are most affected, such as in the Global South and in communities affected by epistemological, environmental or other social justice issues.[11][12] Moreover, there are large gaps in the overall use of open source software and data across the larger climate tech business sector, despite the strong reliance on both

---

[5] Chavelli, Felix, and Sarah Connors. "Analysis of the WGI Contribution to the Sixth Assessment Report: Review of the WGI AR6 References". Zenodo, February 7, 2023. https://doi.org/10.5281/zenodo.7615825.

[6] https://opensource.org/

[7] https://opensource.com/article/19/4/history-mit-license

[8] https://web.archive.org/web/20200503111426/ https://resources.whitesourcesoftware.com/blog-whitesource/top-open-source-licenses-trends-and-predictions

[9] https://innovationgraph.github.com/global-metrics/licenses. GitHub itself has 420 million repositories according to their website, https://github.com/about, (Accessed on October 14, 2025); a large subset of these will be licensed repositories.

[10] Hoffmann, Manuel, Frank Nagle, and Yanuo Zhou. "The Value of Open Source Software." Harvard Business School Working Paper, No. 24-038, January 2024.

[11] *Dosemagen, Shannon; Emilio Velis; Luis Felipe R. Murillo; Evelin Heidel; Michelle Thorne; and Alex Stinson.* "Open Climate Then and Now." Branch Magazine Issue 4.

[12] *Dosemagen, Shannon; Emilio Velis; Luis Felipe R. Murillo; Evelin Heidel; Michelle Thorne; and Alex Stinson.* "Open Climate Now." Branch Magazine Issue 2.

public data created by governments and prior work by scholars.[13] There has been an increased call for more alignment around policies for working in the open, such as using open source technology,[14] either through open science practices or through the Digital Public Goods program,[15] and the global call for a fee-free (diamond) Open Access publishing ecosystem.[16]

Open source work, like science, is built by standing on the shoulders of others. Open source software by definition is reusable and accessible to many people aside from the original creators. This reusability allows for accelerated development, as creators can build upon prior innovations rather than starting from scratch. By analyzing projects for documentation that notes how it is used, and by tracking dependency trees in order to see where projects are part of the shared stack of different projects, it is possible to understand how open source software can be used for climate change without necessitating public releases through official PR channels or after waiting for policy changes from large institutions.

We use this method for identifying and enabling access and use of open source tools for climate change. In this paper, we share OpenSustain.tech, the most comprehensive human-curated catalogue specifically for open source software projects, containing over 2,500 distinct utilities built for addressing the climate crisis. It is a community-driven, open source project focused on cataloguing useful software from both within and outside the academic ecosystem. In this paper, we analyze its origin and goals, its governance model and the contents of the repository. We will also describe how it contributes to the broader shift towards open approaches in addressing the climate crisis.

This ongoing work has been presented in several previous venues. The catalogue itself is publicly available, accessible, and open to edits at OpenSustain.tech. Further, a report in 2023 discussed its progress.[17] Here, we seek to expand on that work.

---

13 Stinson, Alex (*May 05, 2025*). "It's time for you to contribute to the Climate Commons." *Climate Drift.*
14 https://www.unesco.org/en/open-science/about?hub=686
15 https://www.digitalpublicgoods.net/
16 See the 2025 Toluca-Cape Town Declaration: https://sparcopen.org/news/2025/toluca-cape-town-declaration-is-a-milestone-for-recognizing-science-as-a-public-good/
17 Tobias Augspurger, Eirini Malliaraki, Josh Hopkins, and Dan Brown. 2023. The Open Source Sustainability Ecosystem. The Linux Foundation. https://zenodo.org/records/7771633

Why build OpenSustain.tech?

The Strategic Importance of Open Work

Access to information is a measure of power. Traditionally, access has been limited to elites, who depend on information asymmetry to maintain power. The scientific revolution challenged this dynamic by replacing dogma with open methods and shared evidence. Throughout history, from independence movements to modern governance, transparency has functioned as a corrective mechanism against institutional authority that depends on obscured information. Open data and open source code can help in democratizing information flows, creating the same feedback loops that make markets efficient, expose errors, and reward accuracy.

The use of open data to invert asymmetrical power differentials applies directly to climate solutions across multiple sectors. In carbon markets, inconsistent data and unverifiable claims have enabled greenwashing and double-counting.[18] In finance, hidden risks lead to mispriced assets and systemic vulnerability. In cryptocurrency and unregulated sectors, a lack of disclosure enables speculation and fraud. In every case, opacity creates a moral hazard that undermines efficient resource allocation and public trust.

The corrective mechanism is straightforward: bring code and data into the open. When diverse actors with competing interests and actual stakes in outcomes have access to verifiable information, error is exposed and accuracy is rewarded. Scholars across different academic departments and ideological traditions recognize this principle. For instance, Hayek argued that no authority can plan effectively without distributed information.[19] Beck showed that societies cannot manage complex risks without open, ongoing data analysis.[20] Taleb demonstrated that asymmetries in information and consequence, where some actors are insulated from the results of their decisions while others bear the costs, systematically distorts judgment and enables catastrophic failures.[21] All identified the same conclusion: transparency and diverse perspectives are prerequisites to effective systems.

