The Open Source Sustainability Ecosystem

August 2023

Tobias Augspurger
Eirini Malliaraki
Josh Hopkins
Dan Brown

Powered by

OPEN CORRIDOR

SUBAK

Contents

Foreword ... 3

Why Open Source? ... 3

How Technology Can Help Sustainability Efforts ... 4

A Brief History of Open Source Culture in Sustainability ... 4

Open Sustainability Principles ... 7

Transparency and Trust ... 7

Traceable Decision-Making ... 7

Collaborative Innovation ... 8

Localization and Decentralization ... 8

Executive Summary ... 9

Insights ... 11

Ecosystem ... 12

Topics ... 13

Popularity & Longevity ... 15

Programming Languages ... 18

Licenses ... 19

Community ... 20

Development Distribution Score ... 20

Contributors ... 20

Organizational Diversity ... 21

Most Listed Projects ... 22

Geography of Organizations ... 23

Users and Usage ... 24

Dependencies ... 24

End Users ... 25

Monetary Value ... 25

Going Forward ... 26

Funding Models ... 27

Recommendations ... 29

Overall ... 29

Technology ... 31

Collaboration ... 33

Funding ... 35

Acknowledgements ... 37

Appendix ... 38

Methodology ... 38

Quantitative Analysis ... 38

Development Distribution Score ... 39

Qualitative Analysis ... 39

Definition of terms ... 41

Foreword

While technological advances are critical to addressing climate change and environmental sustainability, OSS used in the design, operation, and maintenance of human-engineered systems is rarely considered in its own right.

Open source relies on the openness of source code, the ability to modify it, and the freedom to distribute it. Open source software (OSS) is developed through the collaborative efforts of distributed communities, who work together to improve software solutions and share their modifications with others, resulting in an efficient allocation of resources and more robust and reliable products. It is foundational in establishing a digital commons, supporting open innovation communities, managing software development lifecycles, and informing social practices for managing shared resources.

Achieving sustainability goals, such as the decarbonization of energy and agricultural systems, requires a different approach to traditional models, which have relied heavily on proprietary methods and tools. Open source provides a model that can enable all stakeholders to rapidly, efficiently, and effectively develop and deploy software, hardware, standards, and specifications under a collaborative model to create a cleaner, decarbonized future.

This report examines the existing ecosystem of open source sustainability projects across a variety of areas. Many of the most impactful projects are highlighted in this report, however for a fuller view of the projects available, readers should review the complete Open Sustainable Technology landscape online. The report also strives to identify gaps, which filling would benefit sustainability efforts.

Why Open Source?

Without the high modularity made possible by open source, software development as we know it today would not exist. Modern software development is based on the developments of thousands of small projects that have been released as open source over the last decades. If Package managers and indices, such as PyPA and pip for Python, enabled the simple and secure distribution of software components. This process of knowledge generation, adaptation and transformation has been known in mathematics and the natural sciences for centuries. Now it finds its application in open source software development. This approach is often described with the metaphor of "standing on the shoulders of giants", which is explained in more detail in the chapter Open Sustainability Principles.

Besides core information technology infrastructure, open source software is becoming increasingly critical in environmental science,[1] disaster impact assessment,[2] energy efficiency,[3] and sustainability[4] in general. While technological advances are critical to addressing climate change and environmental sustainability, OSS used in the design, operation, and maintenance of human-engineered systems is rarely considered in its own right. This applies to technologies such as energy production and storage, as well as software that simulates, evaluates and predicts complex natural systems, including the atmosphere, biosphere, and hydrosphere.

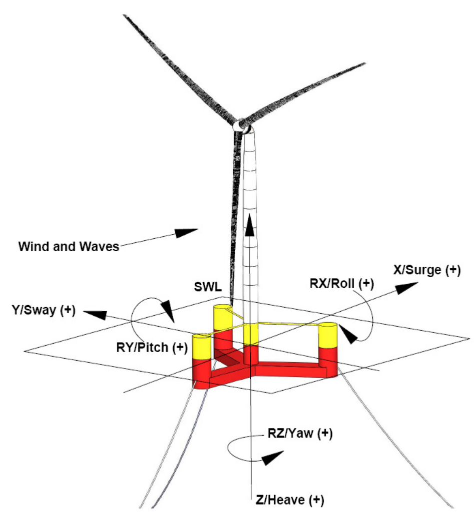


FIG. A - The International Energy Agency is working with the National Renewable Energy Laboratory, Technical University of Denmark and the University of Maine to develop an open source 15-megawatt offshore reference wind turbine design with open simulation and development tools. License: Apache License 2.0

How Technology Can Help Sustainability Efforts

Without observing and simulating vital natural systems, it is impossible to understand their interrelationships, or indicate how humans can preserve this unique, life-protecting world that has emerged in a completely hostile space. Mathematical models, technologies and measurement tools can provide us with independent reporting on the state of the planet, and the extent of human-driven impacts. The knowledge accumulated over decades ensures that people around the world can understand how to preserve vital resources such as fresh water, fertile soil, clean air, and a stable climate. The in-depth understanding of the Earth and its natural, economic and social systems subsequently allows us to ask critical questions and make accurate predictions about human actions and associated impacts such as:

- How does the Earth system respond to global anthropogenic changes?
- What are the risks of exceeding Earth system thresholds? What are the remaining operating budgets? How do we set targets and allocate responsibility?
- What quantity of greenhouse gasses does a product and its supply chain release into the atmosphere?
- How can the value of materials be retained and waste be reduced?
- How does the demand for natural resources impact ecosystems, and how can biodiversity be conserved?
- What are nature's contributions to people, and how do ecosystems services support our economies?
- What are the most effective methods for capturing, storing, and distributing energy?
- What risk does a particular technology pose to humans and the environment in the event of a failure?

Without open data, rigorous science and open software, it is impossible to make evidence-based and verifiable assessments about the feasibility of technologies and practices and their potential impact. The manipulation and withholding of information related to the environmental impact of technologies and unsustainable activities — by state and non-state actors alike — has a long history. From the fossil fuel industry withholding studies about climate change; to manipulation of data and measurements about emissions by the automotive industry; to a lack of communication of important environmental safety information — pollution and environmental disasters go hand in hand with opaque reporting, obfuscated data, and closed-sourced models. To avoid greenwashing and manipulation of environmental information, we need to ensure that claims about the sustainability of technology, the environmental footprints of products, and the impact of human actions are based on quantifiable and verifiable evidence. Likewise, policy development and its real-world implications for people and the planet must be guided through open, science-based approaches.

A Brief History of Open Source Culture in Sustainability

Without the exchange of measurement data between meteorologists, it would have been impossible to achieve accurate weather forecasting in the early days of this scientific field. The origin of the World Meteorological Organization (WMO) stretches back almost 150 years, and international collaboration and data sharing have been central to its activities from the very beginning.

With the advent of the space age, the small number of weather satellites forced nations around the world to share data between different observatories in order to understand the state of the atmosphere and other Earth systems as a whole. The weather forecasts that became possible as a result of this innovation added significant value to all sectors of society. This open, planetary-scale computation project increased our awareness of anthropogenic climate change and our ability to act in this environment.

Despite increasingly strong commercial interests in civil satellite data, a strong culture of open data and open science has been established. 41% of the unclassified Earth observation satellite programs of the 10 largest nations today provide open data.[5] Open access to this data has led to a rapid increase in the number of downloads shown in Fig. B using the example of the Landsat Archive. The creation of the Global Climate Observing System (GCOS), launched at the Second World Climate Conference, also committed parties to support international and intergovernmental sharing of data and analysis.[6] The urgent need for a transition to a more sustainable society, and clear signs of anthropogenic climate change, have created multiple new movements and organizations — all with similar open ethos and mandates across domains:

- In the energy sector, the OpenMod Initiative formed in 2014 with the goal of opening up energy models, so the sector as a whole can improve the quality, transparency, and credibility of its products and create better research and policy.
- In 2017, academic Stefan Pfenninger wrote an article published in Nature urging scientists to "Free The Models" for the energy sector — a call to action supported by influential scientists like Auke Hoekstra and entrepreneurs like Michael Liebreich. This led to the sector pushing toward open data and open source code — a move that will be critical for the transition to a fossil fuel-free economy.
- In 2018, Shuli Goodman founded the LF Energy, an open source organization within the Linux Foundation that enables companies worldwide to develop energy systems related OSS tools collaboratively. Shuli sadly passed in January 2023, but her work continues at LF Energy.
- ClimateChange.ai convenes an open community and provides multiple educational resources around tackling climate change through artificial intelligence — another technological revolution that has been catalyzed by open source culture.
- The Open Climate Community Calls, launched by the well known Appropedia community, created worldwide connections between multiple individuals from different domains — to understand the relationship between openness and climate change.
- OS-Climate created another large-scale Linux Foundation collaboration between major companies to increase transparency and traceability in Environmental, Social, and Governance (ESG) ratings and support sustainable investment through an open source toolchain.
- The Digital Public Goods Alliance is an initiative endorsed by the UN Secretary-General that facilitates the discovery and deployment of digital public goods. Similarly to the Open Sustainable Technology project, it provides an index of valuable open source projects related to the UN Sustainable Development Goals, accompanied by a billion-dollar investment commitment.
- The Coalition for Digital Environment Sustainability (CODES), is an international multi-stakeholder alliance bringing together the scientific community, governmental institutions, NGOs, tech companies, and civil society to champion digital sustainability.

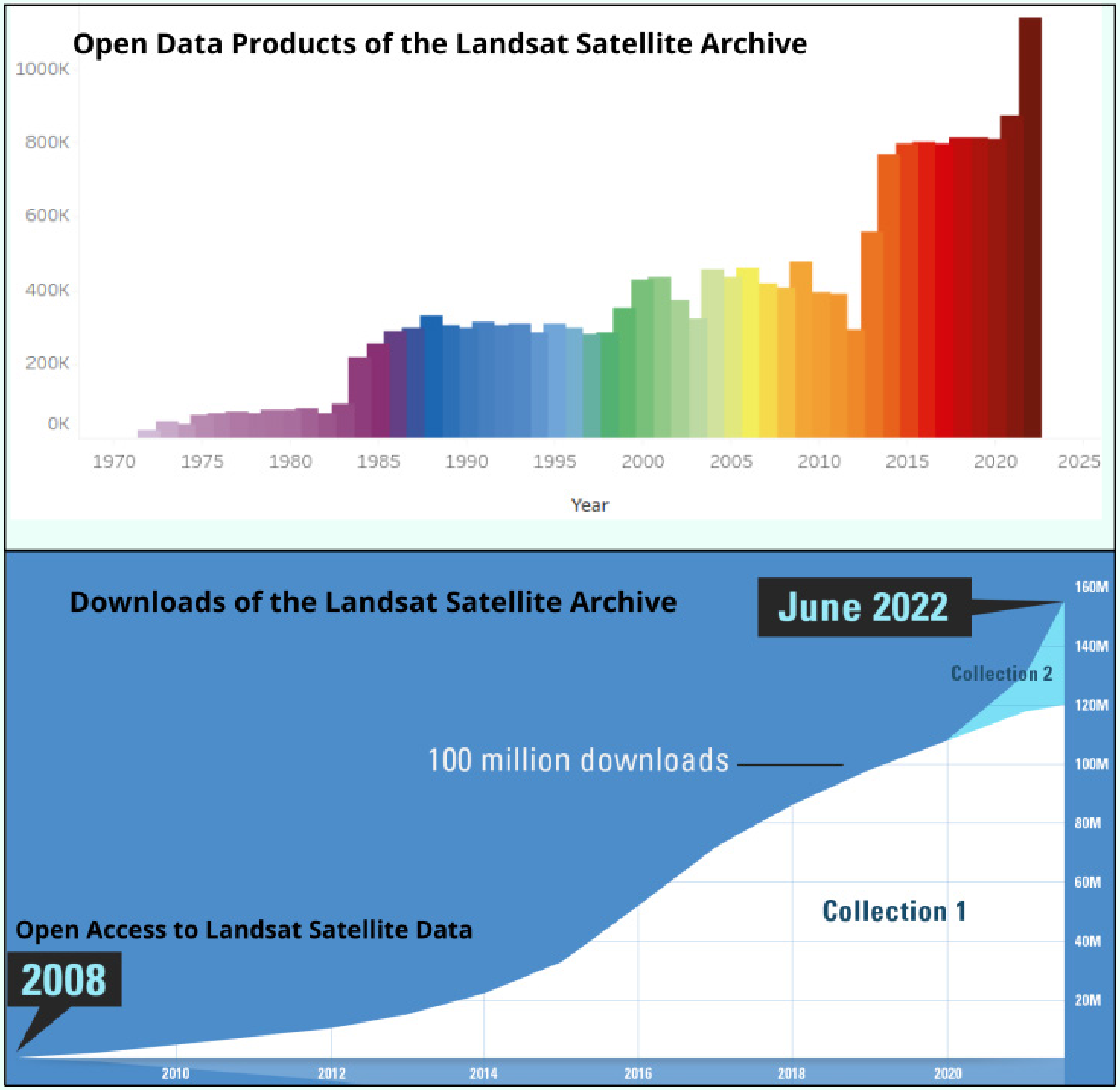


FIG. B - With the disclosure of the Landsat archives, the use of these satellite data increased noticeably Source: Landsat Archive Dashboard

Open Sustainability Principles

Openness provides the basis for collaborative sense-making, enables meaningful consensus — based on an accurate, shared understanding of the state of our planet — provides direction on how to best coordinate our choice-making, and builds capacity for effective action.

Openness is a key indicator of sustainability. It allows different disciplines, organizations and societies to refine their understanding of sustainability systematically and respond to new information effectively. This basic idea of iterative innovation is often associated with the metaphor “standing on the shoulders of giants”, coined by Issac Newton,[1] which means to build on previous discoveries. By making one’s intentions and conclusions transparent and accessible, knowledge can continue to accumulate and be refined over time. It is the integrity and accessibility of information and knowledge that makes “openness” integral to long-term sustainability. Openness provides the basis for collaborative sense-making, enables meaningful consensus — based on an accurate, shared understanding of the state of our planet — provides direction on how to best coordinate our choice-making, and builds capacity for effective action.

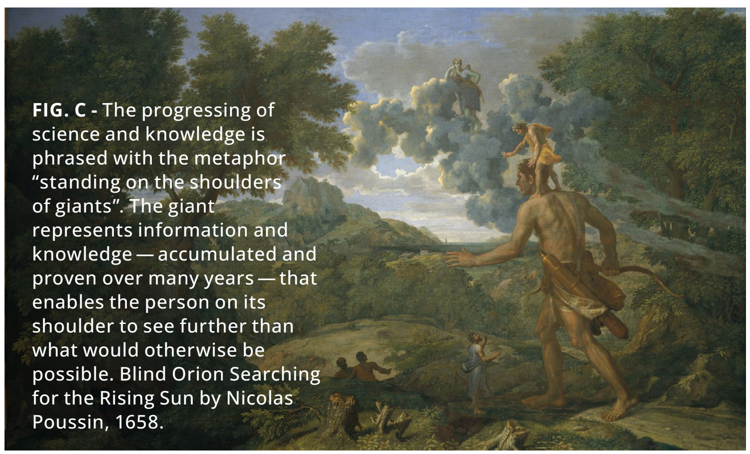
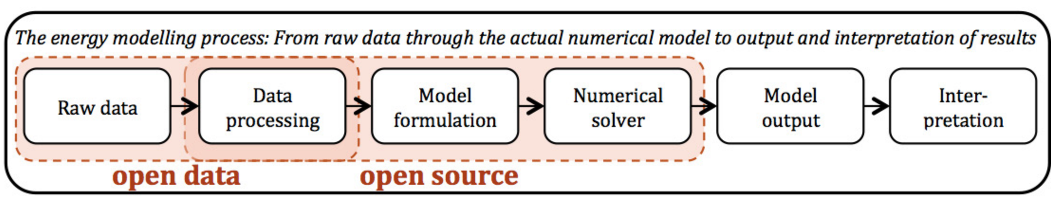


FIG. C - The progressing of science and knowledge is phrased with the metaphor “standing on the shoulders of giants”. The giant represents information and knowledge — accumulated and proven over many years — that enables the person on its shoulder to see further than what would otherwise be possible. Blind Orion Searching for the Rising Sun by Nicolas Poussin, 1658.



