Open Source in Environmental Sustainability

Contents

Preserving climate and natural resources with openness

Tobias Augspurger, Eirini Malliaraki and Josh Hopkins
Report 2023

> *“The struggle to understand and steer the interaction between the bitsphere and the biosphere is the struggle for community in the broadest ecological context.”*
> Ursula M. Franklin, The Real World of Technology (1989)

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 provides 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 organisations were compiled and systematically analysed 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 organisational – 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.

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 modelling, transportation and buildings. Other topics, such as carbon offsets, battery technology, sustainable investment, emission observation and integrated assessment modelling, 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.

Social rankings, derived from software development platforms, show a total of ~127,000 stars across all identified projects, with a median of 42 stars. As a proxy for popularity, the total stars suggest this ecosystem attracts moderate interest amongst open source contributors. Some emerging topics, such as green software, have shown rapid popularity growth in recent years. In comparison, however, the top 27 software projects hosted on Github each have more stars than the entire open source ecosystem in environmental sustainability combined, with a median of ~161,000 stars. Only three of the identified projects have more than 1,000 stars. This highlights a general lack of awareness of this critical ecosystem within the broader developer community and public domain.

Perhaps unsurprising to developers, the programming languages Python and R dominate this open source ecosystem, indicating a strong focus on data analysis, scientific computing, and statistical modelling. The ecosystem maintains an open orientation from a legal perspective, favouring permissive licences such as MIT (26%), allowing others to use and share software with few restrictions, followed by the copyleft licence GPLv3 (17.3%).

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 organisations and startups with open source business models is remarkable, particularly given the rise of similar ventures within other domains. This sectoral imbalance is further highlighted by the growing demand for open science approaches within corporate environmental, social, and governance (ESG) strategies. Geographically, most open source software projects within the ecosystem are based in Europe and North America (64%), with few projects from the Global South. In India and China, notable development efforts appear either underrepresented or nonexistent, despite having large software developer communities.

While capturing the ultimate use of open source technologies remains a challenge, the report highlights key examples of projects which demonstrate alignment and capacity towards positive impact. Many of the identified projects both enable and accelerate sustainable outcomes. These range from the consortium efforts behind models underpinning the well-publicised 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.

Beyond environmental considerations, our findings support recent research suggesting that open source is instrumental in increasing efficiencies through iterative innovation and avoiding the need to "reinvent the wheel". Importantly, building capacity around open information and technology has been shown to have cascading effects across the whole of society, stimulating endogenous growth and the formation of start-ups, as well as boosting labour productivity. However, there are few dedicated funds for open source development and maintenance. There are even fewer funders who have adopted an open source criterion for sustainable investment. Furthermore, while recent policy developments acknowledge the wider potential of open source, there is a lack of cross-sector advocacy and support for this critical ecosystem.

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, nongovernmental organisations, 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 software in environmental sustainability:

- 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”.
- Standardise 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.

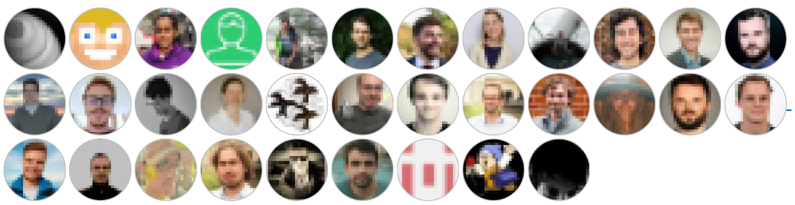
Acknowledgements

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 visualisation.

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:



OpenSustain.tech is a community-driven, non-profit project. However, we would like to thank the organisations that provided us with financial and consulting support.

How to Contribute

More than ever, free and open source projects are enabling citizens, scientists, developers, civil society, industry and government to mitigate climate change. We want to hear from you if you:

- Like to participate in future studies of this form.
- Have experience developing, supporting or systematically using open source software for sustainability applications.
- Want to contribute to OpenSustain.tech by identifying new and missing projects.
- Have experience visualising or processing data with Python and know how to integrate such data into a new website.
- Are a funder and want to support these developer communities via open infrastructure funds, consortia-based support or other collaborative models across institutions and regions.
- Want to help us build any of the recommendations and future directions of OpenSustain.tech.

For these and any other enquiry please reach out via email to Tobias, Eirini or Josh:

Stay Informed

To stay informed about this projects or to connect with the community please join our Gitter Chat, follow us on Twitter, GitHub or LinkedIn.

BibTeX Citation

```
@book{augspurger-malliaraki-hopkins,
    Author     =  {Augspurger, Tobias and Malliaraki, Eirini and Hopkins, Josh},
    Date-Added =  {2023-01-10},
    Title      =  {Open Source in Environmental Sustainability},
    Year       =  {2023}}
```

Table of Contents

Prelude

Gathering

In 2019, Eirini Malliaraki curated a collection of open data in the field of environmental information. Data for Environmental Intelligence gives an overview of open and freely available resources on the state of our natural environment. It was created with the aim of helping developers, journalists, scientists – ultimately, anyone with access to an internet connection – to find information on how our planet's natural systems are changing.

In the same year, Tobias Augspurger curated a list of open source robotic tools for professional robotic development with the Robotic Operation System (ROS). Based on his experience leading the development of a complete robotic software stack at DHL, primarily based on open source tools, he became convinced that *open source will play an increasingly important role in the co-evolution of technology and society*.

Together with the online community *protontypes*, they launched the Open Sustainable Technology project to map all open source software repositories in the field of environmental sustainability. Their goal was to capture a complete picture of this emerging open source ecosystem. It quickly became clear that a wealth of information could be derived from this database that warranted further investigation.

In September 2021, Eirini and Tobias joined forces towards a common goal: to explore and understand the intersection of open source, environmental sustainability, and technology, and to share their findings and insights with the public. Further collaborations were agreed upon to build a global picture of the open source ecosystem. In 2022, Josh Hopkins joined to investigate open source projects in the context of digital public goods for communities and decision-makers, led by sustainability intelligence non-profit organisation Open Corridor.

Fig. 1 - Open source software and Open Data are the foundation of many modern innovation processes. This artwork is generated by the open source artificial intelligence Stable Diffusion.

The Role of Open Source Software

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.

While there is no central authority by design, the open source movement is guided by common principles, such as the Fundamental Principles, defined by the Open Source Initiative. At the same time, the Open Science movement and FAIR Principles have been instrumental in making scientific work transparent and collaborative.

Open source software has transformed industries, allowing for the rapid development of many of the advanced systems which our economies and societies rely upon. Today open source is dominant in 90% of cloud computing,[1] 82% of smartphones,[2] 97% of application development,[3,4] and most artificial intelligence and data science.[5] Linux is by far the most widely used operating system in embedded systems, internet infrastructure, and supercomputers, with all of the top 500 most powerful supercomputers in the world running a Linux distribution.[6] Not surprisingly, 90% of the Fortune Global 500 leverage OSS.[7]. This trend is also evident in venture capital investments in commercial open source software with a total investment of $24B between 2020 and 2022.[8] Several larger companies with annual revenues of more than $100M are on the rise.[9] The total market for open source services is expected to reach $66.84B by 2026.[10]

Tip

Click the image to zoom into the dependency tree.

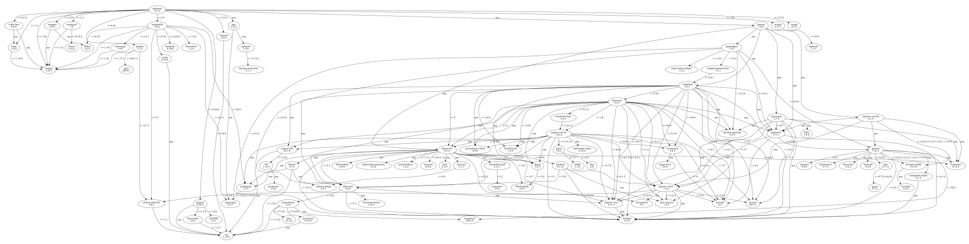


Fig. 2 - The dependency tree of [PyBaMM](), solving physics-based electrochemical DAE models by using state-of-the-art automatic differentiation and numerical solvers. The plot is created with [pipdeptree]()

Without the high modularity made possible by open source, software development as we know it today would not exist. 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. Modern software development is based on the developments of thousands of small projects that have been released as open source over the last decades. Fig. 2 illustrates the connection of a typical open source project with several other releases within the ecosystem.

Besides core information technology infrastructure, open source software is becoming increasingly critical in environmental science,[11] disaster impact assessment,[12] energy efficiency,[13] and sustainability[14] 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 and predicts complex natural systems including the atmosphere, biosphere, and hydrosphere.

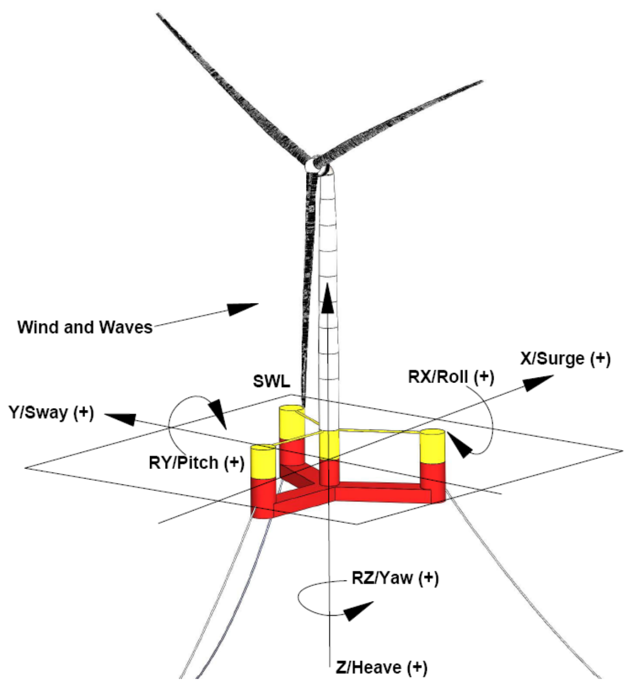


Fig. 3 - The International Energy Agency is working with the National Renewable Energy Laboratory, Technical University of Denmark and the University of Maine to develop an open source 15-megawatt offshore reference wind turbine design with open simulation and development tools. License: Apache License 2.0

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 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 gases 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 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 their potential impact. The manipulation and withholding of information related to the environmental impact of technologies and companies 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.