---

[18] Greenfield, P. (2023). Revealed: more than 90% of rainforest carbon offsets by biggest certifier are worthless, analysis shows. The Guardian. https://www.theguardian.com/environment/2023/jan/18/revealed-forest-carbon-offsets-biggest-provider-worthless-verra-aoe
[19] Hayek, F. A. (1956). The Road to Serfdom (Preface). University of Chicago Press.
[20] Beck, U. (1992). Risk Society: Towards a New Modernity. Sage Publications.
[21] Taleb, N. N. (2007). The Black Swan: The Impact of the Highly Improbable. Random House.

For climate solutions specifically, open methodologies and shared datasets are operating requirements. Decisions made without openness lead to inefficiency, duplication, and erosion of public trust. Openness is a key indicator of measurable sustainability. Without verifiable information about emissions, climate models, methodologies, and trade-offs, we cannot distinguish progress from performance.

Background Context for Open Source and Open Access

Openness is already prevalent in technical fields. Open access and open source were instrumental in the creation of modern computation. In the early 1990s, institutions like CERN and MIT created a digital culture that focused on sharing and reusing standard software to enhance interoperability and ease of software development. This culture later spread to industry, with Sun, Intel, and other major corporations utilizing open source. The meteoric rise of centralized repositories for storing and sharing open source, such as SourceForge and GitHub, afforded mass participation in open source, such that some open source code is now in use in all sectors. A Linux Foundation report from 2022 by the Linux Foundation estimates that between 70-90% of software includes open source components,[22] and a 2025 Black Duck industry report found that 97% of codebases in their study had OSS software components.[23]

Although open source software may be licensed permissively, gathering metrics on usage is difficult. Partially, this is by design: there are no restrictions on end-users for how to use open source software. Even software which mandates that derivative code is licensed similarly, such as with copyleft licenses like GPL, does not have restrictions or mandates on reporting usage. Gathering data of usage relies on proxies.

One such qualitative proxy involves looking at other, similar aspects of open work. Although software has been the most widespread visible success of the open movement outside academic and cultural institutions, much of society has been shaped by the open movement as well.

[22] https://www.linuxfoundation.org/blog/blog/a-summary-of-census-ii-open-source-software-application-libraries-the-world-depends-on

[23] 2025 Open Source Security and Risk Analysis Report

Open access to scholarly literature started in the 90s alongside the early internet, with a focus on sharing and reusing scholarly literature. Following the creation of the OAI-PMH standard and the Creative Commons license framework, open scholarly publishing took off in the early 2000s,[24] highlighted by significant declarations like the Budapest Open Access Initiative.[25] At this time, the open-access model spread across many other sectors: developing freely available educational materials (Open Educational Resources), licensing media at cultural institutions like museums (Open GLAM), academic publications (Open Access to Science), and outputs from governments (Open Government Data).

This broader open movement, alongside its partners in the larger digital rights ecosystem, has entered a new era in the last decade, as the rise of open source, AI and increasingly complex algorithmic curation of the internet has created a large-scale reliance on "open" infrastructures.[26] Even now, tools like Wikipedia or OpenStreetMap that were built as a product of these open knowledge practices have become foundations for AI and other digital infrastructure.[27] Most frameworks for tracking open, such as the Digital Public Goods framework, distinguish between approaches for open access data, publishing, software and increasingly AI models.[28] Each requires evaluating that content against the larger principles, such as FAIR Data[29] or the Open Definition,[30] while also meeting additional criteria for the specific domain or use case.

Sector-specific open access practices are critical because they each have different needs, and sector-specific standards greatly affect the adoption of technology. For example, widespread adoption of open access to medical research on vaccines and genetic data

---

24 https://open-access.network/en/information/open-access-primers/history-of-the-open-access-movement

25 https://www.budapestopenaccessinitiative.org/faq/

26 Eghbal, N. (2016). Roads and bridges. *The Unseen labor behind our digital infrastructure*.

27 Zuckerman, Ethan. "The case for digital public infrastructure." (2020).

28 "Digital Public Goods Standard"

29 Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten JW, da Silva Santos LB, Bourne PE, Bouwman J, Brookes AJ, Clark T, Crosas M, Dillo I, Dumon O, Edmunds S, Evelo CT, Finkers R, Gonzalez-Beltran A, Gray AJ, Groth P, Goble C, Grethe JS, Heringa J, 't Hoen PA, Hooft R, Kuhn T, Kok R, Kok J, Lusher SJ, Martone ME, Mons A, Packer AL, Persson B, Rocca-Serra P, Roos M, van Schaik R, Sansone SA, Schultes E, Sengstag T, Slater T, Strawn G, Swertz MA, Thompson M, van der Lei J, van Mulligen E, Velterop J, Waagmeester A, Wittenburg P, Wolstencroft K, Zhao J, Mons B. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data. 2016 Mar 15;3:160018. doi: 10.1038/sdata.2016.18. Erratum in: Sci Data. 2019 Mar 19;6(1):6. doi: 10.1038/s41597-019-0009-6. PMID: 26978244; PMCID: PMC4792175.