The cross-domain quantitative analyses, discussions, and research in this study have made it possible to identify the following guiding principles, referred to as **Open Sustainability Principles:**

Transparency and Trust

Understanding how human actions affect the environment, and which practices and technologies protect natural resources in the long term, is essential for a sustainable economy. Many solutions are complex, and it is not always obvious whether they will be economically viable, socially inclusive, and environmentally sustainable. Open source approaches allow for public scrutiny and independent auditing. Open data and models create transparent metrics and more comprehensive sustainability impact assessments.

Traceable Decision-Making

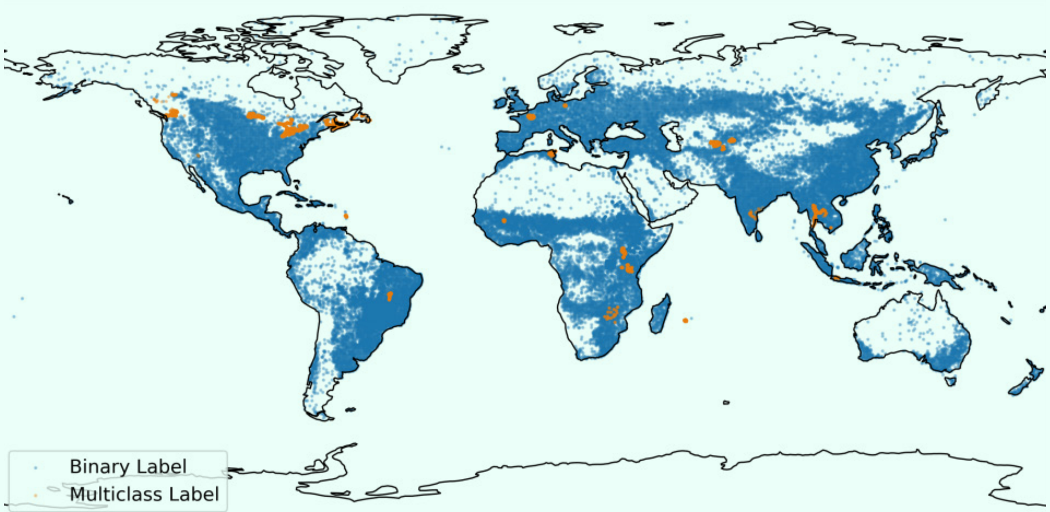
Open source approaches create a common understanding of what is required to deploy various solutions and scientifically assess their environmental impact. With more accurate information, efforts can be focused on the highest-leverage solutions. Open science discloses the assumptions behind these models and measurements and reduces uncertainties in the long term through continuous improvement and observation. Demanding transparency at each stage of the decision-making process (e.g., modeling, forecasting, impact assessment) will help reduce greenwashing and ensure that the environmental impacts of both state and non-state actors are verifiable and traceable.

FIG. D - The openmod manifesto illustrates how traceable decision making is possible in the decarbonization of the energy sector. The openness of data, models and discussions enables traceable knowledge creation for science and society. License: CC 4.0

Collaborative Innovation

Open innovation strategies accelerate collaborative cultures. This translates into significant economic and commercial advantages: the identification of new business opportunities, increased flexibility in developed solutions, enhanced economic diversity and complexity, and the motivation of a broader community to develop an ongoing stream of innovations.[2] Open source approaches also mean each distinct problem only has to be solved once. Frequently, multiple actors invest in open data products, algorithms, and software — leading to duplication, incoherent outcomes, and inefficient use of resources. Through open collaboration, OSS can mitigate inefficiencies in resource allocation and allow for the emergence of order through a decentralized, loosely coordinated approach. OSS also standardized the practice of collaboration between software contributors. This distributes the benefits of cooperative software development towards common standards and higher code quality.

Over the past decade, multiple companies have demonstrated that open businesses can also grow and be financially sustainable. A 2021 study on the economic impact of open source software and hardware concluded that open source technologies injected €65-95 billion into the European economy. Open source significantly boosted small and mid-size enterprises, Europe's most important horizontal economic stakeholders. Growth was reflected in the increased creation of technology startups at more than 650 per year. Commercial open source has seen a massive upswing in the last year with its own conferences, business models, strategies and venture capital funds like OSSC or Open Core Ventures. These developments have already shown a new, open mindset that is driving a systematic shift away from classic proprietary business models — the same mindset can be applied to open sustainable technologies.



Localization and Decentralization

Wealthier economies with greater technological capacity can enable emerging economies to adapt more quickly to climate change by openly licensing technologies. Open source-based participation and collaboration contribute significantly to the global circulation of knowledge about sustainable development. In addition, by sharing information and enabling technologies, under-resourced communities can rapidly build local interdisciplinary capacity, accelerating both digital and sustainability transformations. This is especially important for place-based innovations in mobility, food, and housing. Countries must also be supported to rapidly deploy and adapt open source technologies to meet their infrastructure needs. This has the potential to enhance economic complexity, provide job growth, and drive resource and efficiency gains.

FIG. E - CropHarvest is an open source remote sensing dataset for agriculture with benchmarks. It collects data from a variety of agricultural land use datasets and remote sensing products. License: CC-BY-4.0

Executive Summary

Open source is everywhere. Its culture of transparent and collaborative innovation has transformed modern society, with over 97% of critical digital infrastructure and services depending on it. The role of open source has become increasingly important in addressing environmental challenges. Mathematical models, data and measurement tools, accumulated and shared over decades, have empowered communities worldwide with the understanding needed to preserve Earth's vital resources — fresh water, fertile soil, clean air, and a stable climate. Open cultural and technical approaches are essential for supporting traceable decision-making, building capacity for localisation and customisation, providing new opportunities for participation, and preventing greenwashing by ensuring transparency and trust.

Yet, despite the transformative impact of open source, its potential within environmental sustainability is not well understood. This has resulted in a systemic lack of investment, ultimately limiting our collective capacity in addressing society's most pressing challenges. There is a clear need to accelerate open source efforts to achieve innovation and sustainability at scale. However, a systematic assessment of which projects can be considered critical digital infrastructure or where significant funding and resourcing gaps exist has been lacking.

This report is a meta-analysis, designed to provide the first comprehensive analysis of the open source software ecosystem in the field of sustainability and climate technology. More than one thousand actively-developed open source projects and organizations were compiled and systematically analyzed using qualitative and quantitative methods as part of the Open Sustainable Technology project and its associated database. The analysis covers multiple dimensions — including the technical, social, and organizational — providing an empirical basis for guiding community building, policy development and future investment. We examine the health and vibrancy of this emerging ecosystem, highlighting key risks and opportunities for users, developers, and decision-makers. Finally, we present a shared vision and strategies to accelerate the widespread use and adoption of open source within this increasingly important field. The key objective is to create a vision and develop strategies to accelerate the use of open source in environmental sustainability.

Our findings reveal that while there is wide recognition of the importance of technology in the transition to a more sustainable future, open cultural and technical approaches used in the research, design, operation, and maintenance of human-engineered systems are rarely considered in their own right. Despite the proven transformational potential of open source in other domains, open source still plays a minor role within sustainability as a transformation strategy.

Half of all identified projects are in high-impact, data-rich fields such as climate science, biosphere, energy system modeling, transportation and buildings. Other topics, such as carbon offsets, battery technology, sustainable investment, emission observation and integrated assessment modeling, show few notable developments. Most identified projects are relatively young, with a median age of 4.45 years. Despite this, their open source contributors demonstrate a high level of knowledge and ability to sustain innovative capacity and project longevity, given sufficient resources and support.

Analysis of the distribution of knowledge, work, and project governance reveals that small, open source communities lead most of the development in this ecosystem. On average, open source software projects rely heavily on a single programmer responsible for ~70% of the contributions to a project. This indicates potential contributor risk, which could limit the potential of many of these projects. A sectoral imbalance in open information and knowledge exchange can be seen, with academia and several government agencies contributing significantly to this ecosystem. Meanwhile, the lack of for-profit organizations and startups with open source business models is remarkable, particularly given the rise of ventures within the domains.

Many of the identified projects both enable and accelerate sustainable outcomes. These range from the consortium efforts behind models underpinning the well-publicized assessments of the Intergovernmental Panel on Climate Change (IPCC), to community-driven projects that provide real-time electricity insights to decision-makers across the globe.

Finally, the report presents four guiding principles embodied within this emerging ecosystem, which we define as Open Sustainability Principles: Transparency and Trust, Traceable Decision-Making, Collaborative Innovation, and Localisation and Decentralisation. These principles provide the basis for collaborative sense-making, enable meaningful consensus — based on an accurate, shared understanding of the state of our planet — provide direction on how to best coordinate our choice-making, and build capacity for effective action.

We conclude that digital and sustainable transformation must converge as a digital public good if we are to achieve a safe and equitable corridor for people and the planet. Open sustainability principles can help governments, research institutes, non-governmental organizations, and businesses move rapidly toward decarbonisation and better conserve natural resources and ecosystems. Therefore, we propose several recommendations for building greater capacity and support for open source open source sustainability ecosystem:

- Strengthen the interconnectivity and knowledge exchange of the identified open source communities.
- Build capacity and increase the potential for real-world impact by connecting projects to local use cases.
- Adapt and extend existing projects to underrepresented countries in the Global South.
- Create incubators and other support programmes for open source in environmental sustainability, including dedicated funds that provide core funding for development and maintenance.
- Develop better technical interfaces between platforms, data, models and open source tools across and within sectors to avoid “reinventing the wheel”.
- Standardize environmental data exchange across different levels of government.
- Close the knowledge gap on the environmental impact of state and industries.
- Transform financial institutions through transparent and scientific decision-making for sustainable investments.
- Apply an “Open Source First” criterion when providing funding for sustainable technologies.
- Recognise the contributions of open source in advancing sustainable development on a global scale.

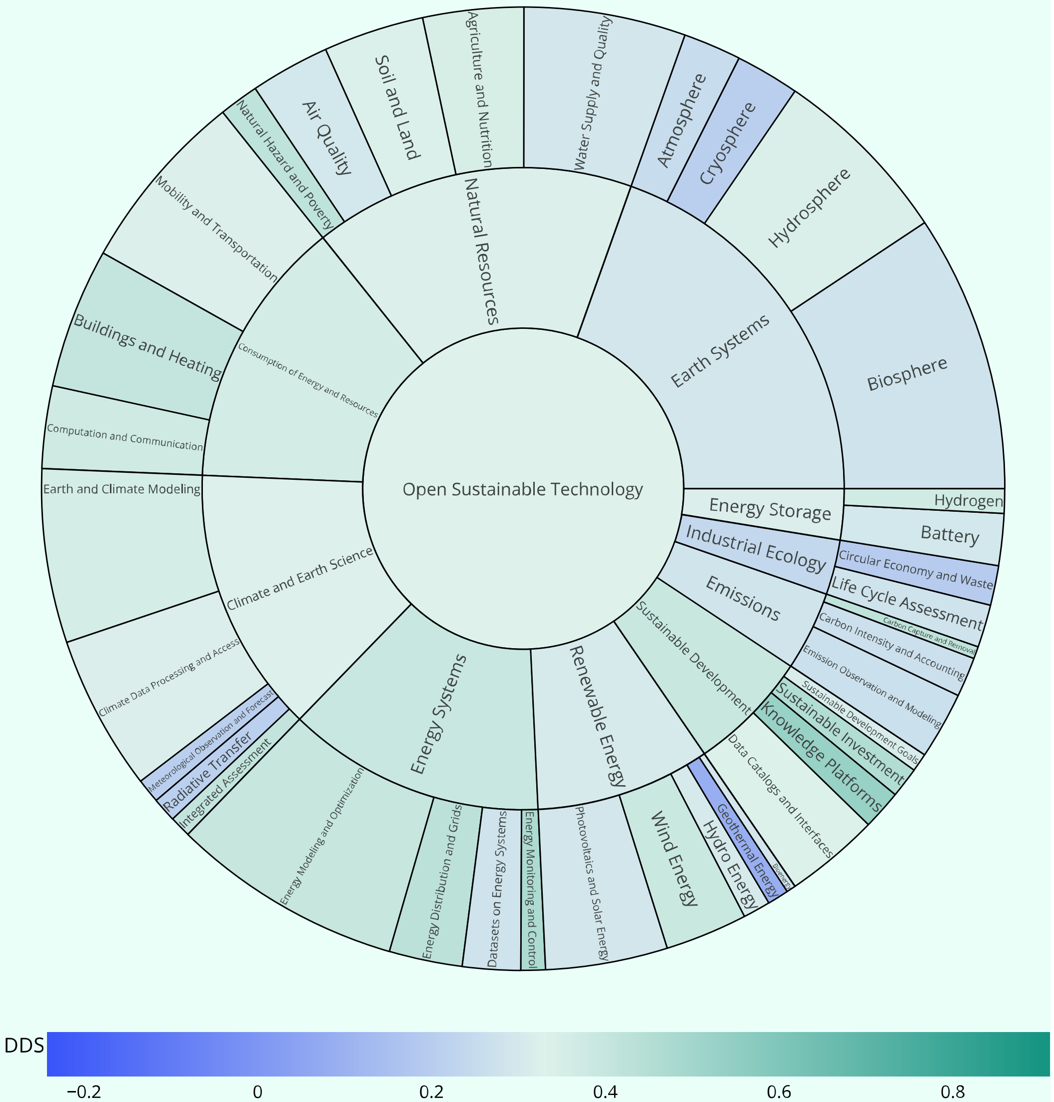
Insights

At the time of writing, 1339 projects have been identified worldwide. Of these, 1188 projects are hosted on GitHub, 27 on GitLab and 125 on other websites or self-hosted Git platforms.

A total of 996 active project repositories have been found. A project is considered active if the public repository has at least one commit or closed issue within the last year. Inactive projects, or those that have become inactive since data collection began two years ago (192), have been excluded from our analysis to prevent distortion of current trends. Unless otherwise noted, all plots in this study refer to active projects.

DIMENSION	VALUE
Total number of projects	1,339
GitHub projects	1,187
GitLab projects	27
Other platforms	125
Number of projects in personal namespace	547
Number of projects in community namespace	792
Total stars of all projects	126,568
Total contributors in all projects	14,628
Active GitHub projects	995
Inactive GitHub projects	192
Projects with contribution guide in %	30.58
Projects with code of conduct in %	13.98
Projects accepting donations in %	3.71
Median number of commits	481
Median stargazers	42
Median stars last year	11
Median Development Distribution Score	0.3043
Median number of contributors	6
Median closed issues last year	8
Median commits last year	23
Median age in years	4.45

FIG. 1 - Overview analysis of total projects analyzed



Ecosystem

Projects are grouped into fields based on their primary topic of focus. While the boundaries often overlap, these fields help to paint a broad landscape and can provide insight into the ecosystem health and complexity of fields relative to each other. The following sunburst diagram shows the relationship between fields, topics, and projects. The color represents the divergence from the mean Development Distribution Score (DDS). DDS measures the distribution of work among developers, with a high value indicating a high distribution of work and, thus, a stronger than average developer community. Discover ecosystem with the interactive sunburst diagram.