A Brief History of Open Source Culture in Sustainability

Open environmental data has a very long but not well-known history. 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 to date. 41% of the unclassified Earth observation satellite programs of the 10 largest nations today provide open data.[15] Open access to this data has led to a rapid increase in the number of downloads shown in Fig. 4 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.[16] The urgent need for a transition to a more sustainable society, and clear signs of anthropogenic climate change, have created multiple new movements and organisations – 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 organisation within the Linux Foundation that enables companies worldwide to develop energy systems related OSS tools collaboratively.
- ClimateChange.ai convenes an open community and provides multiple educational resources around tackling climate change through artificial intelligence – another technological revolution that has been catalysed by open source culture.
- The Open Climate Community Calls, launched by the well known Appropedia commmunity, 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.

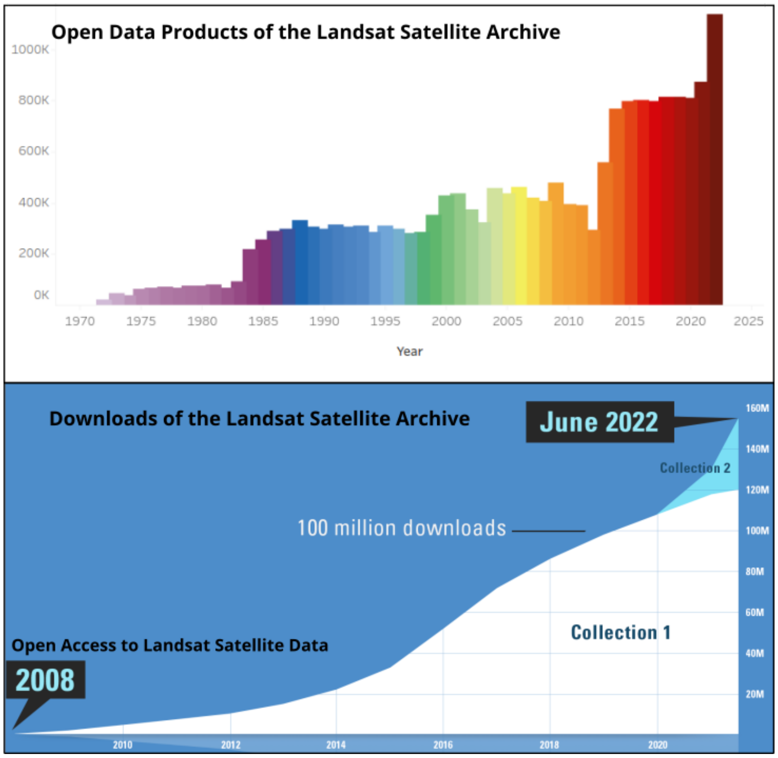


Fig. 4 - With the disclosure of the Landsat archives, the use of these satellite data increased noticeably Source: [Landsat Archive Dashboard]

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.

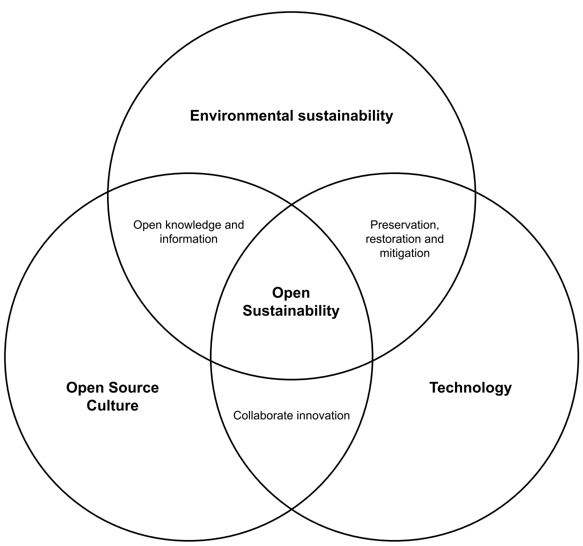


Fig. 5 - Dimensions of open sustainability.

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".[1]

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 contributes 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.[2] 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".[3] 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 minimise 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 analysed in detail within numerous books and publications.[4] A whole ecosystem has emerged with the objective to commercialise 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.[5]

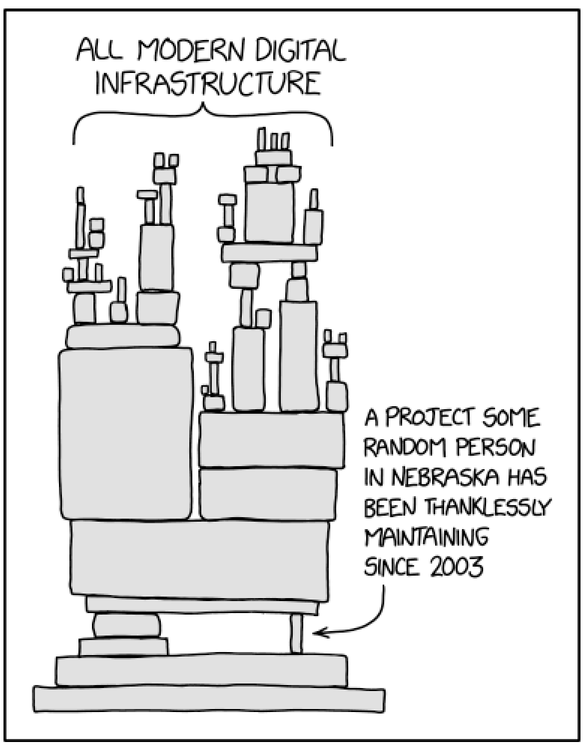


Fig. 6 - A meme illustrating the dependency structures within software shows how even small projects can be of central importance for an entire software ecosystem. Source: xkcd

Objectives

OSS is a public good, yet no one is responsible for stewarding and funding it. While OSS providers contribute immense value to society and the economy, the ecosystem's contributions are largely underappreciated by society at large – resulting in a systemic lack of investment, and ultimately, limiting the long-term sustainability and potential impact of many projects. While there is a general trend toward more widespread use and adoption of open source projects in sustainability, there is no systematic assessment of which projects can be considered critical digital infrastructure, or where significant funding and resourcing gaps exist.

For the past two years, Open Sustainable Technology has mapped the OSS ecosystem in sustainability and climate technology. An OSS ecosystem is defined as a collection of related software projects that co-evolve within the same environment. This study examines the health and state of this emerging ecosystem and evaluates the impact of open source culture on environmental sustainability. **The key objective is to create a vision and develop strategies to accelerate the use of open source in environmental sustainability.** To achieve this, the following objectives are pursued:

- Identify common principles based on shared values within the ecosystem.
- Provide evidence about dynamics and the relationships between projects, technologies, communities, and organisations in the open source ecosystem for environmental sustainability.
- Give developers, communities and users an overview of potential risks and opportunities associated with existing projects, highlighting positive trends and building awareness of critical but neglected areas. Furthermore, help identify projects suitable as building blocks for new initiatives.
- Identify untapped potential and make recommendations on how funding and resources can be effectively allocated to existing projects and ecosystems while fostering growth in emerging fields.
- Highlight opportunities for more systemic collaboration through shared digital infrastructure, standards and interfaces.

Methodology

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

- **General overview** - project themes, growth, age and popularity, and general project maturity.
- **Technology** - preferred programming languages, licence 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, ~1300 **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 organisation.
4. Be structured and documented in a way that allows maintenance, reuse and extendability.
5. Be published under an open source licence.

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 analyse 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.

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 which 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 individuals 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 modelling, 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?*

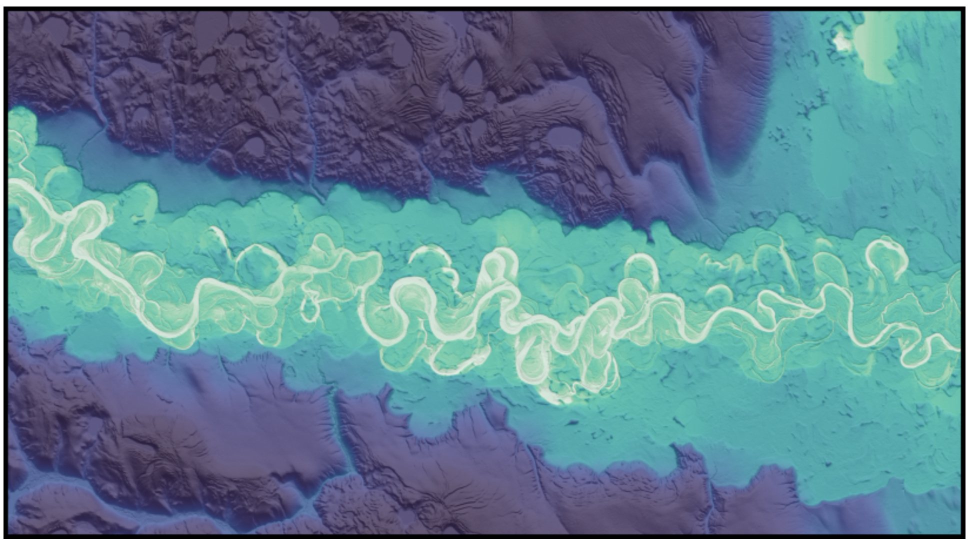


Fig. 7 - Beaver Creek, a tributary to the Yukon River, Alaska, USA. Visualisation created with the open source Python package RiverREM. License: GPL-3.0

Open Sustainability Principles

Openness is a key indicator of sustainability. It allows different disciplines, organisations 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. 8 - 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 scrutinised 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. How can solar cell recycling be done economically? Can carbon capture and storage be a scalable and viable mitigation strategy to decrease carbon dioxide in the atmosphere? Is land used for biofuel crops competing with food production, and what is its effect on biodiversity? Open datasets, models, and access can be critical for transparency and trust in the search for answers to these questions.