30 Created by the Open Knowledge Foundation; see https://opendefinition.org/.

meant that the scientific infrastructure was available for a much more rapid global response to the COVID-19 pandemic. [31]

Open in Climate

For climate, there hasn't been a widespread push for establishing these standards across all of the impacted sectors. In part because of the diversity of the sectors, but also because it affects a number of interdisciplinary areas that mix different expectations for publishing. Additionally, climate technology mitigation and environmental policies affect industries, such as fossil fuels[32] and agriculture,[33] which often rely on obfuscation and proprietary technologies to advance their economic activity. Openness would make these industries accountable and transparent in their statements and decision-making, hindering short-term profits for stakeholders from unsustainable business practices.

In climate science, much of the historical climate data and climate models in use in the academic community are produced by governments that meet many of the standards for open data.[34] The role of the US federal government in creating basic science related to climate and weather, through the National Weather Service and NASA, has made much of this data "open by default" as part of the larger public domain standard applied to US Federal Government Works.[35] However, with the increasing removal of this data from public availability by the Trump administration, the larger climate science ecosystem will need to rely on volunteers, philanthropic non-profit funding, and sustainable for-profit organizations that can collaborate with other Governments around the world. However, such money also frequently flows into closed platforms, such as ClimateTrace, which

[31] See source documented here: https://opendatawatch.com/whats-being-said-resource/data-in-the-time-of-covid-19/
[32] Denial, disinformation, and doublespeak: big oil's evolving efforts to avoid accountability for climate change
[33] Amentae, Tadesse K.; Song, Wei; Wang, Jinjin (2024-06). "Intellectual property rights in the agri-food chains: A systematic review and bibliometric analysis". *World Patent Information*. 77: 102279. doi:10.1016/j.wpi.2024.102279.
[34] See this analysis by Creative Commons in 2023: Campbell,Taylor; Natalie Cheyette, and Wanying Li (2023) "A Landscape Analysis of Open Climate Data." Creative Commons
[35] The Global Effort for Open Access to Environmental Satellite Data https://en.wikipedia.org/wiki/Copyright_status_of_works_by_the_federal_government_of_the_United_Stateguideliness

provides open data but does not facilitate collaborative work on open platforms and methods.[36]

Sectors that are important for addressing climate change and sustainable development goals still heavily depend on closed-source, for-profit software for historical reasons. These utilities are critical sources of climate data and a means to plan and mitigate climate impacts. In recent years, organizations like the Linux Energy Foundation and Open Energy Transition have had to cultivate new open standards for interoperability between energy operators and for energy planning data, respectively. It was this context which led to the partnership resulting in the publication of the previous OpenSustain.tech report, as there was a need to examine the overall ecosystem of available software around open source environmental projects.[37] This need continues, as initiatives like LF Energy continue to see buy-in from major stakeholders in the ecosystem,[38] and as groups such as the GitHub Sustainability Team have worked together with OpenSustain.tech to forward climate goals.[39]

---

[36] See the mixed copyright terms at https://climatetrace.org/terms and the closed-consortium model of creating the collaboration at: https://climatetrace.org/approach
[37] Tobias Augspurger, Eirini Malliaraki, Josh Hopkins, and Dan Brown. 2023. The Open Source Sustainability Ecosystem. The Linux Foundation. https://zenodo.org/records/7771633
[38] Projects - LF Energy. https://lfenergy.org/our-projects/
[39] Climate Action Plan for Developers.
https://github.com/social-impact/focus-areas/environmental-sustainability/climate-action-plan-for-developers

OpenSustain.tech

OpenSustain.tech was established in 2020 through the course of several discussions within the Protontypes community, an organization dedicated to the promotion and support of open and sustainable technology.[40] The transformative impact of open source and data on sectors such as robotics, AI and cloud computing is well documented, and it is evident that there is significant potential for climate technology and environmental science.

As with other open source ecosystems, this one suffers from a problem of discovery and marketing. The lack of online marketing and internet-wide search engines for open source projects makes discovering them a challenge in itself. The large number of low-quality personal open source projects, combined with GitHub prioritizing their GitHub stars metric[41] over activity, means that developers, users, and investors are presented with inactive projects rather than new, promising developments. GitHub stars may be a useful proxy for development,[42] though it is also ripe for abuse.[43] Beyond stars, the frequent use of keywords such as 'climate', 'cloud', 'environment', and 'carbon' within the broader open source ecosystem for projects which are not directly related to the environment is particularly problematic for the open source ecosystem in sustainability, as it contributes to many false positives in searches.