FIG. 2 - All studied projects grouped into the corresponding fields and topics

Projects over time by topic

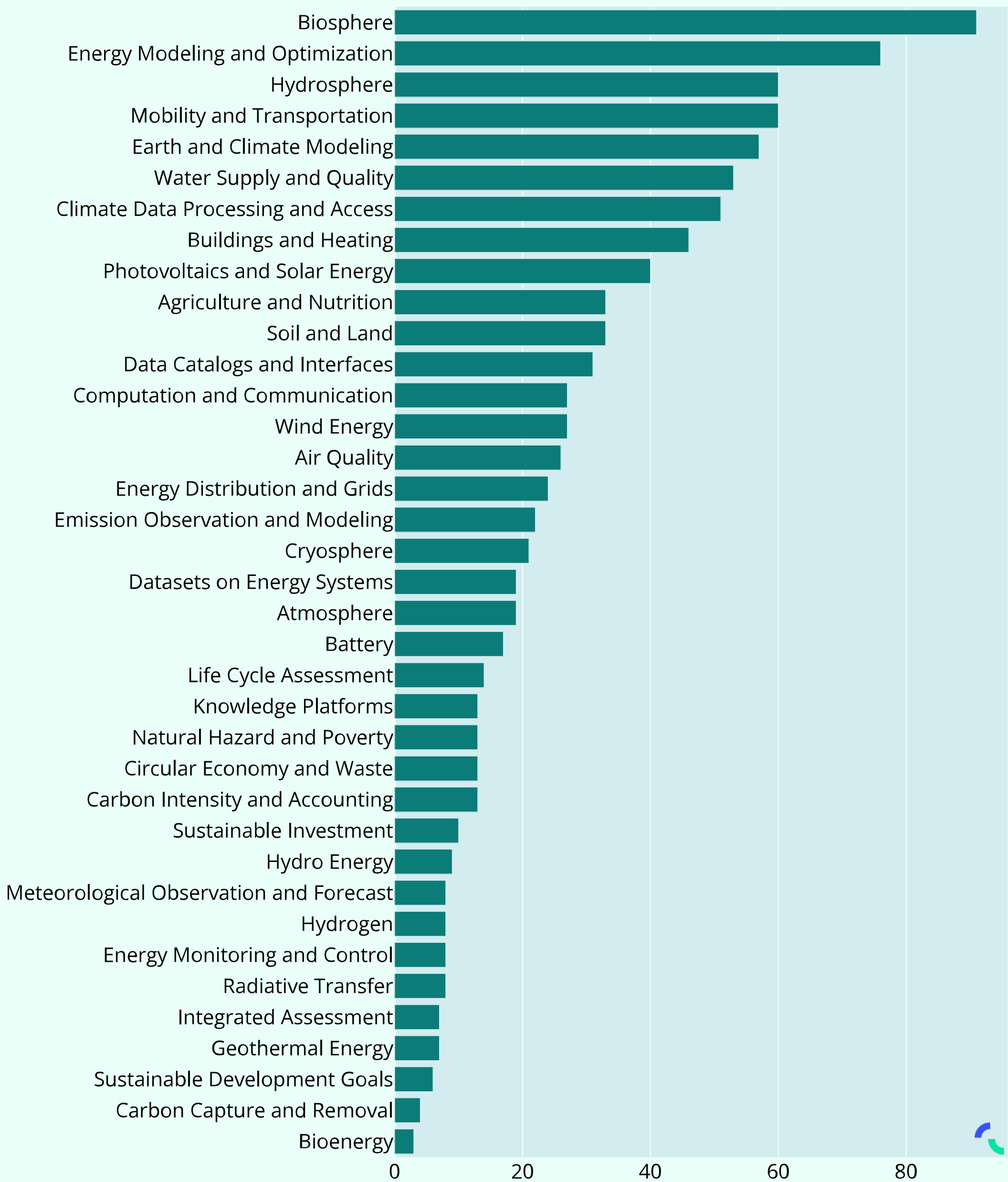
Topics

Project topics have been identified across sectors, technologies and research fields. This mapping process was iterated multiple times as part of the analysis and will continue to evolve as niches develop and emerge. While it may be difficult to compare the scope of the topics directly, the relative size and complexity allows us to identify neglected, vibrant, and emerging areas.

The following scatter plot provides an overview of all projects studied within their respective topics. The size of the circles is proportional to the relative scale of the projects, based on total commits and contributions.

FIG. 3 - Overview of all projects of the last 14 years since the launch of GitHub

Projects within topics



45% of all identified projects can be found within biosphere, hydrosphere, water supply and quality, energy system modeling, mobility and transportation, and buildings and heating. This is likely due to the research maturity of these fields, the multitude of scientific organizations behind them, and the relatively good availability of open data within these fields. We can see strong open source ecosystems, particularly in the field of energy modeling and optimization across various renewable energy technologies, such as photovoltaics or wind energy. However, despite the central importance of batteries for energy storage, only a few OSS projects are under development.

Other areas see a much more limited number of open source projects, such as sustainable investment, representing only 1.15% (a total of 11 projects), and emission observation and modeling, representing only 2.1% (22 projects).

FIG. 4 - Number of individual projects within topics

Topics with low OSS representation include bioenergy, hydrogen, and carbon capture. Despite continuing to attract large amounts of funding from government and investors, the small number of open source projects associated with these technologies makes it difficult to quantify the state of, and potential contributions towards sustainable development. Topics like carbon offsets, carbon credits or climate neutrality disclosures could not be investigated in depth due to a general lack of OSS projects.

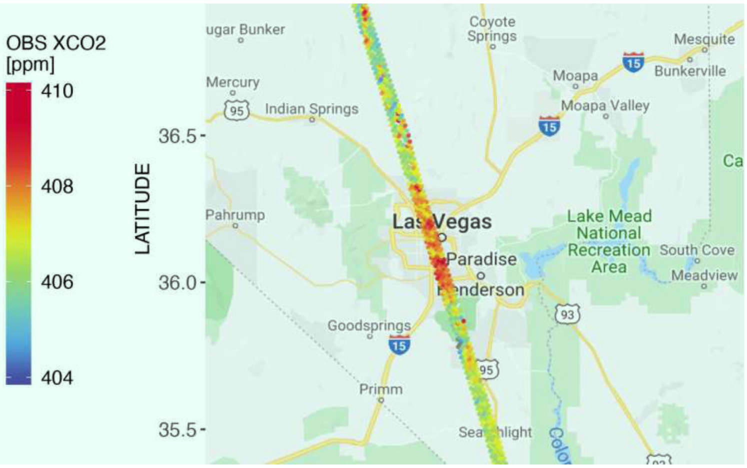


FIG. 5 - The goal oco2peak is to localize CO2 emissions on Earth based on the carbon concentration data measured by the OCO-2 Satellite from NASA. It is one of the few open software tools that have been released in the field of emission observation and modeling.

Popularity & Longevity

The popularity of a particular topic or field can be determined by summing the number of stars across all related projects on GitHub. There are a total of 127,038 stars across all of the identified projects; but a search on GitHub revealed there are 27 projects that have more stars than the entire open source sustainability ecosystem combined. With a median of 42 stars and only 3 projects with more than 1,000 stars, the sustainability ecosystem can be considered relatively unpopular on GitHub.

High growth in popularity can be seen in topics such as energy, transportation, earth observation, and meteorology. These are just a few examples of areas where software innovations are of particular importance, which can explain why open source software is more widely recognized within these communities.

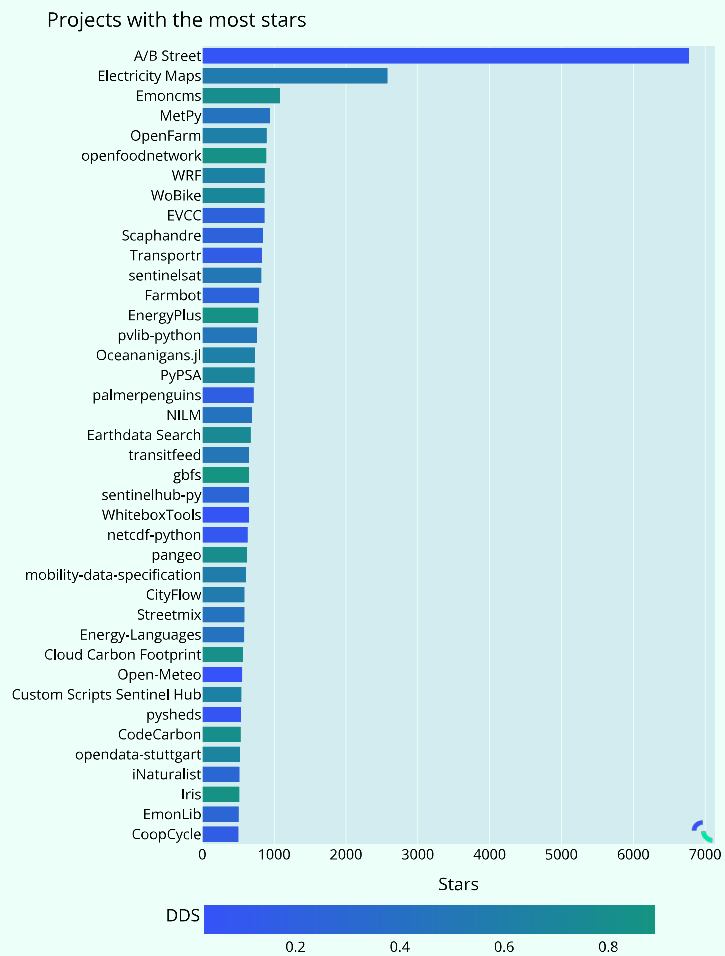


FIG. 6 - Projects with most stars

Here, the accessibility of different domains and applications must also be taken into account. Projects such as A/B Street, Electricity Map, OpenFarm, Open Food Network, Emoncms, StreetMix, or the Farmbot, have relatively lower barriers to entry for software developers and end users. These projects can speak to a wide audience and require less specialist or technical knowledge for contribution and usage. Hence, there is greater awareness in general.

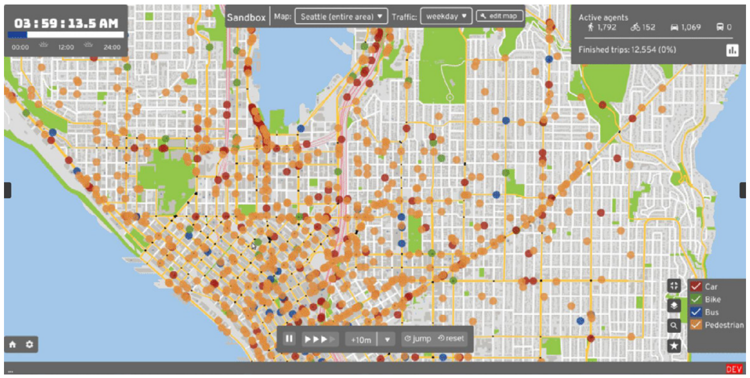
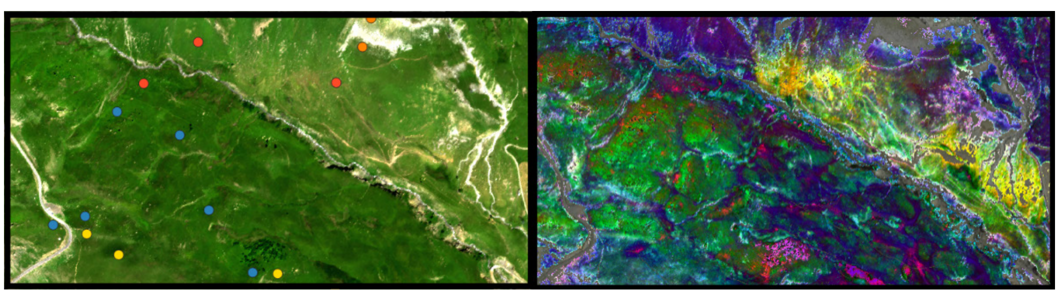


FIG. 7 - A/B Street - Transportation planning and traffic simulation software for creating cities friendlier to walking, biking, and public transit. License: Apache-2.0

On the other hand, popularity should not be conflated with impact. BiodivMapR, for example, has the potential to create a global map of biodiversity based on multispectral satellite images. Projects on sanitary problems or biogas, such as Santiago.jl, are also apparently less popular based on the star metric but can have a significant impact on energy supply and water quality respectively.

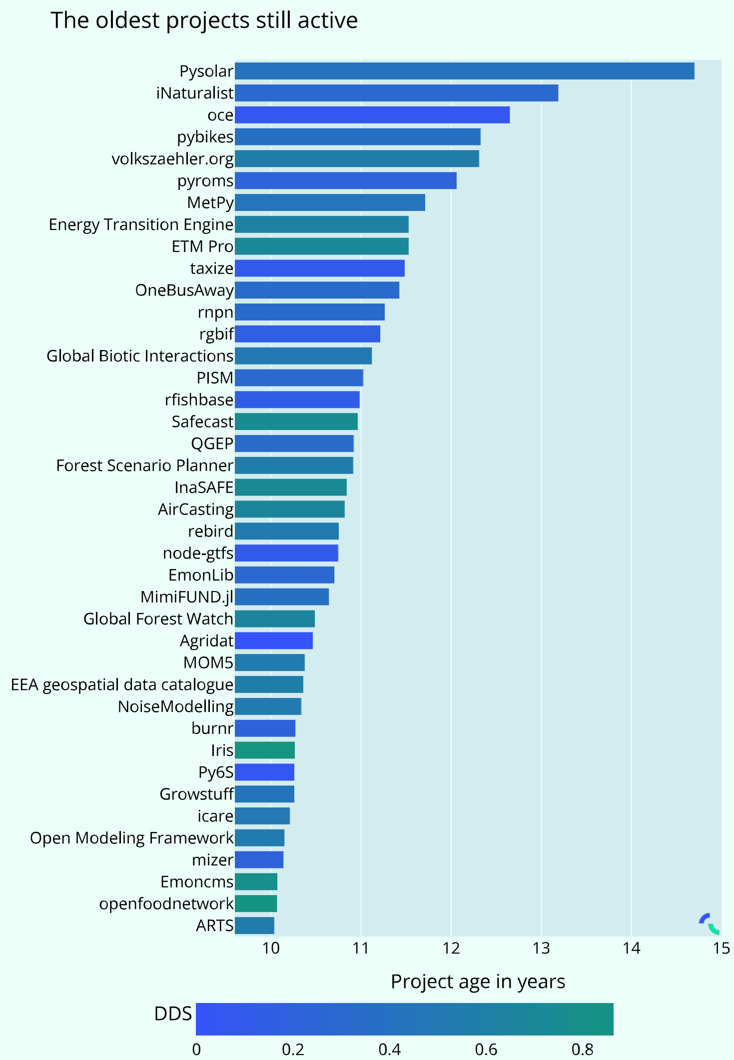


A repository's longevity is also an important health indicator. If a project has been actively used for a long time we can infer that developers and users are interested in its continued development. The age of all projects follows an approximate normal distribution, with a sharp drop for younger projects and a peak at 3.2 years.

A closer look at the age distribution of the repositories indicates a median age of 4.45 years, suggesting that most project development began recently. According to the project age distribution, the number of new projects has decreased in recent years. Further analysis will be required to determine whether this represents a real trend.