Open source approaches also allow for public scrutiny and independent auditing. The ever-increasing resolution of satellite data means that the environmental impact of state and non-state actors can be measured with high granularity. Open data and models create transparent metrics and more comprehensive sustainability impact assessments. For example, open datasets track the locations and companies involved in cement production, iron and steel production, and global coal, oil, and gas supply chains. The platform Carbon Mapper puts pressure on industry by disclosing methane and carbon dioxide sources. Through satellite imagery, impacts on the natural environment such as greenhouse gas emissions, land use change, biodiversity loss, soil drought, or water pollution can be measured with high temporal and spatial resolution. A variety of forest health indicators can be derived from this data, like deforestation risk, wildfire risk, biodiversity or tree detection. The use and improvement of these methods provide independent evidence that habitats are being lost and will ultimately help in the fight to protect these ecosystems.

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 (modelling, forecasting, impact assessment) will help reduce greenwashing and ensure that the environmental impacts of both state and non-state actors are verifiable and traceable.

Open source tools enable meaningful comparisons that guide decision-making. Earth system models, integrated-assessment models, and macro-energy models play a central role in forming plausible scenarios, informing science-based targets and guiding policy development. Satellites and remote sensing tools allow for the observation and validation of such models, helping to refine prior assumptions. Economy-wide tools such as life-cycle assessment and input-output analysis provide methods for accurately estimating the environmental impact associated with a region, sector, product, process, or service. Some systems require real-time optimisation with respect to the consumption of resources. At the sectoral level, platforms like Electricity Maps monitors and predicts the carbon intensity of electrical grids worldwide on an hourly basis. This information is used to shift energy consumption to times when there is a high share of renewable energy sources in the electricity grid, thereby reducing emissions. In the IT industry, developers and businesses use tools like Scaphandre, Cloud Carbon Footprint, and Kube Green to monitor energy consumption, estimate carbon emissions, and inform decisions about the optimal use of cloud infrastructure. Other tools are essential in the research and development of new technologies, and in evaluating current and future performance. Tools such as pvlib provide open libraries for simulating the performance of photovoltaic (PV) energy systems. This, in turn, helps to determine the future value of PV generation projects and shapes the financial and investment risk profile of PV projects.

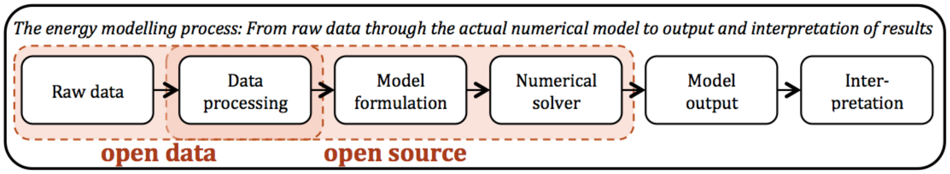


Fig. 9 - The the openmod manifesto illustrates how traceable decition making is possible in the decarbonization of the energy sector. The openness of data, models and dicussions enables traceable knowledge creation for science and society. License: CC 4.0

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] One example is the open online marketplace for local food, Open Food Network. It enables a network of independent online food stores by connecting farmers and food hubs with individuals and local businesses, thus giving them an easier and fairer way to distribute food.

Open source approaches also mean each 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 decentralised, loosely coordinated approach. For example, rather than developing yet another local bike-sharing system, the team behind Citybikes provides an open bike-sharing back-end to third-party developers and researchers. Citybikes is built on pybikes and currently supports over 400 cities. The Citybikes API is the most widely used dataset for building bike-sharing transportation projects.

OSS also standardised the practice of collaboration between software contributors. This distributes the benefits of cooperative software development for common standards and higher code quality. For example, the Mobility Data Specification is a standard that enables right-of-way regulation and two-way communication between mobility companies and local governments. Open interfaces and open architecture make it possible to divide the complexity into different modules, which greatly facilitates cooperation between different developments and organisations.

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 is driving a systematic shift away from classic proprietary business models — **the same mindset can be applied to open sustainable technologies.**

Localisation and Decentralisation

Wealthier economies with greater technological capacity can enable developing countries 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. For example, open agriculture can be adopted by farmers to decrease the environmental impact of their farming techniques. Practices like these enable faster global change that is responsive to local conditions and cultural needs.

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. For example, open source organisations such as NASA Harvest use measurements and forecasts to improve food security and agriculture worldwide. By open-sourcing the data, models, and software tools, the derived knowledge and insights are easily accessible, adaptable, and actionable within the local context. This is critical to building resilience, social equality, and strengthing local economies. The alternative, pursuing external proprietary options, may not be appropriate at the speed and scale required. OSS, on the other hand, can lay a digital foundation based on solutions that are more easily integrated, interoperable, and maintainable. The same applies to open source hardware solutions, which are increasingly important in accelerating decarbonisation efforts in the Global South.

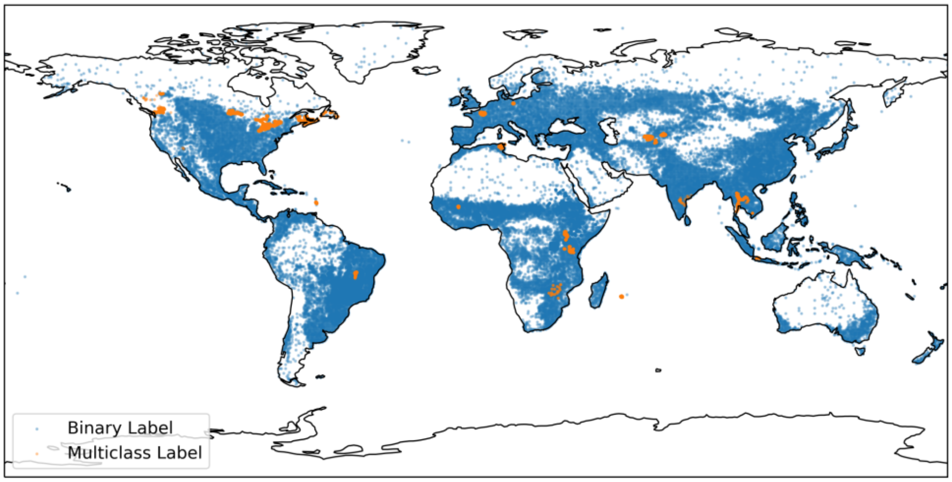


Fig. 10 - CropHarvest is an open source remote sensing dataset for agriculture with benchmarks. It collects data from a variety of agricultural land use datasets and remote sensing products. License: CC-BY-4.0

Overview

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	1339
GitHub projects	1187
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	126568
Total contributors in all projects	14628
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

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 colour represents the Development Distribution Score.

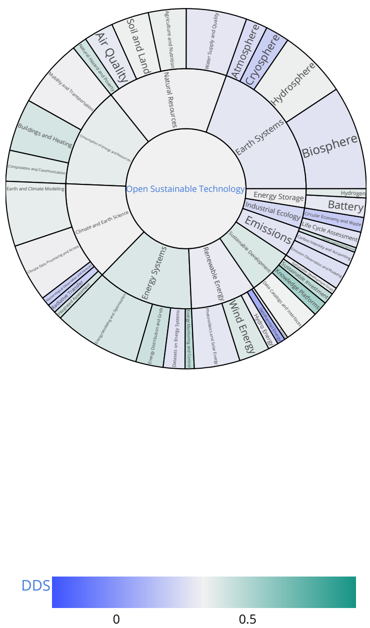


Fig. 11 - All studied projects grouped into the corresponding fields and topics

> **Tip**
>
> The above plot is fully interactive. Drill into fields, topics, and projects with a click! The project name links to the project's repository.

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. The colour bar shows the Development Distribution Score (DDS) as a measure of the distribution of work among the individual developers. A high value indicates a high distribution of work and, thus, a strong developer community. More details about this can be found in chapter [Development Distribution Score](#).