Early discussions aligned on a solution to the problem of difficult-to-find projects – a service that lists strong open source environmental projects, enabling collaboration and reuse. This list-based approach, using human curation, is already popular within the GitHub ecosystem, as shown by the awesome-lists community.[44] Similar developments with a climate focus, such as for Digital Public Goods,[45] focus on large-scale end-user products, covering only a small part of the ecosystem, and do not normally include low-level dependencies.

---

[40] Viewable at https://github.com/protontypes/

[41] This refers to the practice of "starring" or favoriting repositories, not to the GitHub Stars program for highlighting individual developers. https://stars.github.com/

[42] Koch, S., Klein, D., & Johns, M. (2024). The Fault in Our Stars: An Analysis of GitHub Stars as an Importance Metric for Web Source Code. In *Workshop on Measurements, Attacks, and Defenses for the Web (MADWeb)* (Vol. 2024).

[43] He, H., Yang, H., Burckhardt, P., Kapravelos, A., Vasilescu, B., & Kästner, C. (2024). 4.5 Million (Suspected) Fake Stars in GitHub: A Growing Spiral of Popularity Contests, Scams, and Malware. *arXiv preprint arXiv:2412.13459*.

[44] https://github.com/topics/awesome-list

[45] https://www.digitalpublicgoods.net/

The project tries to collect all software that meets the following guideline and is open source. The mission statement is reproduced below:

> *It is a miracle that our planet has developed a unique and stable environment for life in an otherwise indifferent and hostile cosmos. Life has taken billions of years to build up the natural resources we depend on, such as a protective atmosphere, fertile soil, stable weather, and clean drinking water. As a movement to democratize technological development and knowledge creation, open source has the potential to play a central part in preserving this stability. Open Sustainable Technology's mission is to find, list, and share projects that preserve or analyze natural ecosystems through open technology, methods, data, intelligence, knowledge or tools.*[46]

The dataset is stored in several ways: as a large Markdown file, README.md, which is stored in the root repository on GitHub and which renders when users visit the repository; and as a spreadsheet on the site getgrist,[47] or the same spreadsheet which renders on the project website.[48] The spreadsheet is populated from the Markdown list, but includes more information that is automatically added using scripts. Community members can contribute directly by making a Pull Request on GitHub. The data is available under a Creative Commons Attribution 4.0 International license.[49]

The project has grown far beyond the initial list, both in scale and in access points. Because the dataset lives on GitHub and can be called in different interfaces, there are further entry points to engage with the project. For instance, ClimateTriage[50] was set up to surface projects that are looking for contributions, for those who would like to contribute immediately to tooling related to climate change. OpenClimate.fund[51] is a funding platform which aims to distribute funding to the relevant projects. OSS for Climate[52] is a podcast

[46] https://github.com/protontypes/open-sustainable-technology/
[47] https://docs.getgrist.com/doc/gSscJkc5Rb1Rw45gh1o1Yc/p/3
[48] https://opensustain.tech/spreadsheet/
[49] CC-BY 4.0. https://github.com/protontypes/open-sustainable-technology/blob/main/LICENSE.md
[50] https://climatetriage.com/
[51] https://openclimate.fund/
[52] https://ossforclimate.sustainoss.org/

highlighting relevant maintainers and projects, made through a collaboration with the community SustainOSS and the Sustain podcast.[53]

Although several experts in open-source, climate, energy and sustainability participated in the contribution and review of projects, it soon became clear that we could not find the necessary expertise within the small community to judge the quality of all relevant domains. Therefore, a simplified review process was chosen. Rather than measuring project quality itself, the adoption and usage of open source projects within the domain is evaluated. The following main questions need to be answered to review an open source project:

1. Does a project provide an open source license?
2. Does the open source project demonstrate usage and goals that are relevant to environmental sustainability?
3. Does the project show active usage or development in the last year?
4. Does the project have users outside the organization, as evidenced by citations, opened issues, citations or downloads? Is the project documented in such a way that it can easily be reused?

These questions enable our community to evaluate the relevance of a project for the broader sustainability community without domain expertise in all domains. To systematically discover and map new projects, a combination of multiple methodologies has been applied,[54] including a growing dataset of keywords and occasional investigations of key missing topic areas.

Each new project is included in the dataset via a separate Pull Request, which must be reviewed by a core community member. Hard edge cases are discussed within the community and with the author of the pull request.

Currently sectors covered by the community data model focus on two large areas.

[53] https://podcast.sustainoss.org/

[54] https://github.com/protontypes/open-sustainable-technology/blob/main/docs/how_to_identify_projects.md

Tools that focus on human systems:

- Renewable Energy - tools target the entire system design of renewable energy systems
- Energy Storage - tools that target devices and methods for storing energy
- Energy Systems - tools that target every aspect of larger-scale energy systems
- Consumption - tools that model, predict, and optimize the consumption of resources
- Emissions - tools target how to identify, quantify, forecast, and reduce greenhouse gas emissions.
- Industrial Ecology - tools that track the flow of energy and materials from natural resources through their production, use and disposal
- Sustainable Development - tools that facilitate development that meets the needs of the present generation without compromising the ability of future generations to meet their own.