FIG. 8 - biodivMapR - An R package for α- and β-diversity mapping using remotely-sensed images. The left image shows the visible spectra with multiple validation sites. The right image shows the derived biodiversity. License: GPL-3.0

FIG. 9 - Distribution of project age



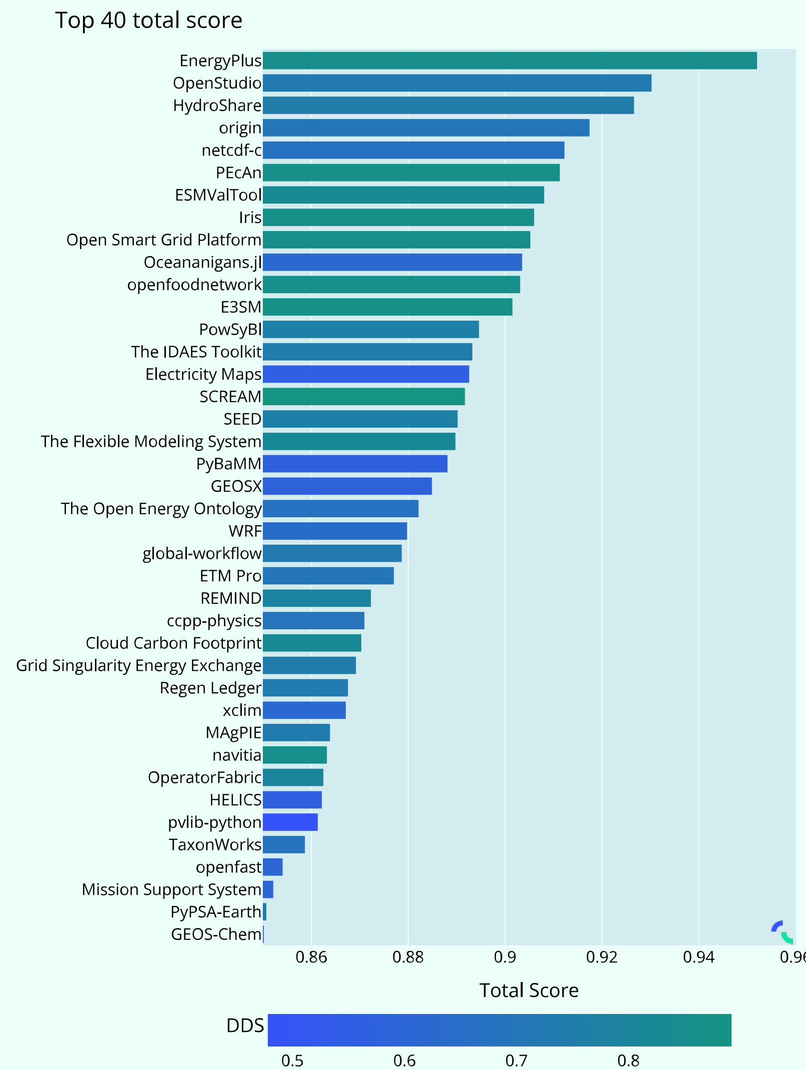


FIG. 10 (LEFT) -
The oldest projects still active

FIG. 11 (RIGHT) -
The 40 Projects with the highest total score

Many of the oldest active projects, such as Pysolar, iNaturalist or OCE, originated in academic communities at a time when the share of other organizations in OSS was still low. This illustrates the grassroots origins of the open source community.

Some of these projects, such as Py6S, agridat or oce have been predominantly developed by a single contributor for over 10 years.

Programming Languages

The types of programming languages utilized provide insight into the skills required of code contributors and the nature of the projects themselves. This metric can help newcomers navigate open source projects, as well as enable project and product managers to gain insight into a project's profile within the context of their own experience and organizations. It can also help inform students as to which programming languages they might focus their efforts on learning, depending on their topic of interest.

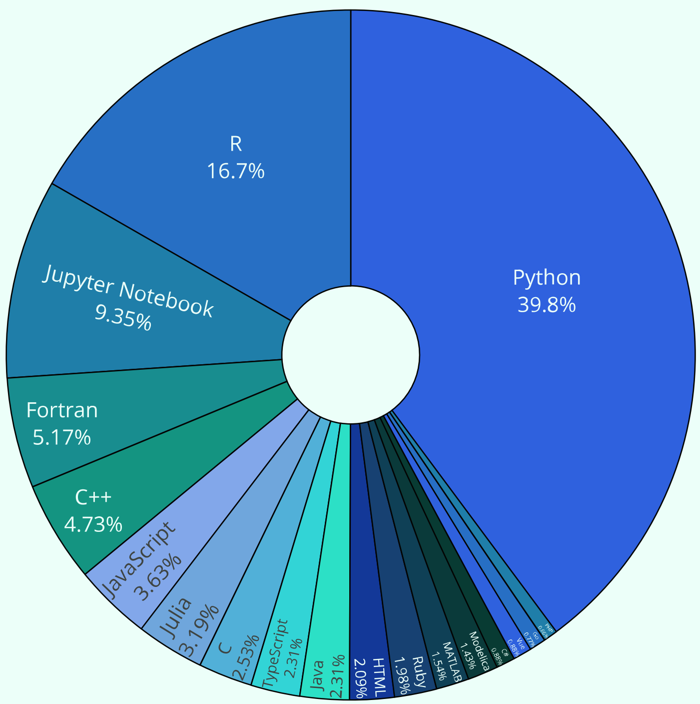


FIG. 12 - Distribution of programming languages

Python dominates the OSS movement for sustainability and is used in 39.8% of all projects, followed by R (16.7%), Jupyter notebooks (9.34%) and other languages like Fortran, C++ and Java. Statistics from GitHut 2.0 and the official numbers of GitHub can provide insights into the programming language usage of open source projects more broadly. Compared to the broader open source ecosystem, it is clear that Python has a significantly higher use within the repositories analyzed, compared to widely popular languages, such as JavaScript. This indicates a strong focus on analyzing large datasets, where Python and Jupyter Notebooks are increasingly dominant, with less focus on the web application side.

The use of R deviates significantly from the wider open source ecosystem. A concentration of software written in R can be found, particularly within the topics of biosphere, hydrosphere, water supply, soil and land use, climate, and food and agriculture. This can be attributed to the high number of data-driven and statistical-related projects within these topics. Despite its advanced age of over 65 years, Fortran is still widely used in the Earth system models applied across hydrosphere, climate and atmosphere fields. This can be explained by the long development time of these projects, and the necessary numerical efficiency of such models for high-performance computing. Julia, a relatively new language, also has a wide range of applications especially in the climate field. For some special use cases, such as building simulation, programming languages like Modelica are frequently used.

Licenses

Licenses ensure that communities behind many of today's most innovative technologies can protect their creations in whatever way they see fit. The open source license determines how openness is defined from a legal perspective.

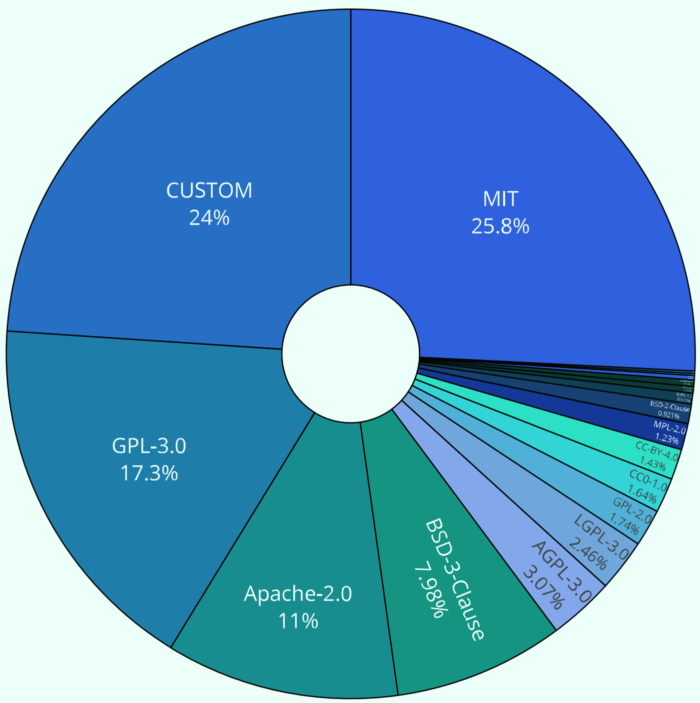


FIG. 13 - Distribution of licenses

Permissive licenses like BSD, Apache and MIT are the most popular in the field of sustainability. The MIT license was the top choice, used in ~26% of the projects. MIT is a short and simple, permissive license. Permissive licenses are easier to use in commercial products, and unlike the copyleft license, they do not enforce the release of all code under a similar license. Permissive licenses, like MIT, create many opportunities for the reuse and commercialization of OSS projects within proprietary software. Projects under this license, on the other hand, may be jeopardized by the use of proprietary products. For example, companies are not required to release bug fixes or enhancements, so improvements cannot be contributed back to the open source project. In some cases, this can lead to developers losing motivation to contribute over time as it becomes apparent that the capacity for the overall development effort is diminishing.

The second most used license was a custom license, used in 24% of the projects. All projects that could not be mapped to standard open source licenses based on the SPDX License List, and those with modifications from the original open source licenses, are considered to be custom licenses.

GPL 3.0 is the third most popular license used by 17.3% of all projects. Permissions under this strong copyleft license require licensing the entire source code of the licensed works and modifications. GPL was created to protect software from becoming proprietary, or private. Copyright and license notices must be preserved and contributors provide an express grant of patent rights. Copyleft licenses are more prevalent in business models that rely on direct monetization of the core project.

Community

Development Distribution Score

The Development Distribution Score (DDS) weighs how development is distributed between project contributors by benchmarking the contributor with the most commits in relation to the other contributors. The distribution of knowledge, work, and governance is critical to a project's long-term viability. When a project or organization undergoes significant social or technological changes (for example, personnel leave a project or can no longer contribute), others require the knowledge and capacity to continue with the initiative. The metric compares a project's reliance on a small number of contributors and, as a result, its resilience to change. Projects with a low DDS appear to be more vulnerable to decisions made by a single organization or developer, which affect not only other developers or users, but also the dependencies to other projects.

Contributors

Open source projects have different types of contributors who undertake various activities ranging from software development, product management, authoring documentation, event planning, community management, and marketing. Analyzing how many people are contributing to a project and who they are is critical to understanding organizational and individual engagement with OSS projects. Best practices from projects with a large contribution base can subsequently be shared and adopted by the broader open source sustainability ecosystem. For the purposes of the analysis, someone is considered a contributor if they have made at least one commit to the Git repository. Figure 14 shows a list of the top 40 projects with the most contributors.

The high DDS among the projects with the most contributors shows the importance of sharing work among a variety of people

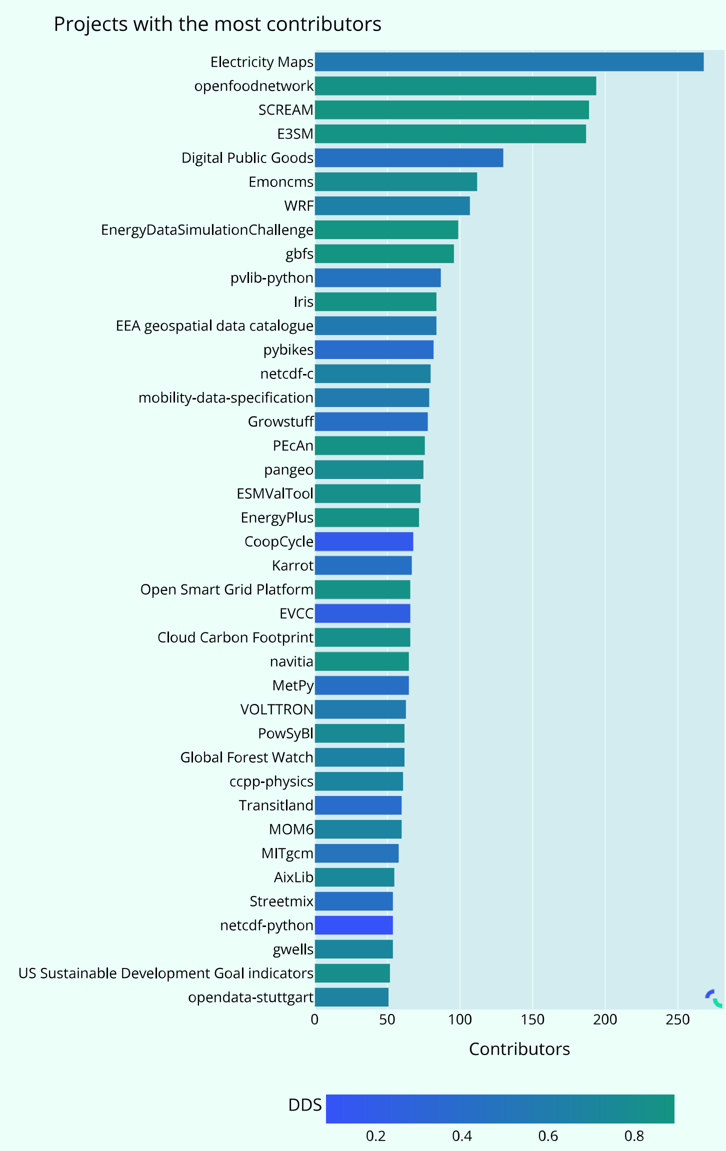


FIG. 14 - Projects with the most contributors

within the community. Only then is it possible to build a large community. Projects that model the atmosphere and climate globally at high resolution can feature many developers — examples being the Energy Exascale Earth System Model (E3SM) and the Simple Cloud-Resolving E3SM Atmosphere Model (SCREAM). Electricity Maps was able to mobilize the greatest number of developers to integrate data on energy consumption and production from local electricity grids on a single global platform.

Organizational Diversity

The details of organizations affiliated with the identified projects have been collected and grouped into six distinct categories:

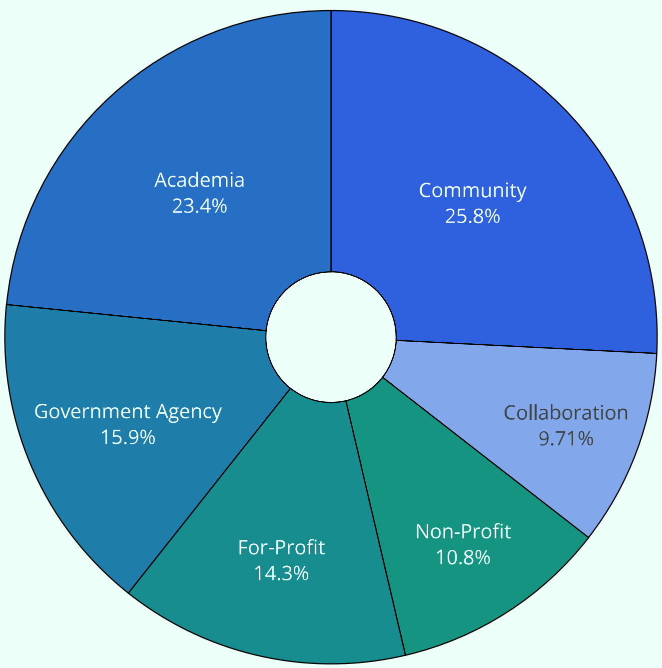


FIG. 15 - The distribution of organizational forms

COMMUNITY-DRIVEN (25.8%): projects which are without institutional affiliation and are led by individual contributors. These projects are also the oldest we have identified. The organizations leading these projects are characterized by a high degree of flexibility and freedom — embodying the origin of the OSS movement itself. The lack of institutional affiliation and support can lead to greater risks in terms of financial sustainability and resourcing.

ACADEMIA (23.4%): projects which are hosted, managed, or developed by academic labs and research institutions. Academic institutions play a critical role in open source sustainability. Universities, particularly research software engineering labs, can provide long-term stability and deep expertise. At the same time, such developments run the risk of bypassing practical usage and losing development capacity due to the high turnover of staff and lack of financial resources within the academic environment.

GOVERNMENT AGENCY (15.9%): projects which are hosted, led or developed by national and/or subnational governments. Open source provides greater control and independence while lowering the risk of vendor or political lock-in, making it easier for governments to plan their digital sustainability futures more holistically. Such projects run the risk of losing resources due to political realignment. The large number of US institutions setting a good example here is remarkable. However, the Canadian Province of British Columbia stands out due to the high number of projects and its own Digital Principles for the Government.