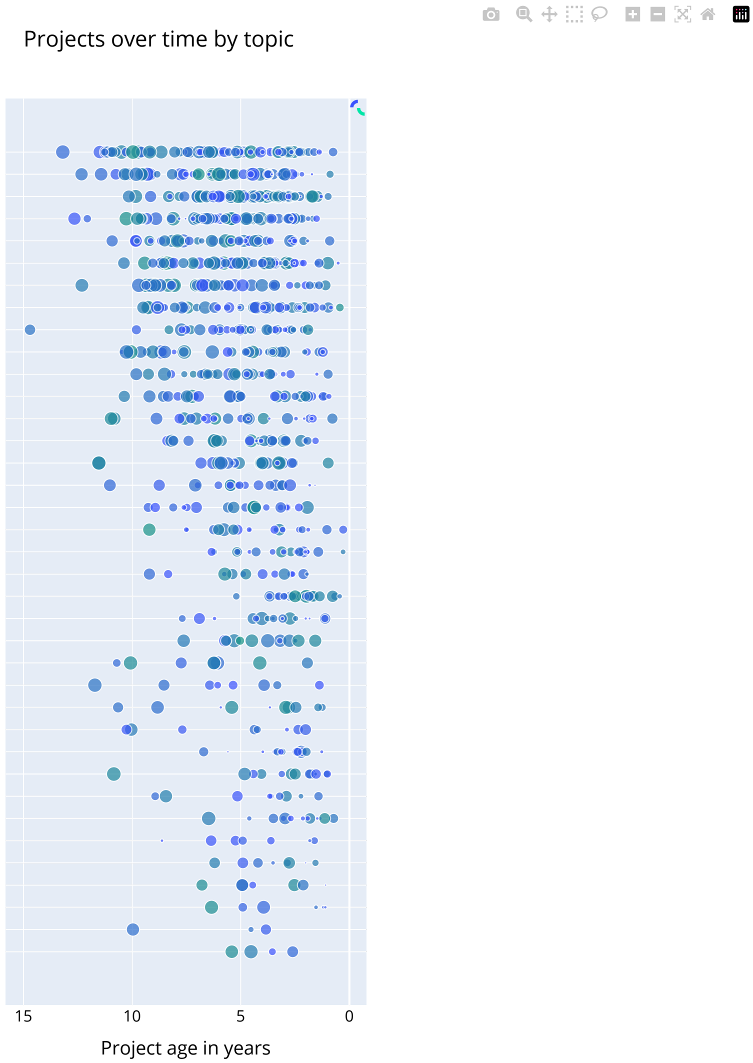


Fig. 12 Overview of all projects of the last 14 years since the launch of GitHub

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

Furthermore, areas where software plays a central role, but only a small number of projects can be identified, are of particular interest. **For example, within sustainable investment, representing only 1.15 % (a total of 11 projects), open source is still a marginal factor.** Despite ongoing discussions about the quality and transparency of ESG (Environmental, Social and Governance) ratings, the field is dominated by proprietary, closed-source frameworks and datasets. The lack of open source and open science in sustainable investment reflects the lack of transparent impact measurement and evaluation, which is key in financing a sustainable transformation.

In emission observation and modelling, only 22 developments have been identified, representing 2.1% of all projects. Despite the significant impact of anthropogenic emissions on the climate, there is a lack of open source tools, platforms, and communities that truly reflect the magnitude of the challenge. A significant business opportunity would exist for an open source community to bring together various emissions monitoring and modelling datasets from around the world on a single platform. A platform like this would be critical for increasing transparency around pressing issues like carbon trading, carbon taxes, and company sustainability assessments. There are new promising developments in this space, such as [The Global Registry of Fossil Fuels](). Moreover, Electricity Maps has successfully demonstrated how this approach works when applied to local energy grids. It has made it possible for hundreds of scientists and developers to collaborate in an open way and combine existing public data into a single digital platform.

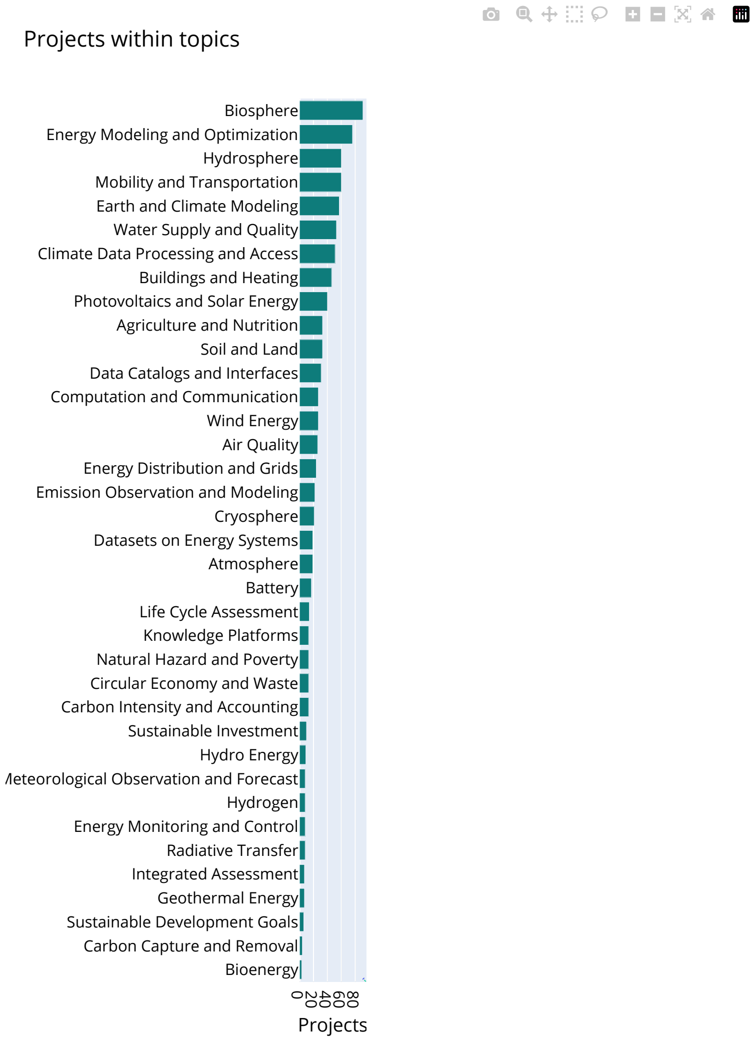


Fig. 13 Number of individual projects within topics

Topics with low OSS representation include bioenergy, hydrogen, and carbon capture. This is likely due to the more nascent nature of these areas and the relatively smaller academic communities working in them. These technologies have a higher degree of uncertainty, with intellectual property closely guarded by a few for-profit companies. 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 – transparently and independently – the state of, and potential contributions towards sustainable development. In the topic of industrial production (with respect to energy and resource consumption) only a single project was discovered even after intensive research. That's why the topic was removed from further investigations.

Lastly, topics like carbon offsets or climate neutrality disclosures could not be investigated in depth due to a general lack of OSS projects. Despite intensive research, no OSS project or organisation (with the exception of CarbonPlan) could be found that provides comprehensive and scientifically sound calculations and methodologies of climate neutrality and carbon offsets claims made by individual companies. All statements about the environmental impact of companies are primarily based on black box algorithms and analyses performed by companies and consultancies, making sustainability claims of Carbon Offsets rather opaque.

Fig. 14 - The goal [oco2peak] is to localise CO2 emissions on Earth based on the 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 modelling.

Popularity

Social software development and version control platforms like GitHub allow users to “star” other users’ repositories. Staring is often used to indicate support, bookmark repositories for later reference, and create curated lists. A repository’s total number of stars is an easy metric to keep track of and is routinely used by the community as an indicator of project popularity.

The popularity of a particular topic or field can be determined by summing the number of stars across all related projects. 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 software in environmental sustainability combined! With a median of 42 stars, the field of sustainable technology can be considered rather unpopular on GitHub.

With only 3 projects that have more than 1000 stars, it is once again becomes clear how little attention OSS receives within environmental sustainability. 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 critical, which explains why open source software is more widely recognised.

> **Tip**
>
> Click the project name and go directly to the repository.

Fig. 15 - 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.

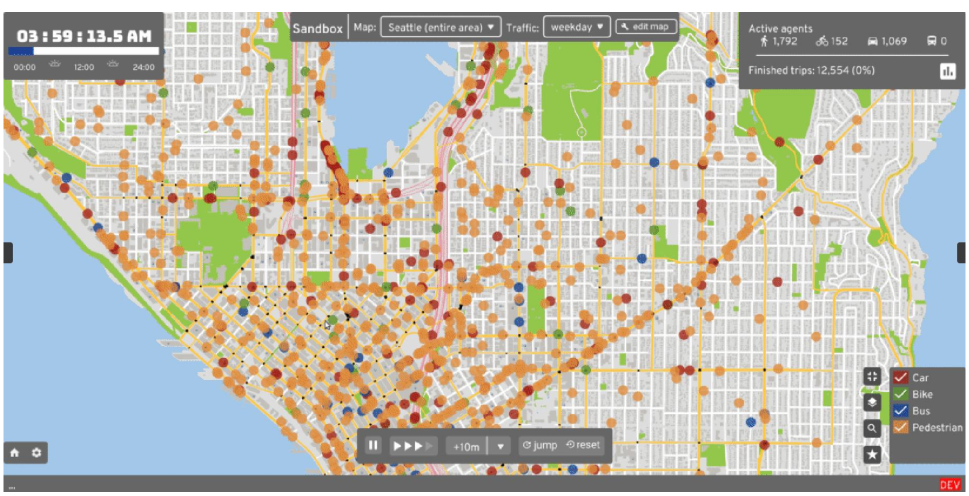


Fig. 16 - [A/B Street]() - Transportation planning and traffic simulation software for creating cities friendlier to walking, biking, and public transit. License: [Apache-2.0]()

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.