Tools that focus on parts of nature and natural resources:

- Biosphere - Ecosystems where life exists, extending from the deep ocean to the lower atmosphere
- Cryosphere - All portions of Earth's surface where water is in solid form
- Hydrosphere - The combined mass of water found on, above, and below the Earth's surface, including oceans, rivers, lakes, groundwater, and atmospheric moisture.
- Atmosphere - The layer of mixed gas surrounding the Earth, retained by gravity.
- Climate Change - The long-term alteration of Earth's climate, including average weather patterns and global temperatures.
- Natural Resources - Resources that stem from nature, including minerals, water, soil, forests, and fossil fuels.

Some open source projects may fit into multiple categories; in that case, the most relevant one was chosen.

Analysis

To systematically analyse the open source ecosystem described by OpenSustain.tech, metadata information from listed Git repositories are gathered using the ecosyste.ms open source intelligence platform developed by Andrew Nesbitt.[55] While this allows us to analyze Git repositories on various open source development platforms such as Codeberg and GitHub, our analysis is limited to open source projects that are maintained within Git repositories. This ensures the consistency of our database. Other websites or platforms that enable open source or open-data collaboration via other platforms are excluded. It is possible that many relevant open source projects, which may be better maintained, more collaborative, and more academic,[56] are excluded from our analysis.

The October release of the datasets reveals a total number of 2497 listed projects, where 2372 projects are hosted on GitHub. 70% of the projects had a commit in the last year and are considered active. At the time of writing the previous report in 2022, we had 1339 total projects with 75% active. For OpenSustain.tech, in total we have 93 contributors with a core community of around 10 people regularly contributing to the project.

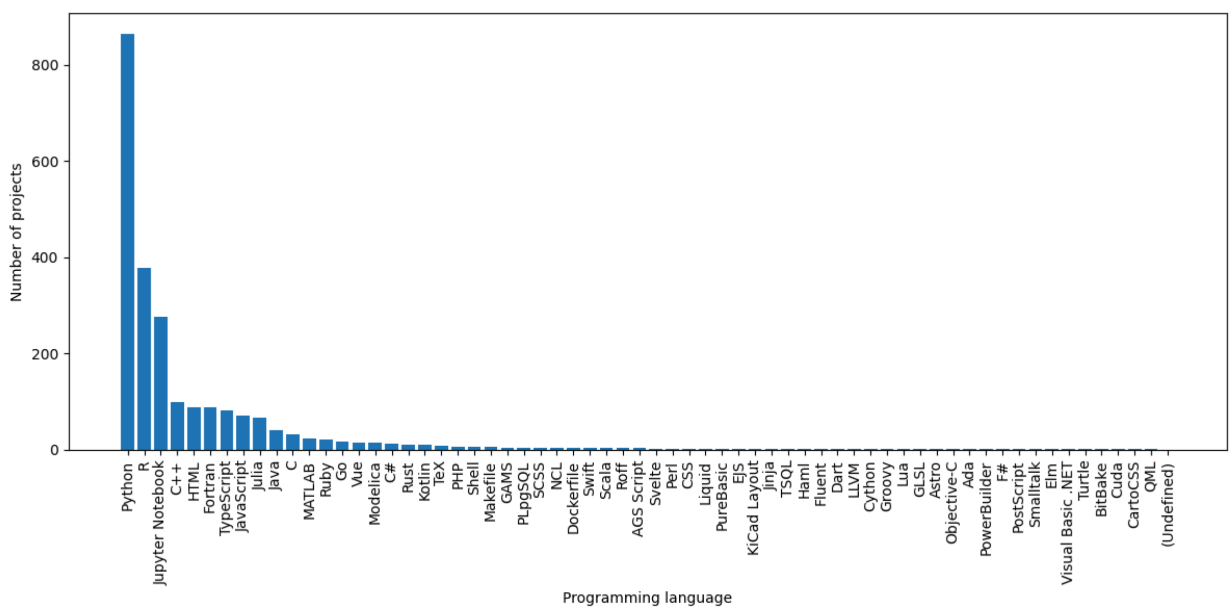

Figure 1. The programming languages used for the projects.

55 https://ecosyste.ms/

56 Trujillo, M. Z., Hébert-Dufresne, L., & Bagrow, J. (2022). The penumbra of open source: projects outside of centralized platforms are longer maintained, more academic and more collaborative. *EPJ Data Science*, *11*(1), 31.