FOR-PROFIT (14.3%): projects which are initiated by private sector entities. Such projects have the potential for rapid growth and offer the opportunity to transfer theoretical knowledge into practical applications. They typically have a high level of resources early on in development but harbor the risk of not being further developed and maintained due to a change in company strategy. Despite the massive capital strength, there are very few good examples of for-profit organizations within open source environmental sustainability. Noteworthy outliers include Electricity Maps, Breakthrough Energy, Vizzuality or Ladybug Tools.

NON-PROFIT (10.8%): projects which are initiated by organizations that do not primarily aim to generate profits for the shareholders but rather pursue charitable goals. For open source development, such an organization can provide the structure for long-term project sustainability if sufficient resources are provided. Several community-oriented organizations became non-profits to provide the legal structure for additional growth of a project or community, such as the Python Software Foundation or the Linux Foundation. Strong nonprofits that take precedence here, are rOpenSci, Reiner Lemoine Institute, OpenClimateFix or Drawdown.

COLLABORATIONS (9.7%): initiatives where projects are hosted, led or developed by a consortium of different actors and institutions. This form of organization is particularly suitable for generating knowledge transfer between different partners. The diversity of different organizations provides the benefit of different perspectives but, without clear leadership, bears the risk of prolonged development cycles. Cooperation at eye level between business, civil society and science have a very high potential to drive sustainable developments for society as a whole. These include collaborations such as Science Based Targets Network (SBTN), California Forest Observatory or the International Building Performance Simulation Association.

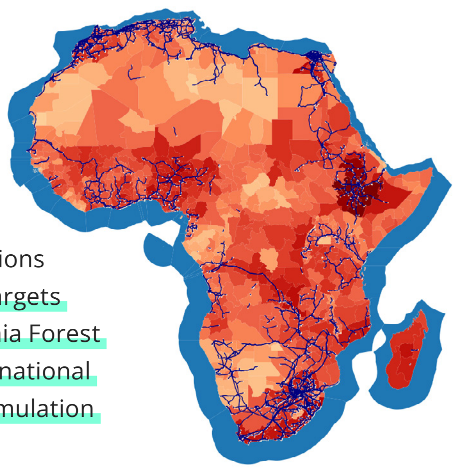


FIG. 16 - PyPSA-Earth is the first open source global energy system model with data in high spatial and temporal resolution. It enables large-scale collaboration by providing a tool that can model the world energy system or any subset of it. License: GPL-3.0

Most Listed Projects

NAMESPACE	COUNTS
rOpenSci	28
National Renewable Energy Laboratory	24
Province of British Columbia	10
oemof community	8
PyPSA	7
National Center for Atmospheric Research	6
European Centre for Medium-Range Weather Forecasts	6
Unidata	6
USGS-R	5
Santander Meteorology Group (UC-CSIC)	5
Climate Modeling Alliance	5
WISDEM	5
Energy Exascale Earth System Model Project	4
advanced network science initiative	4
FZJ-IEK3	4
ncss-tech	4
IBPSA	4
Our World in Data	4
World Resources Institute	4
Open Climate Fix	4
NOAA - Geophysical Fluid Dynamics Laboratory	4
SLAC GISMo	4

FIG. 17 - Organizations with the most listed projects

Geography of Organizations

At a national level, the United States, Germany, France, and the United Kingdom stand out. Despite having more GitHub users than Europe, Asia accounts for only 1.9% of organizations working in OSS for sustainability. Moreover, the absence of Indian communities is notable, with no large organization or project identified, despite a high number of open source developers present. Similarly, very few organizations or projects originate from China, despite a high number of open source developers and a high volume of scientific publications. While there are likely open source developers from underrepresented regions associated with both foreign organizations and global projects with no geographical affiliation, this deviation requires further investigation.

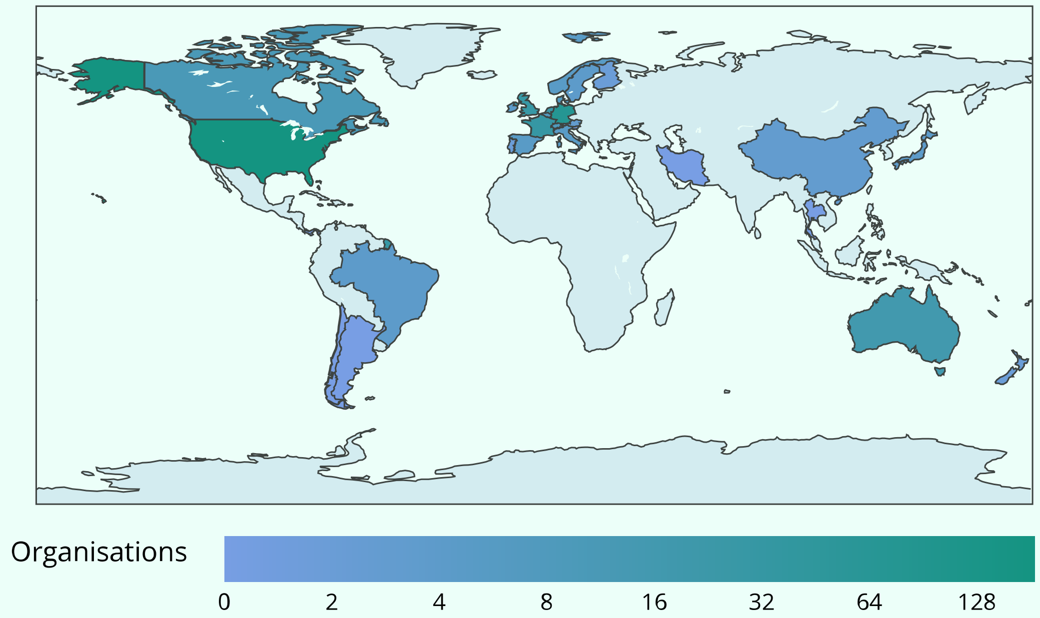


FIG. 18 - Global distribution of organizations

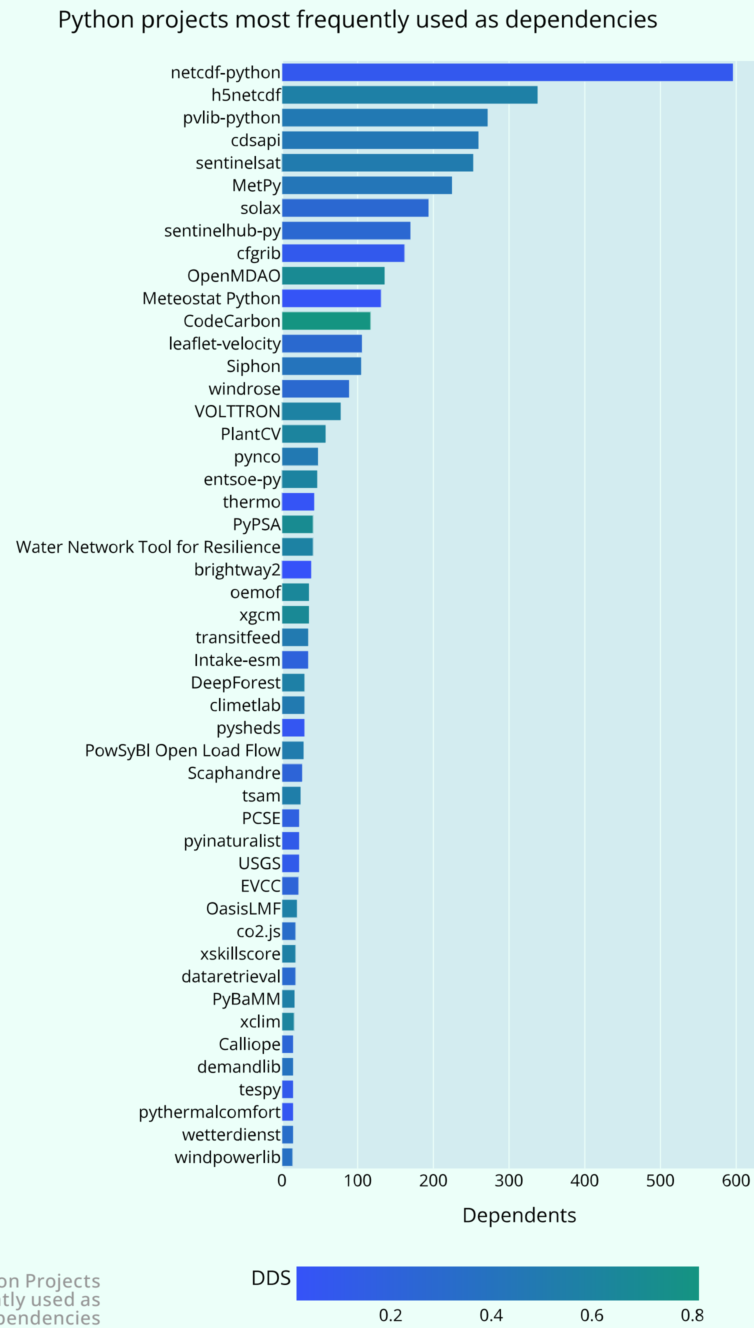


FIG. 19 - Python Projects most frequently used as dependencies

Users and Usage

Recent reports estimate that 95-97%[1,2] of mainstream IT organizations leverage nontrivial OSS assets within their mission-critical IT portfolios, whether they know it or not. Meanwhile, internal government policy is emerging in the USA and EU, “encouraging and leveraging the transformative, innovative and collaborative power of open source, its principles and development practices”[3], with more-and-more government agencies adopting OSS within their operations — including many of the identified projects within this report. When it comes to sustainability, however, there is much potential yet to be realized. While tracking outcomes and impacts associated with OSS remains challenging, the following trends highlight the importance of open source within environmental sustainability for fellow contributors, end users, and greater society.

Dependencies

Since open source is free to acquire and freely available, its ultimate use is difficult to track. Much of the open source usage arises in integrating libraries or APIs as dependencies of other software projects. Importantly, this dependency on OSS will not be apparent to many users; especially in closed-source software, where the dependency on OSS is not always made evident.

Project usage data from public software development and version control platforms remains scarce. GitHub, unfortunately, offers little support in compiling accurate statistics. However, with the limited data obtained from the Python ecosystem, it was possible to identify individual projects with a high circulation but a low DDS score. Here projects like cfgrib, sentinelhub-py or Meteostat stand out. Those projects are widely used and depend highly on the goodwill of a single developer. The median DDS of 0.436 over the 50 most used Python projects indicates that the burden still lies with a few strong contributors leading the development.

The user community of major projects in energy and battery modeling, such as PyBaMM and PyPSA, is split relatively evenly between academia and industry, with fewer users coming from NGOs and independent consultancies. In some cases, industry can drive explosive user growth. For example, over a five-year period, pvlib-python saw thousands of downloads per month. This was driven primarily by several commercial firms who integrated the library into their software products, effectively distributing pvlib-python to their clients.

End Users

Capturing the end users and ultimate use of OSS remains challenging. Due to the complexity of tracing and tracking the cause-effect chain within socio-technical-ecological systems, this report does not attempt to provide an exhaustive evaluation of impact (neither social, economic, nor environmental) of Open Sustainability approaches. Rather, throughout the report, we provide key examples of projects whose outputs indicate a significant capacity towards, and demonstrate alignment with the direction of science-based impact. While many of the identified open source projects contribute to supporting and accelerating sustainable outcomes, some contributions are more evident than others. In this section, we highlight several of these.

A particularly impactful project is pyam, led by the Integrated Assessment Modeling Consortium (IAMC). This project played a fundamental role in the recently published IPCC Special Report on Global Warming of 1.5°C (SR15), producing an ensemble of quantitative, model-based climate change mitigation pathways which underpinning the report. The report has since been cited 106 times and continues to inform governments and decision-makers on the latest climate research and the state of our progress in meeting international climate agreements. Meanwhile, green software management solutions are being adopted at scale within the IT industry. Industry-led non-profit, Green Software Foundation, has seen a high level of backing from tech giants, including Google and Microsoft, who also claim to use these open source tools internally. The Natural Capital Project tool, InVEST, has been applied across sectors worldwide (Africa, India, Central America, USA, China) to answer questions related to land use change, economic valuation of ecosystem services, watershed management, and more. Finally, Global Forest Watch, an open source project led by the World Resource Institute, provides real-time forest monitoring tools to NGOs, governments and decision-makers in an effort to accelerate conservation projects at a global scale. Since its launch in 2014, over 4 million people have visited the platform.

The adoption of open source software and its principles is rarely reflected within official government policy. However, several promising developments can be found. In 2016, the USA Government published a federal source code policy.[4] This policy mandates that at least 20% of custom source code developed by USA federal agencies must be released as OSS and shared between agencies. Likewise, the European Commission's internal Open Source Software Strategy "promotes the sharing and reuse of software solutions, knowledge and expertise, to deliver better European services that benefit society and lower costs to society."[5] While not exclusive to sustainability, this policy development acknowledges the potential for impact in this area.

Monetary Value

While open source business models, including the commonly used premium model, are unlikely to capture the total value that OSS contributes to society, recent research into the economic contributions of open source can provide a rough indication. A 2021 study on the economic impact of open source software and hardware concluded that open source technologies injected €65-95 billion into the European economy. Open source significantly boosted small and mid-size enterprises — Europe's most important horizontal economic stakeholders — reflected in the increased creation of more than 650 technology startups per year. While it is currently unknown to what extent open source has contributed to environmental sustainability, or what its potential economic impact is, we anticipate the monetary value to be several orders of magnitude greater.

Going Forward

Capturing the use of open source projects, and understanding its users and the value they provide to, and obtain from the ecosystem, presents important yet significant challenges. No standards yet exist for measuring the time and resources saved by "standing on the shoulders of giants" and to what degree the added height extended one's view or reach. In future studies, it will be essential to use the various metrics and methods we have created in a targeted way. For example, by identifying projects that have high usage, fast growth and low DDS, we can conduct targeted interviews to pinpoint key users and usage patterns, assess direct and indirect impact, and evaluate ways of supporting effective projects. Until this time, we have considered several proxy methods to help paint a picture of what is possible.

Funding Models

The developers and organizations in this field are supported by a diverse set of funding models. OSS projects benefit from academic and research funding, private sector investment, and government grants. Each project has a different mix of funding models due to the unique structure and ecosystem they are part of.

> "It's a little bit bizarre that there are no government grants available for open source software for public transit."
>
> -DEVELOPER OF POPULAR TRANSPORTATION PLANNING AND TRAFFIC SIMULATION SOFTWARE

Only 3.8% of the projects officially offer GitHub users the option to donate. This is particularly difficult to explain due to the high prevalence of community-based projects. There may be a lack of interest because of the project's low usage and popularity, which in turn means that the developers aren't expecting financial support.

Public sector funders and research councils, together with academic and national labs, play a critical role in sustaining the open source sustainability ecosystem. Universities are not only knowledge providers but can also bring stability, independence and access to developer resources via their Research Software Engineering teams.

However, this ecosystem has few dedicated funds for OSS development and maintenance. Large academic projects prioritize academic research and dedicate small amounts of funding to open source software development. Where there are funds, such as UKRI's Infrastructure Fund in the UK or Horizon Europe, they primarily support new feature development and frequently not at the scale that supports significant strategic software developments. In academic labs, developers and research software engineers are trying to squeeze in person-months in multiple concurrent projects and make this open research infrastructure reusable across many projects. Furthermore, since features are often used across many projects, assigning maintenance or user support resources to any particular OSS development can be complex.

Non-profit and think tank OSS projects face a similar challenge, whereby funding for a new project is relatively easy to obtain, while virtually impossible to raise funding for maintaining software and data infrastructure. Very few funders focus on the use and maintenance phase of the project development lifecycle.