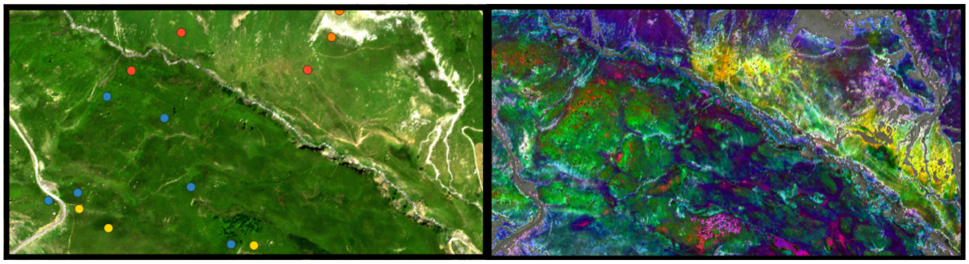


Fig. 17 - biodivMapR - An R package for α- and β-diversity mapping using remotely-sensed images. The left image shows the visible spectra with multiple validation sites. the right image shows the derived biodiversity.

Age

A repository's age is 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 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. The decline of more junior projects within the dataset can be explained in several ways:

- Repositories meet the criteria for inclusion in this study after a certain period of time. This applies in particular to the documentation and the use of the project by third parties.
- It is common practice to make a project public and advertise it after sufficient development time has elapsed.
- The age considered is the project age and not the repository age. GitHub has become increasingly popular in the last 10 years. However, many Git repositories are older and have moved to the platform during this period.
- There are fewer new developments within the subject area.

However, 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. 18 - Distribution of project age

Some of the oldest projects that are still active such as Pysolar, iNaturalist or OCE with an age of more than 10 years originate in academic communities, when the share of other organisations 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 person for over 10 years.

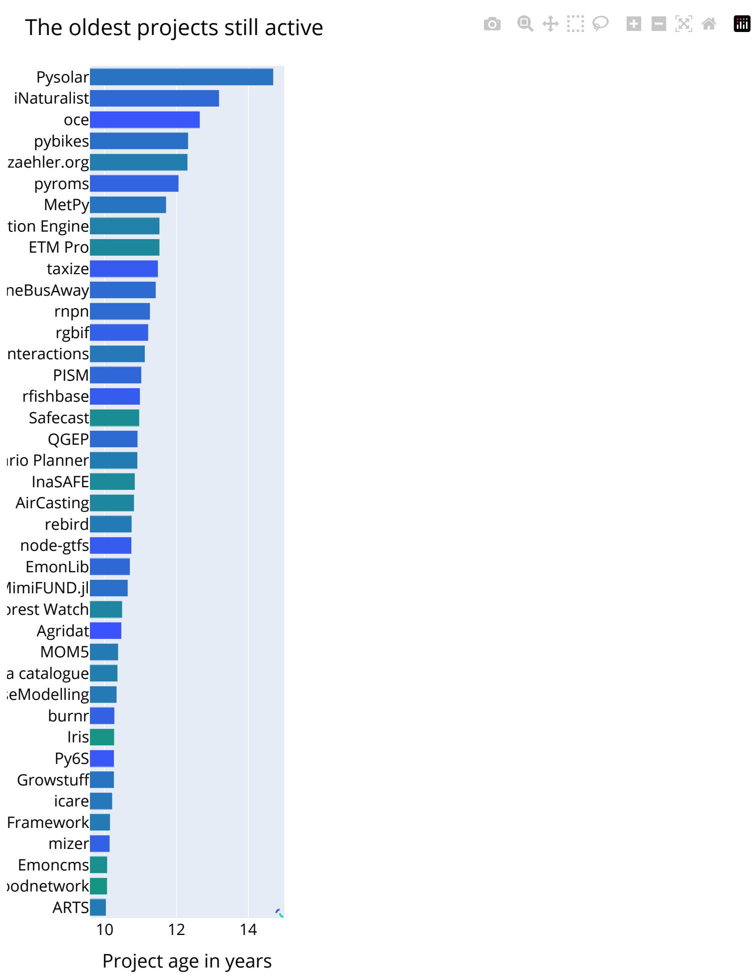


Fig. 19 - The oldest projects still active

Growth

Growth plays an important role in the development of projects and the greater ecosystem. We can detect some emerging topics by examining the growth rates of stars over the last year. The topic of green software is growing in popularity, also known as “green computing”. Its concern is the lowering of energy consumption, carbon intensity, and environmental impact of programming languages and applications. With over 97% of the world’s applications now using open source, this clearly affects energy usage and efficiency worldwide.[1]

Projects like Kepler, kube-green, Scaphandre, Cloud Carbon Footprint, Software Carbon Intensity Specification, and CodeCarbon demonstrating the growth within this new topic. This trend is partially driven by new collaborative communities, such as the Green Software Foundation, which is supported by major software companies. The ambitious climate targets of companies like Google and Microsoft may be a contributing factor to this trend. Levels of access and participation may also play a role, as green software typically has a lower barrier to entry for general software developers than other topics.

Tip

Click on the project names to go directly to the repositories.

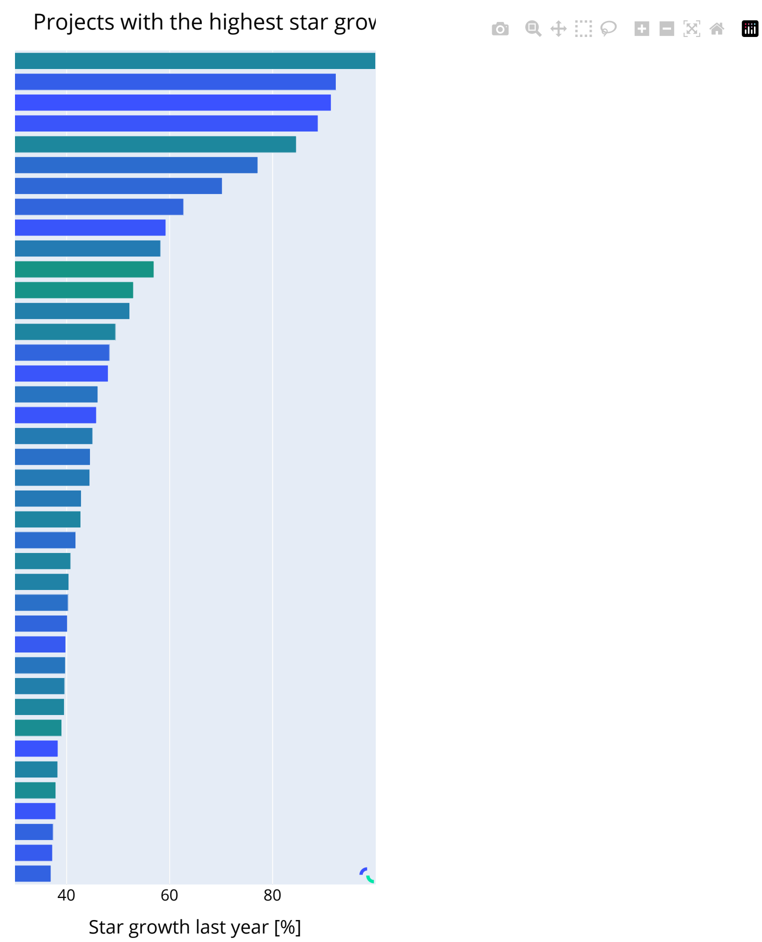


Fig. 20 - The 40 projects with highest star growth

Another way to look at project growth is in relation to total commits over the past year. The number of commits created within a given timeframe highly depends on the project type. For example, a web platform with a large codebase tends to go hand-in-hand with a high volume of commits compared with projects centred on data analysis. Meanwhile, software tools that are dependencies of many other projects are more likely to see a high frequency of commits and sustained growth, as there is an incentive to maintain the codebase to avoid bugs and breaking changes. However, for this reason, high growth can also be considered a risk. Within this ecosystem, both relatively young and large, widely known projects can be seen with high commit growth. These include xclim, a library of derived climate variables; InVEST, an Natural Capital Project model used to “map and value the goods and services from nature that sustain and fulfil human life”; and Energy Sparks, which is designed to help schools improve their energy efficiency. However, as with other growth metrics, this metric can be misleading for a number of reasons and may not reflect the quality of growth.

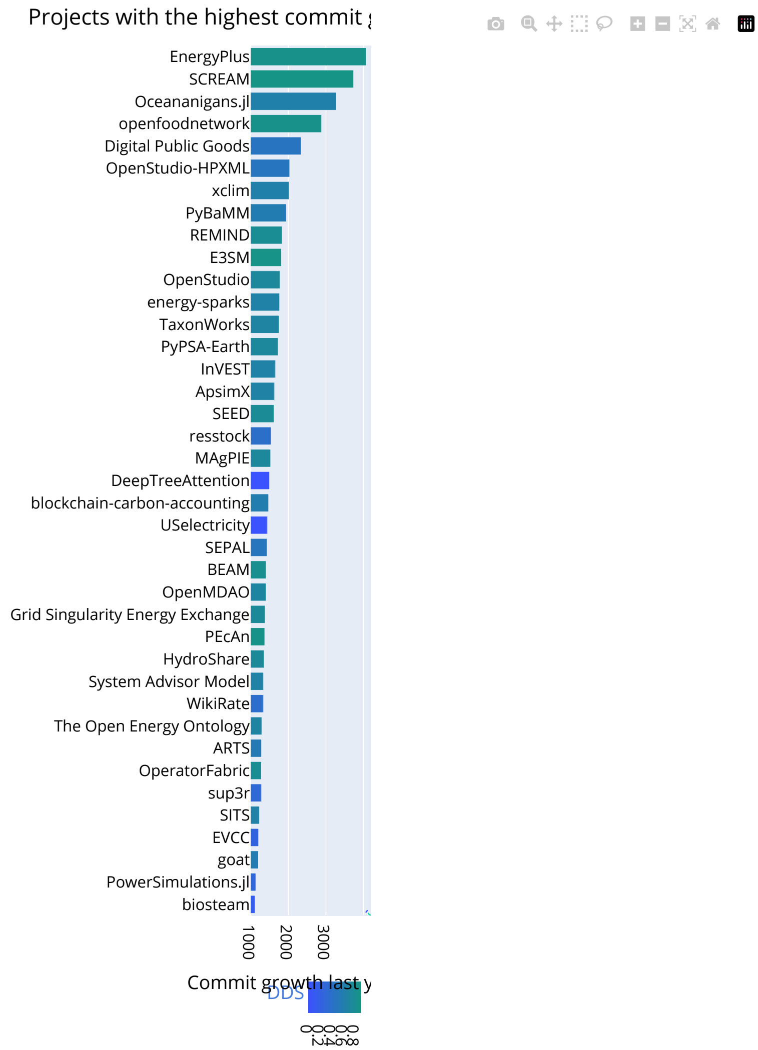


Fig. 21 - The 40 projects with highest commit growth

Ranking

Ranking all projects by a total score can provide a much deeper understanding of the ecosystem. While quantifying the state and health of a project remains challenging, using a multidimensional index creates a more comprehensive picture. A repository's total score is a composite index of three dimensions: **size**, **community** and **activity**. Each dimension contains several indicators, represented by an index used to rank projects relative to each other. Each dimension index is referred to as a score.