The majority of the projects are written in Python, as shown in Figure 1. This is unsurprising, as Python is the most popular coding language globally.[57] Of those projects, the majority had more than five contributors, as seen in Figure 2. This suggests that projects which are most likely to be gathered and added to indexes like OpenSustain.tech are more likely to have multiple stakeholders. Other popular coding languages include Jupyter Notebooks and R, which are widely used for analysis of data, and C++ and Fortran used for system development and frontend application languages like HTML, TypeScript and JavaScript.

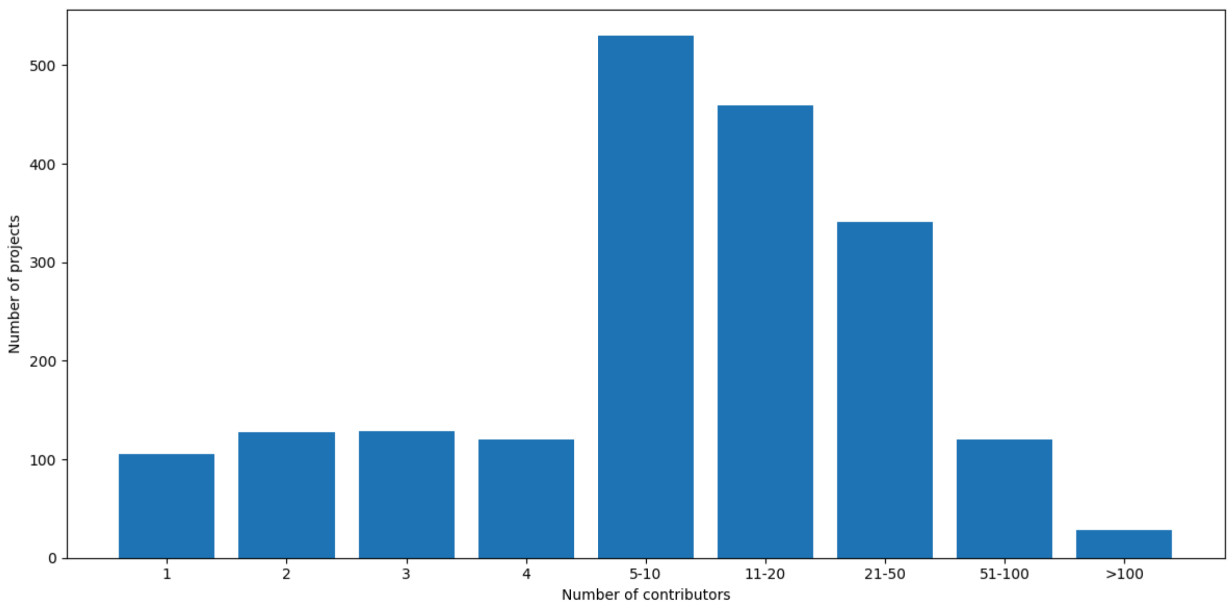

Figure 2. Number of contributors per project. Note that most projects have more than 5 contributors, suggesting that we are collecting projects mostly used by multiple stakeholders.

To analyse the growth of the ecosystem as a whole, the age of all projects are compared with the dataset of 2022 as can be seen in Figure 3:

[57] https://www.tiobe.com

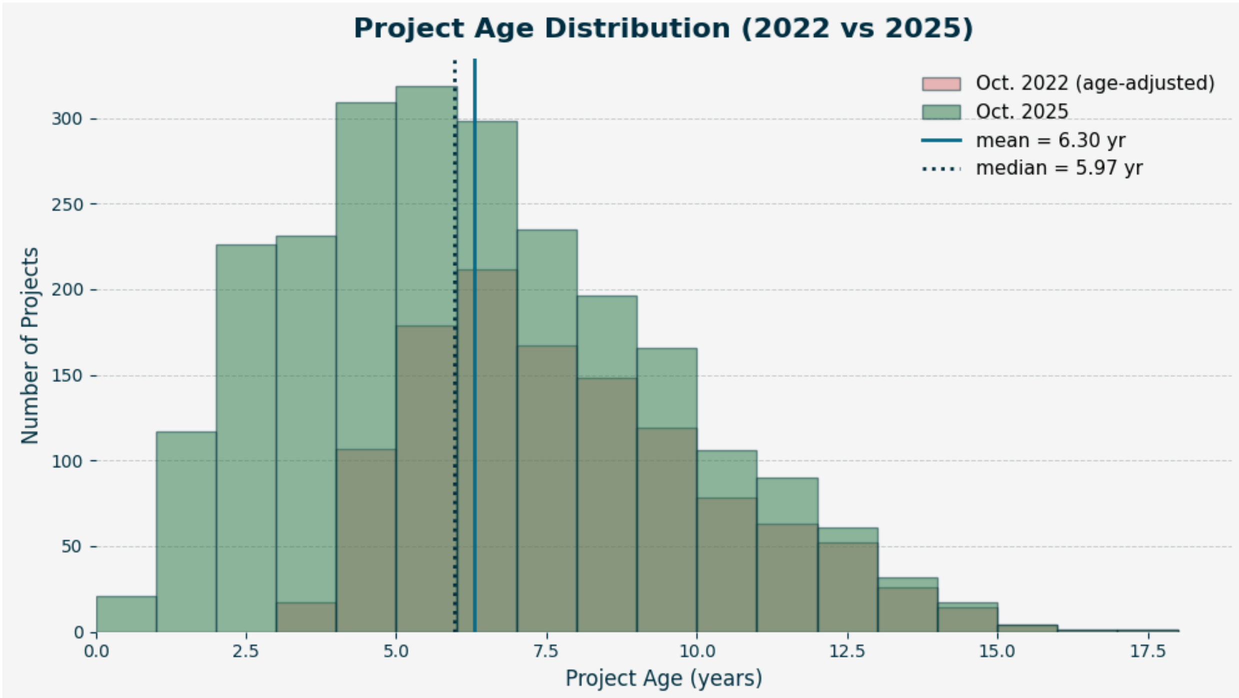

Figure 3. The project age distribution, comparing 2022 to 2025.

The age distribution shows that most of the newly discovered projects that fulfil the requirements of the dataset were started in the last 5 years. The total number of projects also doubled, which may point to a growing ecosystem, although this is not compared here to growth in OSS as a whole. The drop in new projects that are less than three years old may be explained by the time it takes for a project to be ready for large-scale adoption, and by how long it takes us to discover these projects.

The age of projects in different subcategories reveals trends for the last 3 years. Energy system modelling, meteorological observation and forecast and computation and communication (green software) have many new projects which are quickly adopted. Despite their central role in sustainable transformation, topics related to industry, such as emission observation and modelling, or sustainable investment, are still not showing strong new or existing development.

In order to open up a new section of technological domains, a substantial number of open projects with clear, sustainable goals need to be discovered. However, technologies such as nuclear power have not yet been included, as there are incentives for not making these projects public or for obscuring their general use. These developments can often be used for military purposes as well as for the widespread destruction of natural resources, which makes disclosure problematic.

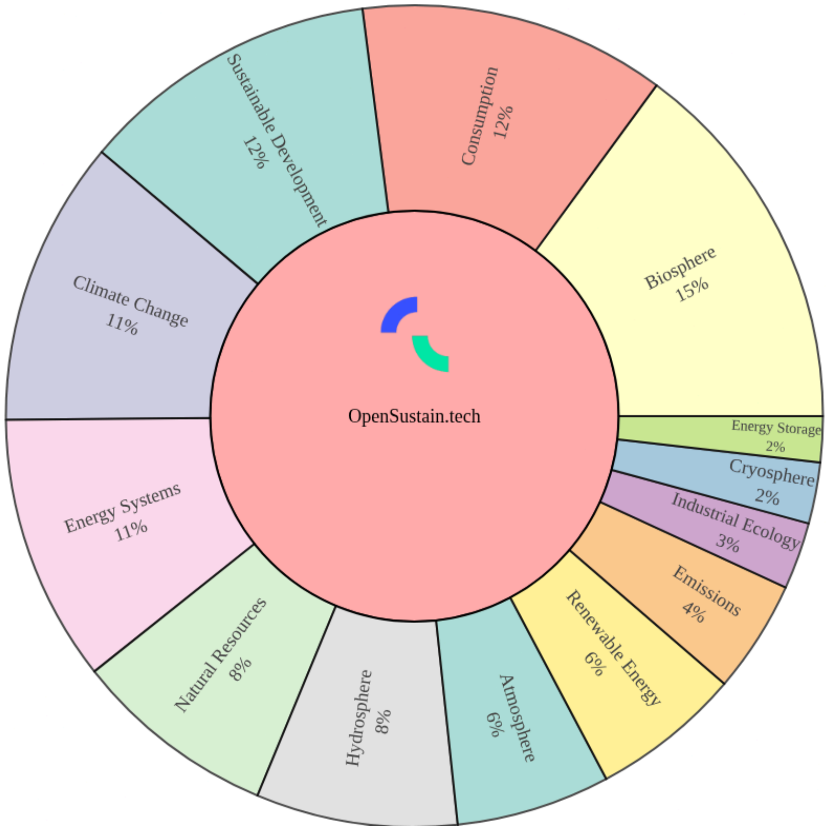

Figure 5. Projects by Category.

Another issue with finding new projects is the limited time and energy for contributors. In total, 106 users contributed PRs to the repository. However, the vast majority of PRs (936 out of 1113, or 84% of total) to the database were made by Tobias Augspurger, the founder. 7 other contributors contributed 6–16 PRs each (amounting to 0.5-1.5% of the PRs, each), while the remaining 123 (11%) PRs were contributed by a larger number of contributors with under 5 contributions each.

Dependencies and Science Score

The projects in the database are part of a wider ecosystem of open source projects. Dependencies in open source are projects which are imported directly into code bases, allowing reuse without redevelopment. This is part of what makes open source such a powerful mechanism, as developers can benefit legally from prior work directly.