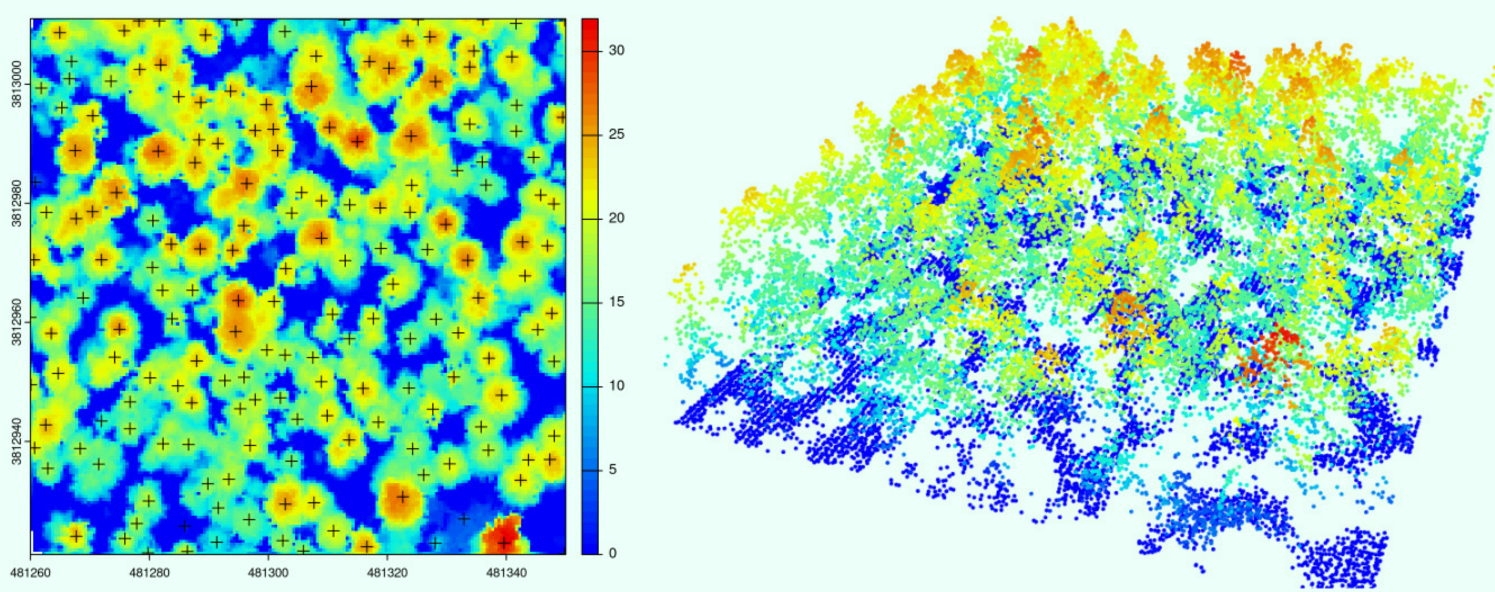
Projects with larger user bases (e.g. openFAST) have a mix of investments, grants and income from industry. This type of project, however, is more likely to incur technical debt, which will increase the costs of both ongoing maintenance and future development. These more mature projects require more significant funding from industry and governments. In many cases, however, a different business structure and legal entity would be required if a lab-led project wanted to accept donations or seek alternative grant funding. This is currently not feasible in government-funded labs, research councils, and academic labs.

Collaboration between industry and academia brings new challenges and opportunities. Academic open source software that has commercial potential typically requires intensive and long-term financial support, which cannot be provided from within academia. Hybrid business models such as Open Core, where the backbone of the codebase is open source, are becoming increasingly popular. Many developers we spoke with had prior experience with Open Core and are acutely aware of some of the trade-offs with this approach. For example, project maintainers want to support open source without commercialization being at the expense of the communities. At the same time, maintainers are aware that their users and contributors may be uncomfortable if they feel companies exploit their open-source contributions for profit. Choosing the suitable open source business model, licenses, developer involvement, and organizational form is crucial to ensure successful commercial application and community development.

Industry-consulting models have had mixed success so far. In some circumstances (like green software in data center operations), open-sourcing the code is a hard requirement from a developer's perspective. In this case, the response from industrial partners has been positive. In other instances, companies wanted to keep the project used to model their systems (e.g., batteries) closed source for several years, as they were worried their competitors would use them for free. However, it may be counterproductive to maintain closed sourced culture while attempting to rapidly scale sustainable outcomes.

Finally, the notion of "soft infrastructure" through digital public goods is gaining traction, which takes a whole-of-society perspective on the impact of OSS. Government agencies fund cross-sector initiatives through research labs that, in turn, collaborate with industry partners to generate new public-purpose value. When these projects can demonstrate the cascading impacts of open source on society more broadly, it becomes easier for these projects to raise funding.

FIG. 20 - lidR enables the segmentation of individual trees based on airborne LiDAR data. Thus, changes and the condition of a forest can be determined precisely.



Recommendations

Overall

Based on the insights above, we propose recommendations for effectively supporting and building capacity within the open source sustainability ecosystem:

ENHANCE COLLABORATION BETWEEN STATE AND NON-STATE ACTORS. State and regional governments can accelerate their digital and sustainability transformations by collaborating more closely with industry, academia and NGOs toward shared objectives. This requires effective transition brokers to guide and facilitate multi-stakeholder alliances from a politically neutral standpoint, and help create the necessary preconditions for change to emerge across scales and system boundaries. Open Sustainability Principles can be applied here in order to:

- Generate advocacy and support
- Support cross-sector and cross-agency coordination
- Promote and guide innovation funding
- Bridge knowledge and information gaps between municipal, regional and state governments
- Identify challenges and opportunities for effective action
- Build capacity for local community groups to take effective action
- Guide stakeholders towards science-based decision-making

CLOSE THE KNOWLEDGE GAP ON THE ENVIRONMENTAL IMPACT OF INDUSTRY. Our analysis has revealed that only a few companies are willing to engage in open scientific dialogue concerning their environmental impact. Despite many companies having net zero commitments and an increased adoption of science-based targets based on credible emission reduction pathways, a lack of clarity and transparency surrounding the measurement, reporting, and verification remains. While standardized calculation methodologies for emission reporting are emerging,[1] companies often rely on estimates which do not fully and transparently reflect their decarbonization efforts. For instance, emissions from industrial facilities are often self-reported or assessed by external parties using unverifiable survey methods, which are prone to error and obfuscation. There is a clear need for an open source measurement, reporting, and verification (MRV) framework that provides a standardized mechanism for aligning sustainability efforts with science-based outcomes. Open source approaches can improve transparency and process in terms of methodological development, provide transparency regarding methods and pathways adopted, and enhance reporting data quality and verifiability. Incorporating openness and scientific verifiability into sustainability assessments and product disclosure statements is one way to accelerate collaborative innovation efforts. Initiatives such as the SBTN, WRI, and UNRISD are well positioned to engage the open source community to accelerate these efforts.

ADAPT AND EXTEND EXISTING OSS TO UNDERREPRESENTED COUNTRIES IN THE GLOBAL SOUTH. Communities in emerging economies have limited involvement in developing open knowledge, software, and data related to their natural environment. Therefore, building and empowering local communities to use open source tools for local technological transformation has far-reaching potential. In fast-growing nations, such as India, many developers have the potential to use OSS to understand their local environments better and thus protect essential ecosystems. Scholarships that teach young people how to use open source to understand the natural environment could greatly benefit the climate and the planet. Such investments are resource efficient and easily transferable worldwide. The Education section on OpenSustain.tech can help build the foundation of the knowledge, projects and people needed. Earthlabs and Pythia already demonstrate what such a format can look like within the field of Earth science.

ESTABLISH AN OPEN EARTH INTELLIGENCE INCUBATOR. Despite the high synergy between open source culture and enabling sustainable technologies, there are few incubators or support programs specializing in open source project development in environmental sustainability. While a handful of individual programs focus on nonprofits and academia, there is a general lack of funding, knowledge, and support for cross-sector collaboration. This trend also explains the low proportion of academic open source projects that are commercialized via open source business models. Such an incubator can fill significant gaps in the ecosystem, such as:

- Identifying opportunities for new ventures and non-profits that creatively combine existing ecosystem components to create more public-purpose value.
- Incentivising the development of new ventures and projects that combine existing modules and packages into effective open source products.
- Providing project support including, marketing, design, funding packages, business model design and community management.
- Building networks with potential users such as cities, government agencies, and nature reserves.

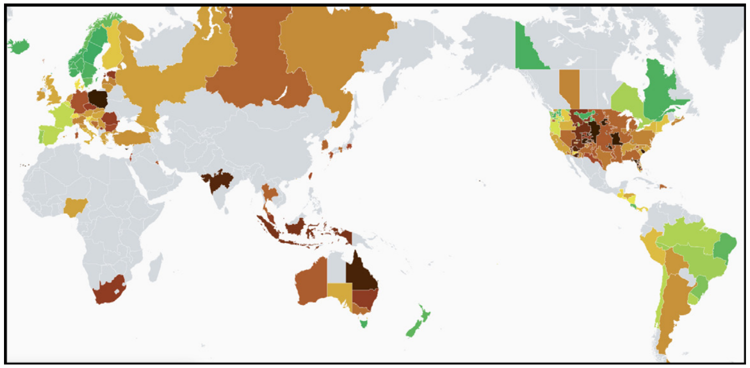


FIG. 21 - Electricity Maps provides time-resolved historical data and forecasts about the amount of renewable energy in the electrical grid for countries all around the globe. Energy consumption can be adjusted in such a way that as few emissions as possible are produced. It provides both a solid business model and an open source community behind the project. License: MIT

TRANSFORM FINANCIAL INSTITUTIONS THROUGH TRANSPARENT AND SCIENTIFIC DECISION-MAKING FOR SUSTAINABLE INVESTMENTS. Banks, rating agencies and investors are at the center of our economies, allocating resources to what is deemed sustainable and viable. However, our analysis has shown OSS and open data is the exception in this sector. The financial industry is still far from deploying sustainable investment funds which are based on open scientific methods. However, such an approach has the potential to create greenwashing-free investment opportunities for private and public investors. Financial products funded on open and evidence-based practices will likely gain investors' trust in creating both sustainable and favorable returns. Banks and rating agencies would thus become critical users and potential contributors to various environmentally relevant open source projects. OS-Climate already represents a significant milestone here.

APPLY AN "OPEN SOURCE FIRST" CRITERION WHEN PROVIDING FUNDING FOR SUSTAINABLE TECHNOLOGIES. Our findings show that open source can have a significant impact on sustainable choices and technology diffusion. However, when it comes to financing sustainable technology projects, open source is often not a decisive investment criterion. A fundamental rethink needs to take place here. Openness must be recognized as a key indicator for sustainable development. In particular, the investment of public funds can help to reverse this trend and ensure that such investments directly benefit the general public in the long term. Government policy that prioritizes open source within public research and development is critical to ensuring publicly funded outputs do not end up constrained by the intellectual property of universities or companies, but rather returned to the commons as public goods.

Technology

MONITOR ENVIRONMENTAL SUSTAINABILITY THROUGH OPEN EARTH OBSERVATION AND OPEN SOURCE DATA PROCESSING. The ever-increasing spatial and spectral resolution of open satellite data will make it possible to identify not only large-scale environmental impacts but also their drivers. This will allow governments, investors and citizens to better understand and predict changes in biodiversity, deforestation, water stress, pollution and many other features related to the health of the environment based on scientifically robust evidence. Satellite missions like TRUTHS will bring the traceability and precision of satellite data to a level which becomes legally robust. Key actors can thus more easily assess asset-level risks, and measure upstream and downstream environmental impacts at a planetary scale. Public-private partnerships are emerging to validate and further develop previous estimates of models using global measurement data. Projects such as Spatial Finance Initiative and WWF have already recognized this potential for informing robust ESG metrics for companies. The development of an open source reporting framework and toolchain for determining the environmental, social, and corporate governance implications of a company's value chain, based on satellite data and spatially explicit supply chain information, has immense potential for companies, governments and investors.

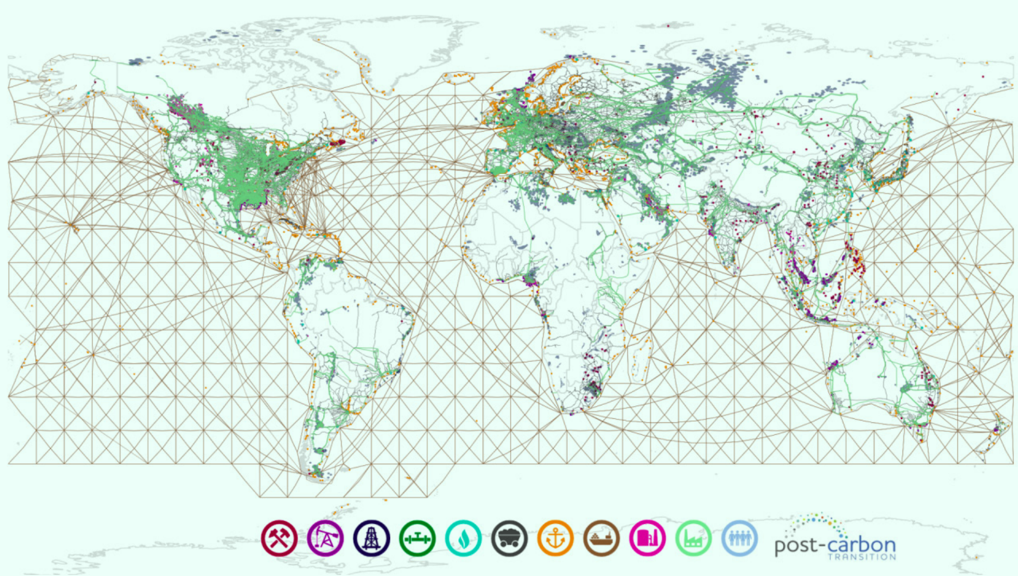
CREATE BETTER TECHNICAL INTERFACES AND MIDDLEWARE BETWEEN PLATFORMS AND TOOLS. Most open source projects within environmental sustainability are used in isolation and are rarely integrated. This leads to increased fragmentation and design waste ("reinventing of the wheel"), ultimately inhibiting the innovation process. There is much that can be learned from other open source ecosystems, such as the Robotic Operating System, where a high degree of modularity has led to increased collaboration between different open source communities from various subject areas. This has been made possible through common interfaces, workshops, community meetings, applications and standard architectures. The same mindset can be applied to different areas within sustainability. Many monolithic projects and platforms do not offer the necessary flexibility for different types of applications. However, a digital Earth twin combining different existing open source projects offers the advantage of leveraging existing knowledge and communities. Cities and communities play a special role here, as a multitude of environmentally-relevant spheres come together to form concrete applications with immediate policy implications. Great potential exists in urban applications where a modular, cross-domain operating system for environmental sustainability enables multi-scale interoperability between processes and services. For subfields such as energy system modeling, earth observation and geosciences, developments such as PyPSA, Julia Climate, Radiant.Earth, WhiteBox, and Pangeo are already pointing in the right direction.