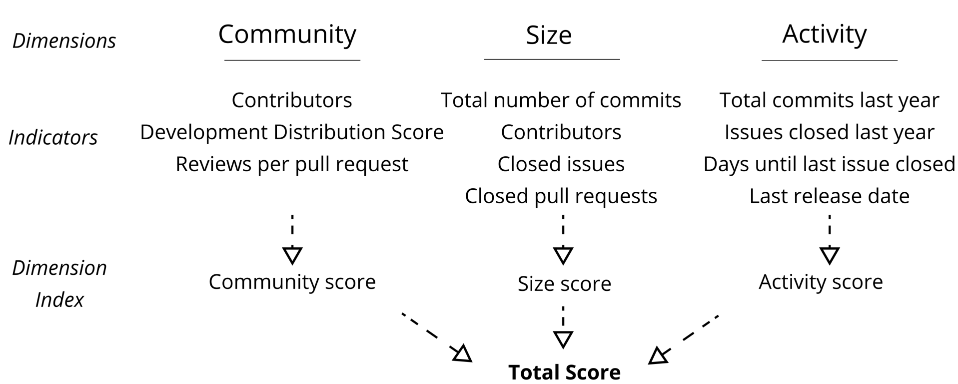


Fig. 22 - The relationship between the dimensions and indicators within the ranking.

For example, the activity score ranks each project using an index composed of *Total Commits Last Year*, *Issues Closed Last Year*, *Day Until Last Issue Closed*, and *Last Release Data*, which is normalised by 1. The weighted sum of all scores (activity, community, and size) is referred to as the “total score”. The figure above shows the relationship between the dimensions and indicators within the ranking (for implementation detail, see this code cell.)

Unlike stars, which can provide insight into a project’s overall popularity, ranking by total score unveils unpopular but otherwise strong projects. For example, larger projects like EnergyPlus suddenly rises to the top. However, as with any index, there are limitations. In this case, monolithic software developments have a higher probability of achieving a high score, meaning that projects which rely more on a modular approach (i.e., projects distributed across multiple repositories) may be significantly underrepresented.

The ranking of the projects according to their activity rating highlights the young projects that are still developing rapidly. DeepTreeAttention in particular stands out here, which is mainly developed by a single person. Other young projects such as Ozon3, cmip6-downscaling or PowerSimulations.jl show not only a very high activity but the DDS also shows a strong growth of the community of this young projects.

The real value of such health analytics comes into play when development and community data is combined with usage data. Unfortunately, this data is currently only available to a limited extent via Python dependencies. Further work is required to extend usage metrics to include other software package managers and survey methods.

> **Tip**
>
> Click the project name and go directly to the repository.