There are 2,497 projects within OST. In the global code ecosystem, as measured by ecosyste.ms, 27,482 projects use these projects as dependencies. Many of the projects in

OST will be user-interfacing projects, such as dashboards or webapps, which will not be able to be used as direct dependencies in other software. To be included in other code and measured, dependencies must be registered with a package manager such as npm.[58] The ecosystem projects which rely on OST projects will thus rely on a subset of all OST projects. Different ecosystem projects may rely on the same dependencies. When non-unique projects are excluded, there are 7,776 unique repositories which rely on OST projects,[59] and 2,715 unique packages.[60] Packages which are formally published on package managers will always form a subset of repositories, as not all repositories are published as packages. Finally, 370 OST packages depend in turn upon other OST packages.

OpenSustain.tech projects also depend upon other packages which are not yet included in OST, either because they are too broad in scope or because they have failed other criteria for inclusion. For all of the OST projects, there are 52,559 total dependencies on 14,518 unique packages (again, 370 of which are part of OST).[61]

Ecosyste.ms also performs a Science Score for gathered packages.[62] This score gathers data on linked repository data, documentation, Citation.cff files, academic institutional GitHub organizations, academic email signatures in Git commits, Codemeta files, and scientific mentions to the Journal of Open Source Software (JOSS),[63] Zenodo,[64] academic papers, and the inclusion of DOIs.

For the 2,536 most recent projects in OST,[65] the average score was 43.07, with a median score of 39. 1,829 (72.1%) of the projects had a Zenodo.json file; 1,335 (52.6%) had academic emails in their commit messages; 959 (37.8%) had DOI references and 859 (33.9%) had academic publication links. 343 (13.5%) had Citation.cff files, which is a relatively new file format that standardizes how open source software should be cited.[66] 230 (9.1%) of the projects had an institutional organizational owner. 172 (6.8%) of the projects were linked to JOSS papers.

---

58 http://npmjs.com/
59 https://gist.github.com/andrew/7dac6525f2b93197bc17640133d0af3d
60 https://gist.github.com/andrew/055da5951c643ad3645fa13548eec9c1
61 https://gist.github.com/andrew/b4850f068878b276c15828eac7b7b2d7
62 https://science.ecosyste.ms/
63 Journal of Open Source Software. https://joss.theoj.org/
64 Zenodo. https://zenodo.org/
65 As of October 27 2025, when the Science Score script was run.
66 https://citation-file-format.github.io/

Summary and Outlook

Our analysis shows that the open source software ecosystem in climate and sustainability is still healthy and growing, despite the political and economic challenges of the last three years. The work ahead is difficult. Transparency requires stewardship. Maintaining open standards, datasets, and software demands coordination and expertise. The individuals best equipped for this work already exist: open source maintainers who build, review, and safeguard critical infrastructure across the digital ecosystem, often without compensation and driven by commitment to the public good. By embedding openness at the foundation of climate solutions, we can ensure that transparency is not an afterthought but a structural feature. This alignment of market logic, democratic accountability, and collaborative practice will hopefully create climate solutions that are more trustworthy, more adaptable, and more resilient.

For OpenSustain.tech, after over five years of intense mapping efforts, we have reached a saturation point in the discovery of new projects. This makes it extremely challenging to find new candidates for the dataset, even with a set of new discovery methodologies like Large Language Model (LLM) enabled search.

Due to the advanced complexity of the project, continuing to build the project with purely unpaid community support reveals new challenges. The needs of maintaining the dataset, software tools and downstream services like ClimateTriage makes new development challenging. The newly created governance and partnership with ecosyste.ms helped to split the responsibilities and workloads to multiple stakeholders.

The diverse community behind the projects has also enabled the development of a multitude of innovative ideas to further support the ecosystem in this area. Many discussions and ideas can be found in our issues board. Here are some selections of future ideas:

1. Systematically analyzing all environment-related scientific papers for Git repositories and data repositories to better understand how software is used in published scholarship.
2. Analysing the usage of projects in detail using downloads, issues and citations.
3. Analysing the organization dataset derived from the project's dataset.
4. Creating further recommendations and blueprints how open source and open data can drive systematic transformation for sustainable development.

5. Creation of a stronger network to promote new developments and the Open Sustainability Principles.
6. Organization of a conference bringing together various open source projects in this field.
7. Organization of a summer school / educational platforms / hackathons to combine existing developments, increasing interoperability and collaboration between projects.
8. Extending the Analytics dashboard to show the state of the ecosystem with every new data release.

The future development of OpenSustain.tech will depend on support from volunteers and collaborators, as well as financial backing.

Acknowledgements

We would like to thank and acknowledge the many contributors to OpenSustain.tech. We would also like to acknowledge the help and support of partner organizations, such as the GitHub Sustainability Team, which have offered advice and helped bring OST to a wider audience.