APPLY OPEN SUSTAINABILITY PRINCIPLES TO HARDWARE AND DESIGN BLUEPRINTS. Our findings show that only a few open source hardware projects have a strong connection to sustainability. Although strong communities behind open hardware exist, such as Open Source Ecology, Open Hardware Observatory or Appropedia, the technical requirements for developing and maintaining hardware products are often beyond the capabilities of individual actors, presenting significant challenges. This makes sector-wide approaches and collaborative development and operating models essential to ensuring circular design principles are embodied within each phase of a product's lifecycle. Using this approach, open source hardware has the potential to revolutionize the way we think about the design, production, distribution, maintenance, and end-of-life of physical goods — improving performance, reliability, and cost-effectiveness. One way to ease the distribution and scalability of sustainable hardware solutions is to develop open source design blueprints, digital twins and embedded software. Biosphere Solar is one such organization recognizing this potential in an attempt to produce the world's first circular solar panel, combining an open source solar photovoltaic (PV) design and collaborative business model. With the advent of RISC-V, a fully open-source processor architecture and ecosystem, this approach is already transforming the computing industry. Projects exploring similar potential in sustainability include Libre Solar, OpenEnergyMonitor, OpenEVSE and FarmBot. However, only the IEA-15-240-RWT open source wind turbine demonstrates the strategy and scale required to transform an entire sector. This highlights the vast potential for businesses, governments, researchers, and industry to embrace cross-sector opportunities to accelerate the open hardware and software ecosystem towards sustainable outcomes.

DEVELOP OPEN DATA COMMONS IN CONJUNCTION WITH OPEN SOURCE CODE. All interviewed developers and contributors have an intimate understanding of the data landscape, as well as the quality, provenance and accessibility or lack of open data in their respective fields. While open data platforms, such as Zenodo, are central to big data management within the open science movement, links within topics are often sparse and lack an ontology to allow for easy discovery. Organizations such as Subak can serve as a critical link, investing in and stewarding missing data across various topics, and enabling technologies within environmental sustainability. Establishing connections between open source code and datasets can help create horizontal applications and thus make them more applicable worldwide. Likewise, a systematic analysis of the datasets associated with the identified open source projects can help close the gap between project outputs, usage and open source development.

STANDARDIZE ENVIRONMENTAL DATA AND MODELS USING OPEN SOURCE ACROSS DIFFERENT LEVELS OF GOVERNMENT. The standardization of data structures and APIs using open source approaches can contribute significantly to ensuring that data about our natural and built environment are interoperable and deliver valuable insights. The province of British Columbia is a pioneer in this regard, delivering a variety of open source and open data developments on water supply, wetlands, and air quality. Such standardized environmental data is vital not only for scientific analysis but also for intelligent monitoring and optimisation of public utilities and services. However, the potential for such data can only be realized if the collection and provision conform to common standards. The Smart Data Models project, supported by the FIWARE Foundation, is one example of such an open standard. Such open data models play a vital role in the technical foundation required for standards-based innovation and procurement, while ensuring trusted exchange and data sovereignty within and across sectors, places and organizations. Consolidation and dissemination of environmental insights across different provinces, cities, and municipalities can break down information siloing, and contribute significantly to a shared knowledge base for decision-makers, researchers, businesses and citizens alike.

UNDERSTAND RESOURCES MOVING THROUGH SUPPLY CHAINS USING OPEN SOURCE APPROACHES. Digitisation enables visibility into supply chains, allowing stakeholders to track and trace the flow of materials and goods, understand their movement from point to point, and measure their impact. For instance, product passports linked to QR codes and RFID systems allow companies and consumers to track a product throughout its lifespan, improving transparency for a wide range of resources — from consumables to construction materials. These tools can provide information on the state of a product at any given point in its lifecycle, such as when a plastic container arrives at a recycling facility, or information about a building component currently in-use. More detailed traceability can be achieved by generating a "digital twin" of a product, including information about its material composition, disassembly instructions, and labor practices. In order to create circular and efficient supply chains, such visibility is essential. Open standards such as the circularity.ID are pointing in the right direction. However, there is a lack of open infrastructure optimized for the digital age. The convergence of accounting frameworks, such as the System of Environmental Economic Accounting (SEEA), Resource-Event-Agent (REA), and other industrial ecology approaches, together with decentralized messaging protocols, show promise in linking actors and system layers in a secure way. We encourage the adoption of open principles and technologies in providing the level of traceability and visibility required to ensure supply chains are efficient, fair and sustainable.



CREATE OPEN COMMUNITIES AND FRAMEWORKS FOR MONITORING GREENHOUSE GAS EMISSIONS THROUGH REMOTE SENSING. Free access to satellite data from Sentinel-5P, GOSAT, GOSAT2, OCO1 or OCO2 could provide the basis for such collaboration. However, in many cases, the integration of these data into a common format is costly. There is also a lack of powerful and open transport models that can trace these emission measurements back to individual point sources. Although open repositories such as STILT and X-STILT exist for these models, open licenses and communities are lacking. The first important steps in the right direction are shown by Emissions API and oco2peak. Monitoring point emissions through open and traceable satellite data and models has the long-term potential to prevent the obscuring of major emission sources and to attribute them to polluters.

Collaboration

FURTHER STRENGTHEN THE INTERCONNECTIVITY OF COMMUNITIES. A flourishing ecosystem of knowledge sharing and exchange between communities relies on networks and platforms where projects, ideas and news can be openly distributed. Prior to the formation of OpenSustain.Tech, many organizations, users and developers found it difficult to navigate active projects in and across topics in OSS for sustainability. Strengthening the interconnectivity of communities can alleviate contributor risk, provide opportunities for greater participation, and build capacity for more effective collaboration. The emergence and convergence of diverse groups play a vital role in consolidating and further developing shared values, visions, mindsets and ideas embodied within Open Sustainability Principles.

FIG. 22 - Asset-level Transition Risk in the Global Coal, Oil, and Gas Supply Chains. License: MIT

PROVIDE MAINTAINERS WITH TRAINING AND SUPPORT TO PRESERVE OPEN SOURCE PROJECTS. Maintainers must be trained in both software development as well as community and project management. Many lead developers are experts in their fields, however often lack the knowledge and experience required to build and maintain healthy communities around their projects, with marketing and other entrepreneurial activities often being neglected. In many cases, it is assumed that a high-quality project will quickly find users and contributors, which is not necessarily the case in practice. Unfortunately, there is a lack of training materials and opportunities tailored to the unique characteristics of OSS in sustainability. Providing documented approaches and training and support can help ensure that maintainers can build the community capacity required to preserve open source projects in the long term. One way to facilitate this is by establishing an open source program office within an organization.

CONNECT PROJECTS TO LOCAL USE CASES. Cities can play a central role in the application of various open source technologies as they translate research and development into action on the ground. However, the knowledge and the skills required to leverage the full potential of open source approaches are often lacking at this level. At the same time, scientists and community groups working directly with local authorities will be key in bridging gaps between research, urban application, and policy development. All parties can benefit significantly from such a partnership. Local authorities can provide researchers and civic society with open data about the natural and built environment through a secure and standardized interface, while researchers can share resources and insights for local application and integration. Such open science approaches can enhance cross-sectoral collaborations, essential to answering important research questions and the continuous improvement of open source technologies. These multi-stakeholder alliances can build trust and transparency in the exchange of information, enhance data-driven decision-making, and foster a co-creation process with citizens and businesses based on local conditions and needs.

MAINTAIN AND DEFEND AN OPEN ORIENTATION WITHIN ACADEMIA. While many of the projects identified within this study are the outputs of publicly funded research, many research institutions are yet to embrace open science and promote open source as the default position. While the protection of intellectual property remains common practice, research constantly demonstrates that regulatory and legal obstacles actively hampers innovation.[2,3] However, there is a clear desire among academics to abandon outdated intellectual property models. The majority of American and Canadian academics support an open orientation, encouraging universities to establish open source endowed chairs. Western Sydney University is one of the first academic institutions to make an official open source commitment of this kind. With more universities incorporating the Sustainable Development Goals into their internal strategies, we encourage both academics and administrators to evaluate their contribution towards these Global Goals within the context of open source and public-purpose value, particularly with respect to environmental sustainability. In light of this, we recommend that all publicly funded research within the fields of sustainability be made open access and open source by default, for the benefit of the people and the planet.

SUPPORT THE USE OF OPEN SOURCE PRODUCTS AND SOFTWARE DEVELOPMENT WITHIN GOVERNMENTS. Open source approaches can give government agencies better control over technologies and sustainable outcomes. While the adoption of open source software and its principles is rarely reflected within official government policy, several public institutions have official measures in place to ensure the effective use of OSS. In 2016, the USA Government published a federal source code policy, The People's Code. This policy mandates at least 20% of custom source code developed by USA federal agencies must be released as OSS and shared between agencies. Likewise, the European Commission's internal Open Source Software Strategy "promotes the sharing and reuse of software solutions, knowledge and expertise, to deliver better European services that benefit society and lower costs to society.". The benefits of using OSS products and open source software development across government towards sustainable applications are vast. Direct benefits include reducing the total cost of ownership, preventing costly vendor lock-in (at the expense to the taxpayer), improving digital autonomy, enhancing multi-scale and cross-agency interoperability, bolstering the security of infrastructure and digital services, and enhancing the co-design of digital services through community participation. The Open Source Software Guideline, published by the Queensland Government, highlights the expected benefits of using and developing open source software within government, and provides information for agencies considering adopting a similar approach.

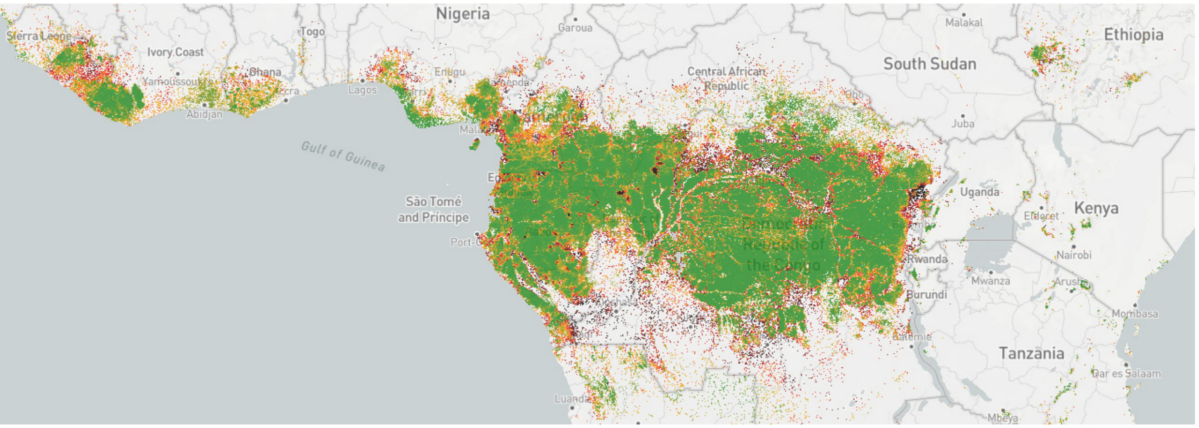


FIG. 23 - The forestatrisk Python package can be used to model the tropical deforestation spatially, predict the spatial risk of deforestation, and forecast the future forest cover in the tropics. License: GPL-3.0

Funding

DEPLOY DEDICATED FUNDS FOR CORE DEVELOPMENT AND MAINTENANCE. Stable and secure funding for core product development is foundational for maintaining ongoing, high-quality project support. Maintenance funding for OSS for sustainability should be given extra attention. This may include allocating funds to:

- Develop documentation, code libraries, training, or other resources that make it easier for more people to manage and develop and extend existing technology
- Develop robust solutions to recurrent short-term problems such as solving legacy problems
- Complete migration from legacy code
- Improve the design and user experience of technology
- Establish or adopt technical standards

The insights, methods and tools developed within the framework of this study can contribute significantly to making such investments transparent and targeted.

EMBED OPEN SOURCE PRINCIPLES WITHIN PHILANTHROPIC AND IMPACT INVESTMENT. Organizations across sectors are increasingly recognizing their own reliance on open source, with over 97% of modern digital infrastructure leveraging open source code.[4] There is also a realization of the material benefits organization and societies reap from adopting and engaging in open source — both directly, as well the co-benefits of collaborative innovation. From the largest health management information system in the world to pandemic tractability and climate change adaptation tools, the scale and impact of open source is vast. However, direct contributions to this growing shared resource is rarely considered as a desired outcome within philanthropic and social impact circles. Our findings reveal that a bottom-up convergence of digital and sustainability transformations are crucial to bridging the digital divide and achieving an inclusive, regenerative and circular economy. We encourage philanthropists and impact investors to recognize the transformative properties of open source and its cascading effects, particularly within environmental sustainability, and embody these principles within their funding criteria. Examples of funders embracing this collaborative orientation, seeking to maximize the impact of their resources, include Climate Subak, the Open Technology Fund, and the community-driven platform, Open Collective.

MAKE OPEN SOURCE A PRIORITY FOR GOVERNMENT INVESTMENT. Governments at all levels can play a central role in building capacity for OSS in sustainability ecosystems and beyond by making open source a priority when developing internal systems, within the procurement process, and when funding research and development where software and hardware are outputs. Importantly, building interdisciplinary and transdisciplinary capacity around open source information and technology has cascading effects across business, academia, industry and the community; stimulating endogenous growth,[5,6] labor productivity,[7] and the formation of start-ups.[8] Increased economic complexity and resource efficiency gains are also evident.[9,10,11] OSS presents clear environmental, economic, and social advantages, and should therefore be seen as a key component of effective digital innovation or sustainability strategies. With many governments acknowledging the Sustainable Development Goals within their internal policy, we encourage policy and decision makers to assess their contribution towards these Global goals within the context of open source and public-purpose value, particularly with respect to environmental sustainability. New alliances, such as Digital Public Goods Alliance, play an important role in generating a shared understanding of the benefits of supporting and fostering open source, together with governments at scale.

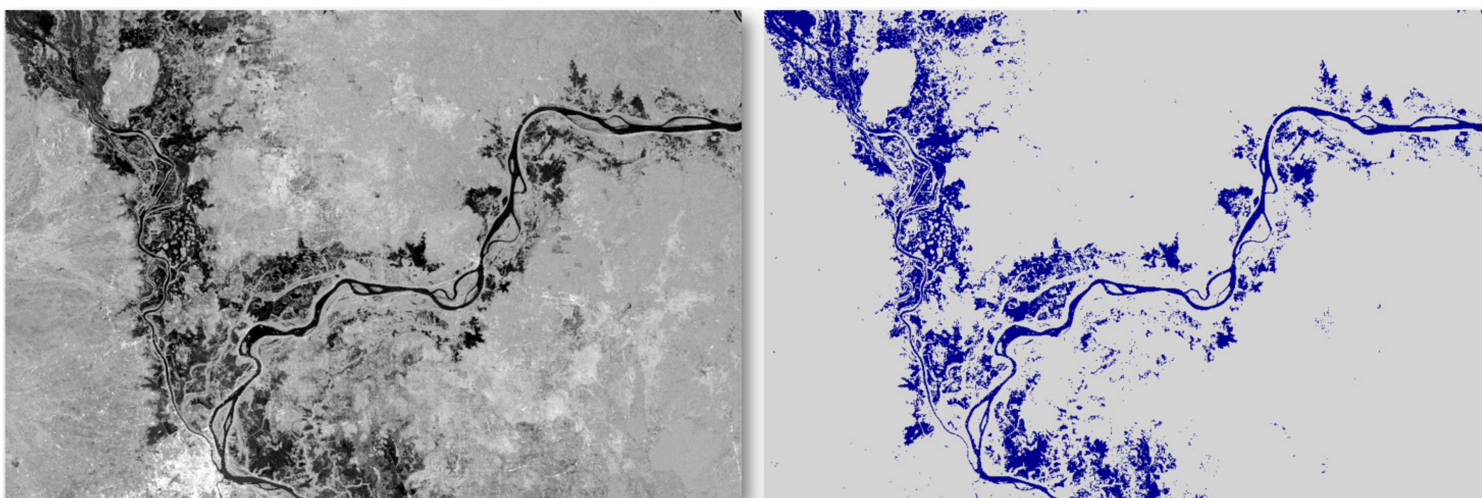


FIG. 24 - HYDRAFloods is an open source Python application for downloading, processing, and delivering flood maps derived from remote sensing data. License: GPL-3.0

Acknowledgements

Thanks for all the valuable insights and interviews:
Jackson Hoffart (Rocky Mountain Institute), Daniel Huppmann (International Institute for Applied Systems Analysis), Benoit Petit (Hubblo), Ryan Abernathey (Earthmover), Tom Brown (PyPSA), Clifford Hansen (pvlib), Valentin Sulzer (PyBaMM), Jenni Rinker (DTU Wind and Energy Systems), Julia Wagemann (Independent), Dustin Carlino (A/B Street), Shuli Goodman (Linux Foundation Energy), Rafael Mudafort (US National Renewable Energy Laboratory), Joe Hamman (CarbonPlan), Sylwester Arabas (PySDM), Trystan Lea (Open Energy Monitor)

Special thanks to Katarzyna Kulma for reviewing and contributing to the data analysis and visualization.