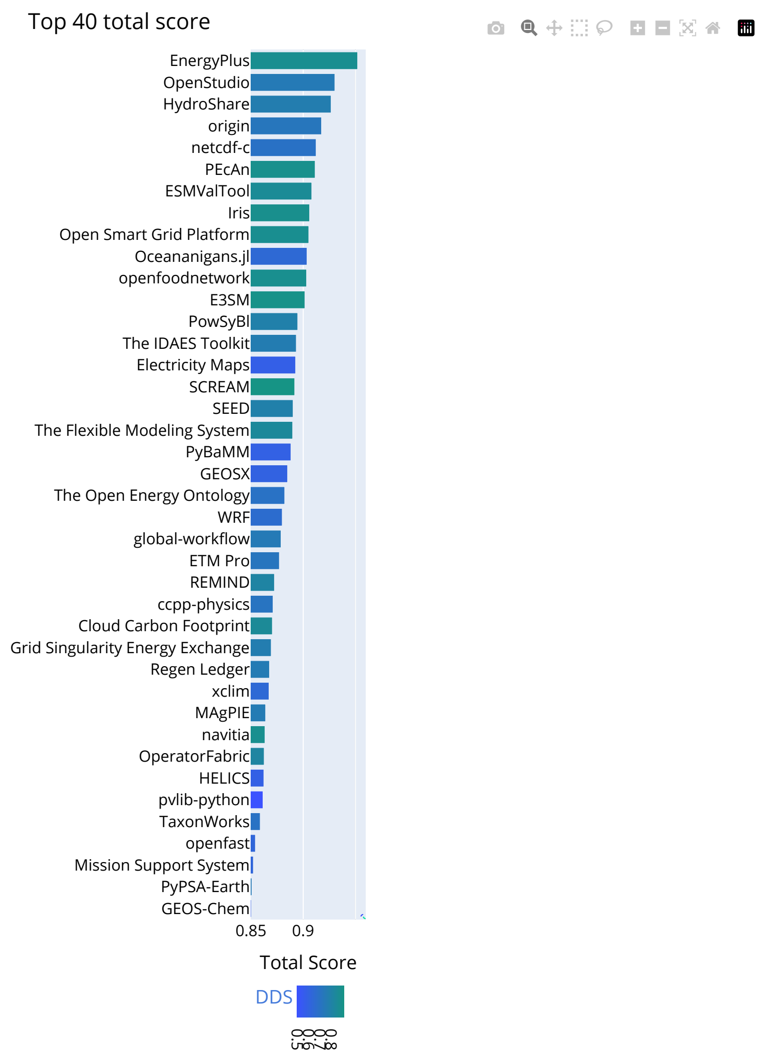


Fig. 23 - The 40 Projects with the highest total score

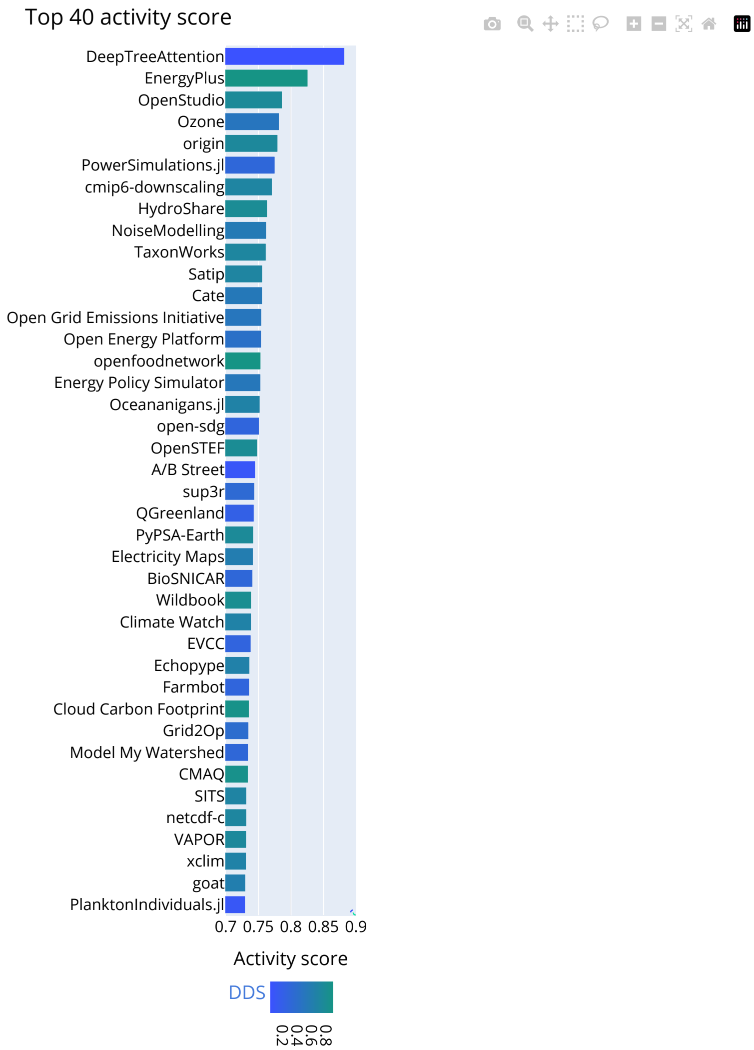


Fig. 24 - The 40 Projects with the highest activity score

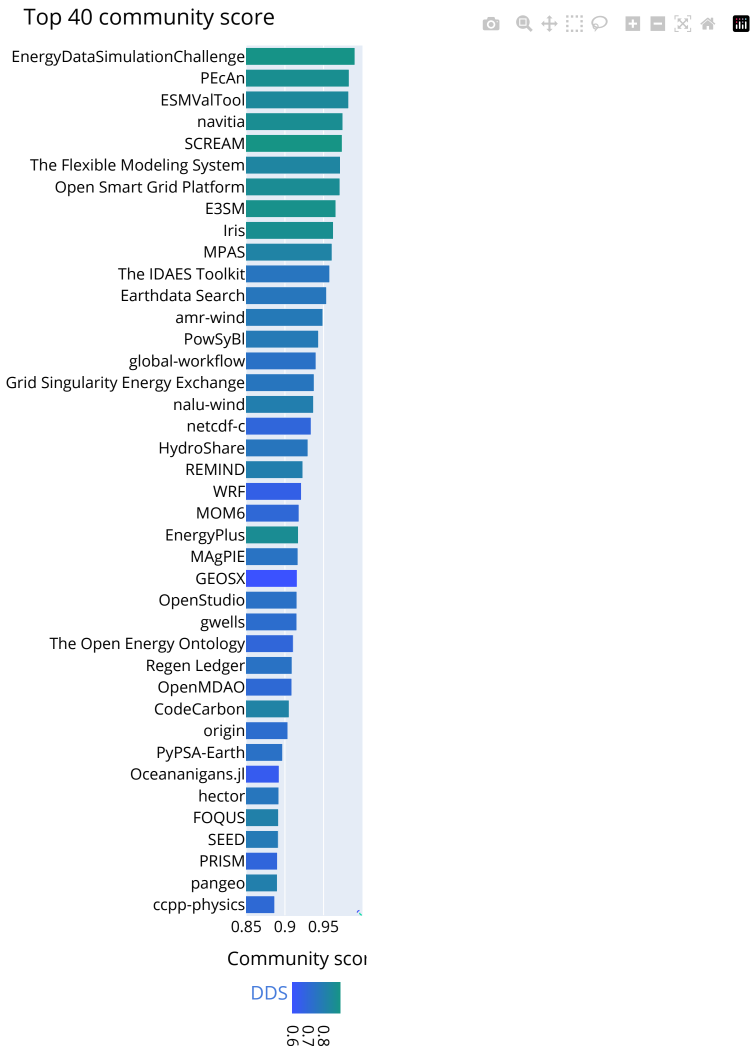


Fig. 25 - The 40 Projects with the highest community score

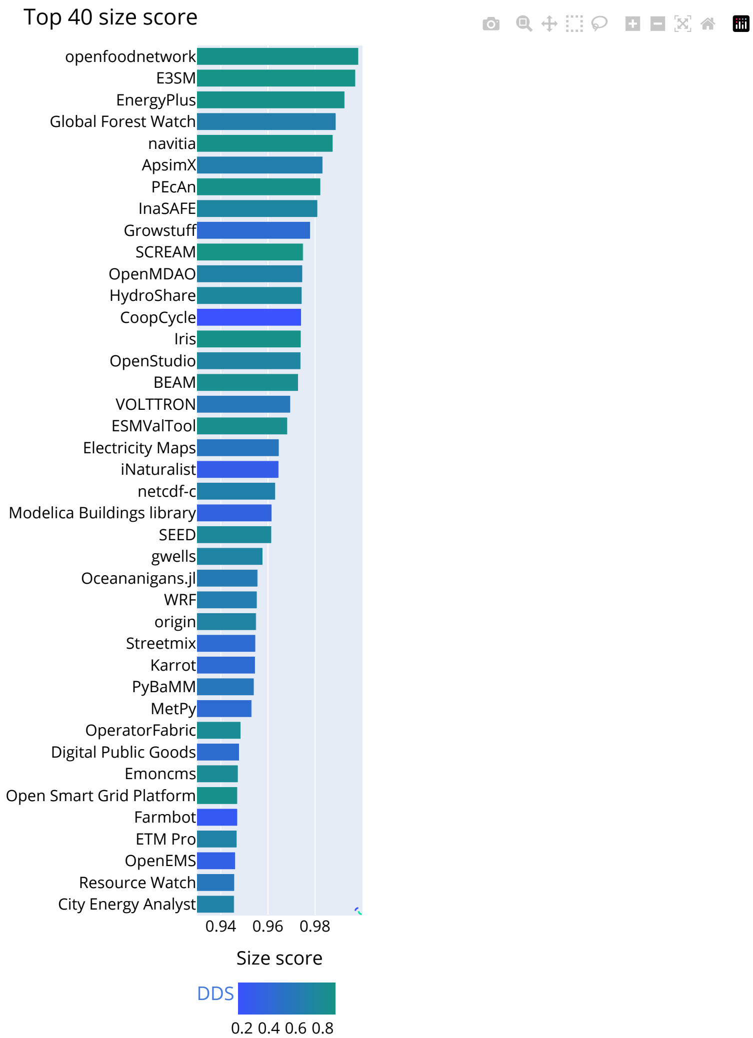


Fig. 26 - The 40 Projects with the highest size score

Dimensions and Calculations

```
# Each project is ranked according to different indicators in the dimensions of community, activity and size.
# A value of 1 represents the highest rank and 0 the lowest.
# The individual values are summed up within the dimensions to create the scores for the different
dimensions.
df_active["activity"] = (
    df_active["total_commits_last_year"].rank(pct=True)
    + df_active["issues_closed_last_year"].rank(pct=True)
    + df_active["days_until_last_issue_closed"].rank(pct=True)
    + df_active["last_released_date"].rank(pct=True, na_option="top")
) / 4

df_active["community"] = (
    df_active["contributors"].rank(pct=True)
    + df_active["development_distribution_score"].rank(pct=True)
    + df_active["reviews_per_pr"].rank(pct=True)
) / 3

df_active["size"] = (
    df_active["total_number_of_commits"].rank(pct=True)
    + df_active["contributors"].rank(pct=True)
    + df_active["closed_issues"].rank(pct=True)
    + df_active["closed_pullrequests"].rank(pct=True)
) / 4

# The scores are summed up and normalised so that 1 represents the largest total score.
df_active["total_score"] = (
    df_active["activity"] / df_active["activity"].max()
    + df_active["community"] / df_active["community"].max()
    + df_active["size"] / df_active["size"].max()
) / 3
```

Programming Languages

The number and kinds of programming languages 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 the project’s profile within the context of their own experience and organisations. It can also help inform students on which programming languages they might focus their efforts on learning, depending on their topic of interest.

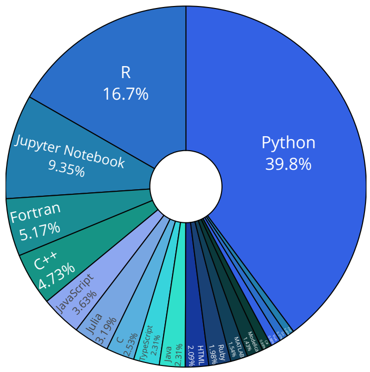


Fig. 27 - Distribution of programming languages

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 or the official numbers of GitHub provide insights into the programming language usage of Open Source projects. Compared to the broader open source ecosystem, it is clear that Python has a significantly higher use within the repositories analysed, compared to widely popular languages, such as JavaScript. This indicates a strong focus on analysing large datasets, where Python and Jupyter Notebooks are increasingly dominant, with less focus on the web application side. Python, in particular, is the language of choice in projects in energy modelling, biosphere, hydrosphere, wind energy, buildings and heating. Python is considered to be an energy-inefficient programming language. However, in practice, computationally intensive operations are typically offloaded to energy-efficient processes (e.g., via a C-API) using libraries such as NumPy.

The use of R deviates significantly from other statistics and has a high prevalence within the software world. A concentration of R developments 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 statistical-related projects within these topics, and the low number of general web development projects within the field of Sustainable Development. 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. For some special use cases, such as building simulation, programming languages like Modelica are frequently used.