Thanks to the contributions and support of:
Tjark Döring, Alejandro Aristi, Victoria Lo, Nithiya Streethran, Felix Dietze, Johannes Karoff, Joe Torreggiani, Malgorzata Augspurger, Miriam Winter, Lars Oliver Zlotos and Nick Fiege.

This report would have not been possible without people from all around the world contributing to OpenSustain.tech.

Appendix

Methodology

Developing and maintaining open source software relies on an integration of technical, social and organizational factors. A mixed methods study was therefore designed to understand all aspects of this socio technological system. We collected quantitative and qualitative data concurrently and used one source's findings as cross-validation for the other. Thousands of projects were analyzed across six dimensions:

- **GENERAL OVERVIEW** — project themes, growth, age and popularity, and general project maturity.
- **TECHNOLOGY** — preferred programming languages, license use, and technical challenges facing contributors.
- **COMMUNITY** — community composition and participation, and overall activity.
- **ECOSYSTEM COLLABORATIONS** — cooperation across and within sectors and disciplines.
- **FINANCIAL SUSTAINABILITY** — business models and funding mechanisms.

Quantitative Analysis

From October 2020 until August 2022, ~1,300 actively developed open source projects were crowdsourced and curated. All entries were selected based on the contribution guide of OpenSustain.tech. For a project to be listed, it must:

1. Follow at least one aspect of the Open Sustainability Principles.
2. Be instrumental to the preservation and restoration of natural ecosystems, support climate change mitigation or adaptation, or enable environmental sustainability more broadly, through open technology, methods, data, knowledge, intelligence or tools.
3. Be used by others outside the core project or organization.
4. Be structured and documented in a way that allows mainte-nance, reuse and extendability.
5. Be published under an open source license.

The project dataset is entirely machine-generated based on data from OpenSustain.tech and metadata from the GitHub API. Due to the limitation of the database, we assume that a project takes place in a single repository. If multiple repositories could be identified as belonging to a single project, the main repository was identified and listed. For this reason, the terms repository and project are used analogously in this study. When entering the projects, care was taken to use the main repository of a project. The methodology and code used to parse and analyze the data are available in the AwesomeCure repository. The scripts for generating the plots can be found in the repository of the study. Several strategies were tested to include as many projects as possible in data collection using multiple keywords related to sustainable technology and environmental sustainability:

- Searching OSS platforms like Gitlab, GitHub, Bitbucket or Zenodo.
- Mining scientific papers for terms like git, and searching in each paper's keyword dictionary.
- Using search engines for OSS, such as Libraries.io, PyPi or rdrr.io.
- Investigating OSS related journals such as the Journal of Open Source Software.
- Crowdsourcing input, and interviewing people working in relevant domains.
- Browsing stars awarded by developers in the field.

Despite extensive research and a comprehensive use of complementary strategies, this database is only representative of a subset of projects within this domain, and should therefore not be considered exhaustive. We must acknowledge that several technological developments relevant to OSS in environmental sustainability are not directly related to outcomes in environmental sustainability. Instead, they provide the technical foundation that enables this software. The lines here are frequently blurred, and the extent to which a development contributes to environmental sustainability is difficult to predict. For example, a geoscience development that contributes to the exploitation of oil fields could also contribute to the exploitation of geothermal energy.

Furthermore, some important attributes, such as the number of clones and downloads could not be collected via the GitHub API. Projects on self-hosted Git platforms are also more difficult to find. While not currently supported, future re-analysis will consider other platforms, such as GitLab, more extensively. The number of open projects which reuse an open source software project is another important attribute. We were able to obtain this data for Python projects via GitHub using a web crawler. Other data, such as user permissions, cannot be viewed in individual projects without additional authorisation. For this reason, governance structures in most open source projects are difficult to determine. Individuals who did not contribute code were excluded from our analyses. Even if such contributions are critical, they cannot be obtained through the GitHub API at this time.

Development Distribution Score

In this study, a proxy is developed to quantify the Bus Factor, the Development Distribution Score (DDS). The DDS weighs how the development is distributed between project contributors by benchmarking the individual with the most commits in relation to the other developers. DDS seeks to measure how distributed the knowledge, work, and governance is in relation to other projects. The metric compares a project's reliance on a small group of contributors and, therefore, its resilience to change. It can be seen as a lead indicator of health and complexity, whereby the greater the diversity of knowledge accumulated and distributed in a community, the greater the community's resilience and productive capabilities. More details about this can be found in chapter Development Distribution Score.

Qualitative Analysis

OSS communities are keystone actors of OSS projects. They are typically initiated by an individual or groups with specific needs often not met by existing solutions. We conducted 15 interviews with developers and contributors from projects of various sizes and fields, including environmental economics, sustainable finance, climate and earth science, energy system modeling, renewable energy, batteries and transportation. Because we used a concurrent rather than sequential triangulation strategy, we had the opportunity to revisit and enhance our model to account for information revealed by the interviews. We drew inspiration from the questions asked in Roadwork ahead to enquire about the developers' challenges, incentives, needs, the financial viability of their projects, and the barriers that have hampered the development of best practices. We asked about:

- **TRAJECTORIES AND POSITIONS IN OSS —** What are you working on, and how did you end up there? Where do you see your project related to the broader open sustainable ecosystem?
- **TECHNOLOGY & SUPPORT —** What open datasets and tools are you missing? What do you need to maintain your project in the near term?
- **COMMUNITY —** How many users does your project have, and how is this metric tracked? What efforts are made to build a diverse developer base? What is the developer model? How are developers retained?
- **COLLABORATION —** In which field would you like to see more collaboration?
- **FINANCIAL SUSTAINABILITY —** What efforts are being made to reach your definition of sustainability? What are your sources of funding or sponsorship?
- **FUTURE OUTLOOK —** Do you see your project being widely used by your community in the future? What are the top 5 open sustainable projects on your radar?

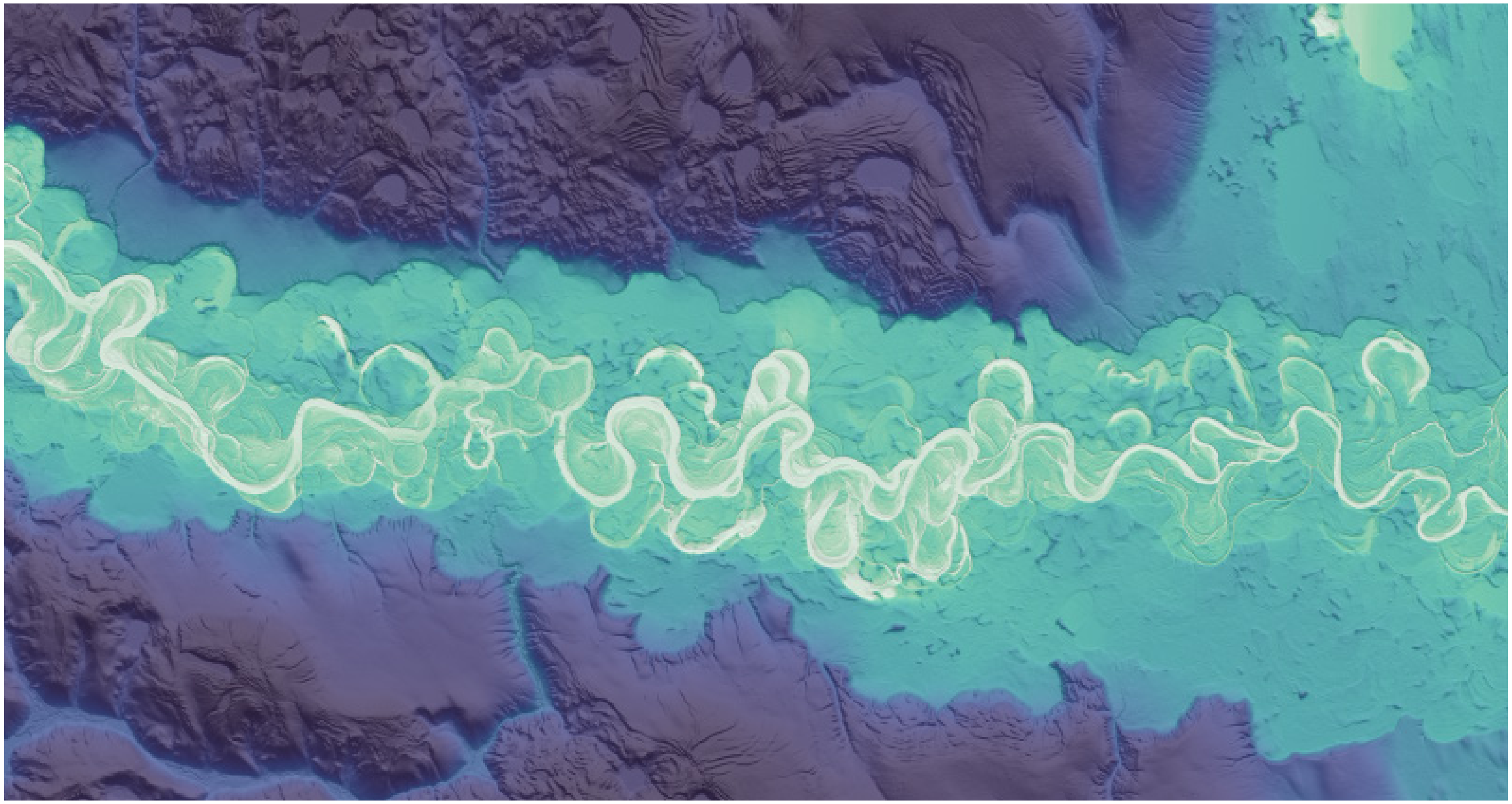


FIG. 25 - Beaver Creek, a tributary to the Yukon River, Alaska, USA. Visualization created with the open source Python package RiverREM. License: GPL-3.0

Definition of terms

When it comes to complex (socio-technical-ecological) systems, many different perspectives can be taken depending on the context. This study examines the relationship between three intersecting dimensions — open source culture, technology, and environmental sustainability — from the perspective of 'open sustainability'. This term must be clearly distinguished from other similar concepts relevant to this study.

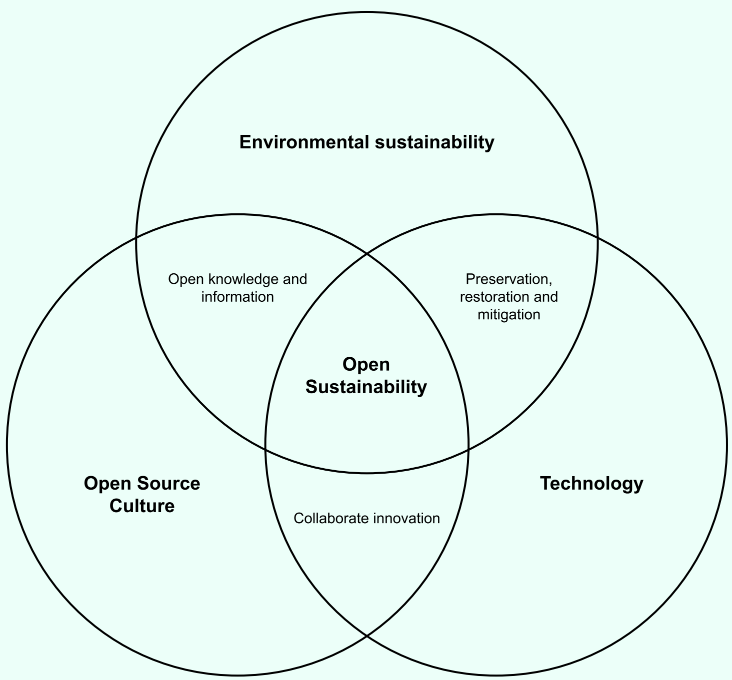


FIG. 26 - Dimensions of open sustainability

SUSTAINABILITY is a concept that is concerned with meeting the needs of the present without compromising the resources of future generations. Sustainable systems are those that "meet the current needs of many individuals involved in producing, deciding, and using a commons without compromising the ability of future generations also to meet their needs".[1] In this sense, sustainability is achieved "as long as the average rate of withdrawal does not exceed the average rate of replenishment".[2]

OPEN SUSTAINABILITY refers to the use of open cultural and technical approaches towards sustainable outcomes. The concept examines how open source culture, technology and methods contribute to all three dimensions of sustainability — ecological, economic and social. It is concerned with how openness as a philosophy is instrumental to sustainability through the acceleration of transparent and collaborative innovation. Within this study, this concept is dealt with in detail, with a particular focus on environmental sustainability.

Previous definitions of open sustainability are broadly focused on open innovation within the context of sustainable development, often purely from an internal corporate perspective.[3] We consider open source as an essential component of transparent and collaborative innovation, therefore, provide a more narrow definition that explicitly embodies open source culture and methods, without restricting its application to any one actor. Open source appropriate technology (OSAT), coined by the Open Design movement, is another related term. Appropriate technology — considered to be largely sustainable, small and appropriate — is within the scope of open sustainability. However, open sustainability is not prescriptive concerning the design specifications or implementation details of technologies. Instead, it defines methods and guiding principles from which similar technological attributes can emerge.

DIGITAL SUSTAINABILITY is another related term, defined as the "process of applying social, economic, and environmental stewardship principles to digital products, services, and data delivered via the internet".[4] It is often included within digital transformation strategies. While digital sustainability is an aspect of open sustainability, open sustainability is predicated on openness and makes this explicit. Furthermore, digital sustainability is constrained to digital products, whereas open sustainability is not.

SUSTAINABLE TECHNOLOGY is a broad term most commonly used to refer to clean energy sources and systems that minimize environmental impact. For example, so-called "green software", which is concerned with lowering the energy consumption and carbon intensity of cloud computing, would be considered a sustainable technology. Whereas an energy-intensive climate model may not be sustainable in its operation, it can provide critical insight that informs sustainable decision-making with vast implications.

OPEN SOURCE SUSTAINABILITY refers to the ability of an open source project and its community to sustain themselves over a long period of time and to adapt the project to new circumstances and technologies. A strong community and governance structure is central to delivering bug fixes and new features.

Open source sustainability is widely known and analyzed in detail within numerous books and publications.[5] A whole ecosystem has emerged with the objective to commercialize the support, risk analysis and funding of open source software. Within the Linux Foundation, a new community emerged to determine the health state of an open source project: Community Health Analytics Open Source Software (CHAOSS). Another strong community that has formed in this space is Sustain. This work is important, as large, well-known projects typically receive more donations than small ones. However, small modular projects can be critical to the global digital infrastructure. Even though large companies use these projects, donations are small compared to the development resources that are required.[6]

August 2023

To reference this work, please cite as follows:

Tobias Augspurger, Eirini Malliaraki, Josh Hopkins, Dan Brown, “The Open Source Sustainability Ecosystem, The Linux Foundation, August 2023.