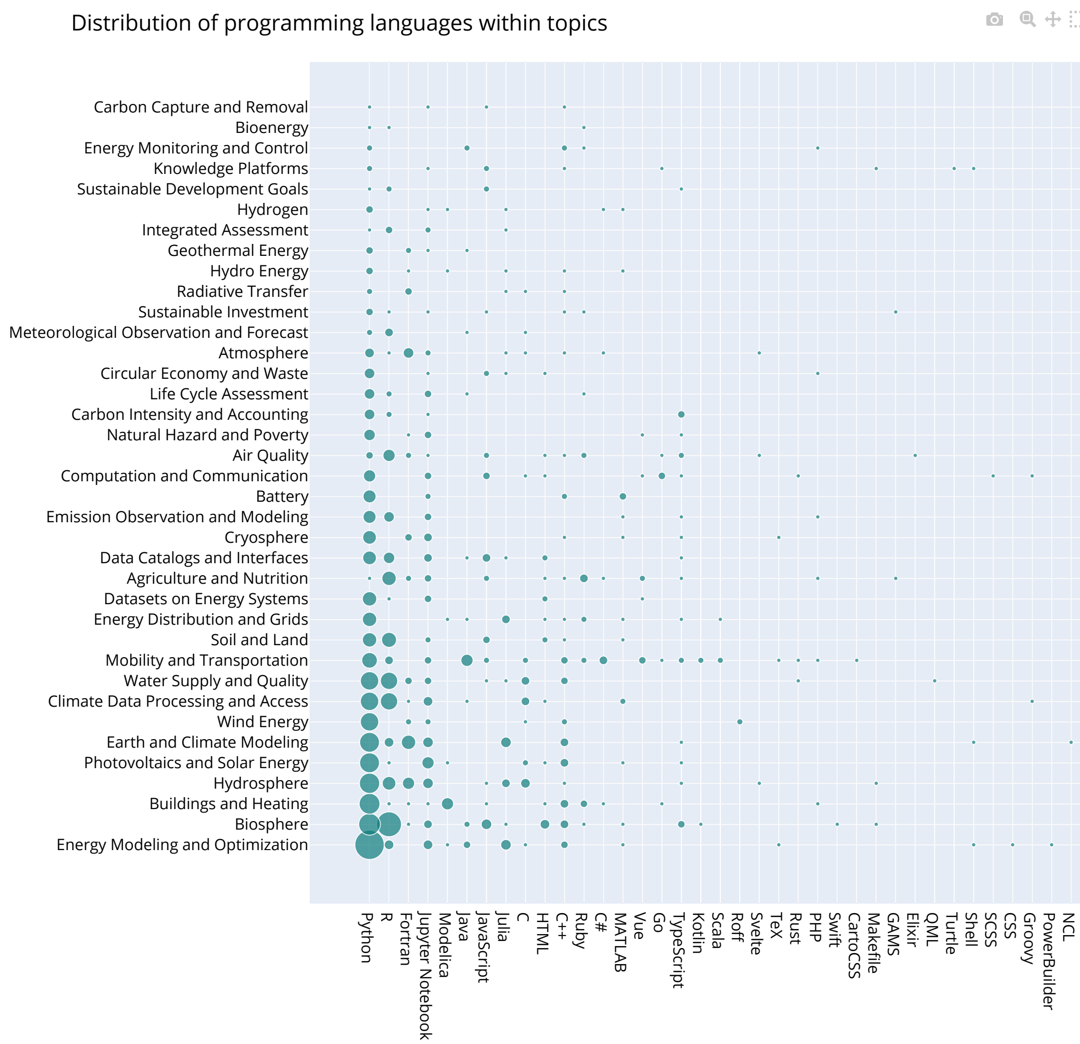


Fig. 28 - Distribution of programming languages within topics

Licences

Licences ensure that communities behind many of today’s most innovative technologies can protect their creations in whatever way they see fit. The open source licence determines how openness is defined from a legal perspective. Depending on whether one wants to license a small software tool, a larger framework, datasets or an internet service, the choice of licence can determine the long-term strategic direction. In this way restrictions, conditions and guarantees are defined by the authors under which use and derivations are possible. Especially for commercial use, licensing criteria play an important role. Websites like ChooseALicense can help you choose the right licence. The distribution of licences reveals the orientation of the open source ecosystem towards more permissive licences or copyleft licences.

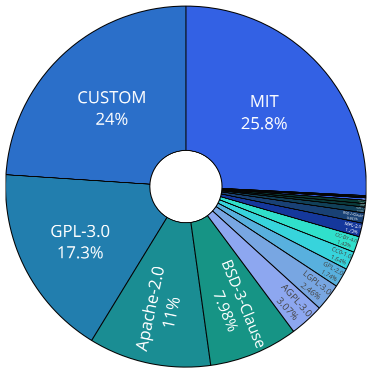


Fig. 29 - Distribution of licences

Permissive licences like BSD, Apache and MIT are the most popular in the field of sustainability. The MIT licence was the top choice, used in ~26% of the projects. MIT is a short and simple permissive licence. Permissive licences are easier to use in commercial products, and unlike the copyleft licence, they do not enforce the release of all code under a similar licence. Permissive licences like MIT create many opportunities for the reuse and commercialisation of OSS projects within proprietary software. Projects under this licence, on the other hand, may be jeopardised by the use of proprietary products. 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 licence was a Custom licence, used in 24% of the projects. All projects that could not be mapped to standard open source licences based on the SPDX License List are considered custom licences. Licences with modifications from the original open source licences also fall under this category.

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

Community

Establishing communities of developers and users brings open source projects closer to long-term viability. Developers and project leaders should be deliberate about attracting the type of community that is best suited to support the OSS project. To understand how organisations and people contribute to the community, we focus on two community types:

- **The community of contributors**. Software developers, architects, and designers who write code, create and update documentation, and contribute in a multitude of ways.
- **The community of users**. End users who find value in an OSS product, provide feedback and expertise, and can act as evangelists for the project.

The analysis of the communities is divided into several chapters:

Development Distribution Score

The distribution of responsibility and workload among contributors is essential to a stable community. A common way to estimate the risk resulting from information and capabilities not being shared among team members is the so-called Bus Factor. *The "bus factor" is the minimum number of team members that have to suddenly disappear from a project before the project stalls due to lack of knowledgeable or competent personnel.*

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 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 organisation undergoes significant social or technological changes (for example, personnel leave a project or can no longer contribute), others have 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 organisation or developer, which affect not only other developers or users, but also the dependencies to other projects.

The commits of the strongest contributor are measured in relation to the total number of commits. Although commits are not an absolute measure of an individual's performance within a project, they do reflect working relationships after a certain period of development. Furthermore, it makes it possible to assess the status of a project without having to make direct comparisons to other projects. The DDS value is calculated using the following formula:

$$DDS = 1 - \frac{n_{MaxCommitsSingleContributor}}{n_{totalCommits}}$$

For instance, a DDS of 0.1 means that 90% of the transfers come from a single developer. Without the high engagement of that individual, it will become challenging for the rest of the community to maintain and further develop the existing code base. The following table shows the statistical median of the DDS on the whole dataset.

Group	Median DDS
All projects	0.304
Active projects in personal namespace	0.136
Active organisation projects	0.404
Active projects	0.335

Fig. 30 - Median Development Distribution Score within various groups of projects

Across all active and inactive projects, the median DDS is at 0.304. This means that most open source projects depend heavily on a single developer contributing 70% of the commits to a project. For inactive projects, this value drops to 0.136 while active projects have a DDS of 0.335. The highest values are shown for projects within GitHub organisations, with a DDS of 0.405. The top 50 projects (ranked by stars) have a median DDS of 0.415 and the development communities with the most contributors have the highest DDS value of 0.688 **demonstrating that workload is more evenly distributed among individuals in a large development community**.

In particular, the difference between inactive and active projects makes it clear that the DDS is an important indicator for the longevity of an open source project. However, a high DDS is not advantageous in every case. Brooks' Law is an observation about software project management according to which "adding manpower to a late software project makes it later". Especially for projects of high complexity, very large team sizes quickly lead to overhead and communication problems. Under these conditions, the distribution of work between many can become problematic. One solution to this is to split software projects into modular components that can be managed by smaller groups. This approach is known as the Unix philosophy in software development.

The following scatter diagram shows the distribution of DDS within different topics. Each circle represents a project and the size of the circles is scaled relative to the size score.

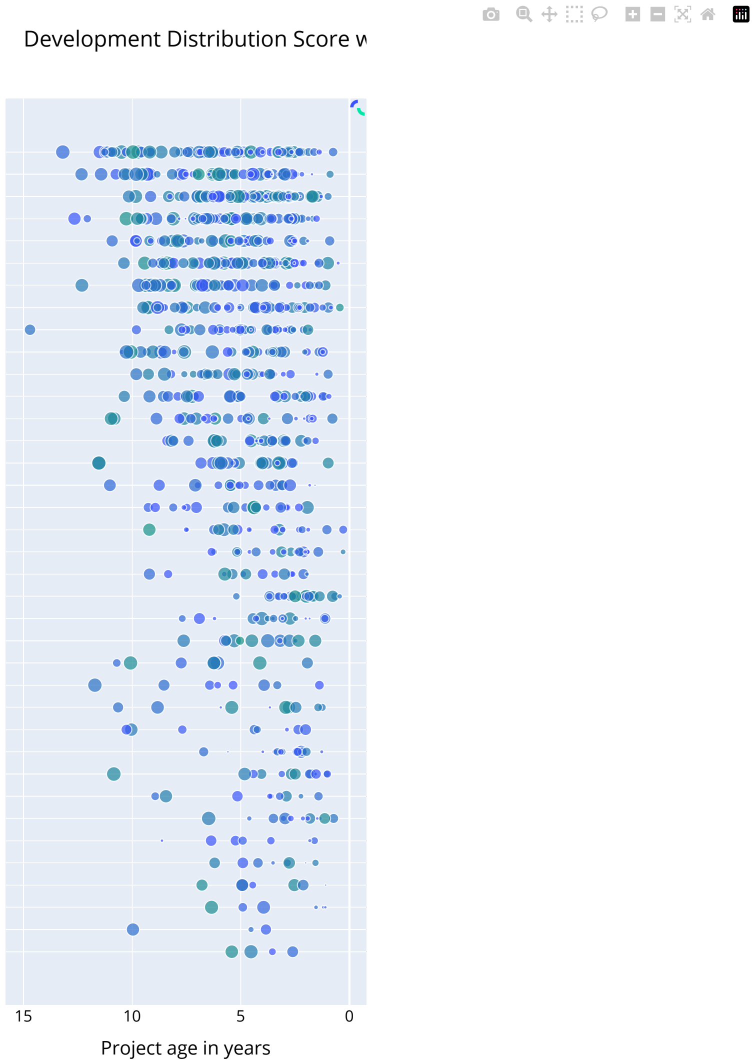


Fig. 31 - Development Distribution Score within topics

Contributors

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

Building a strong community of developers and users is the key factor in ensuring the longevity of open source projects. Depending on the project scope and size, it is necessary to continue to attract and integrate active developers in OSS organisations. The project team must create both the culture and the technical infrastructure to effectively participate in the larger community. Given the right conditions, newcomers can progress from their initial contributions to becoming core developers or maintainers. Below is a list of the top 40 projects with the most contributors.

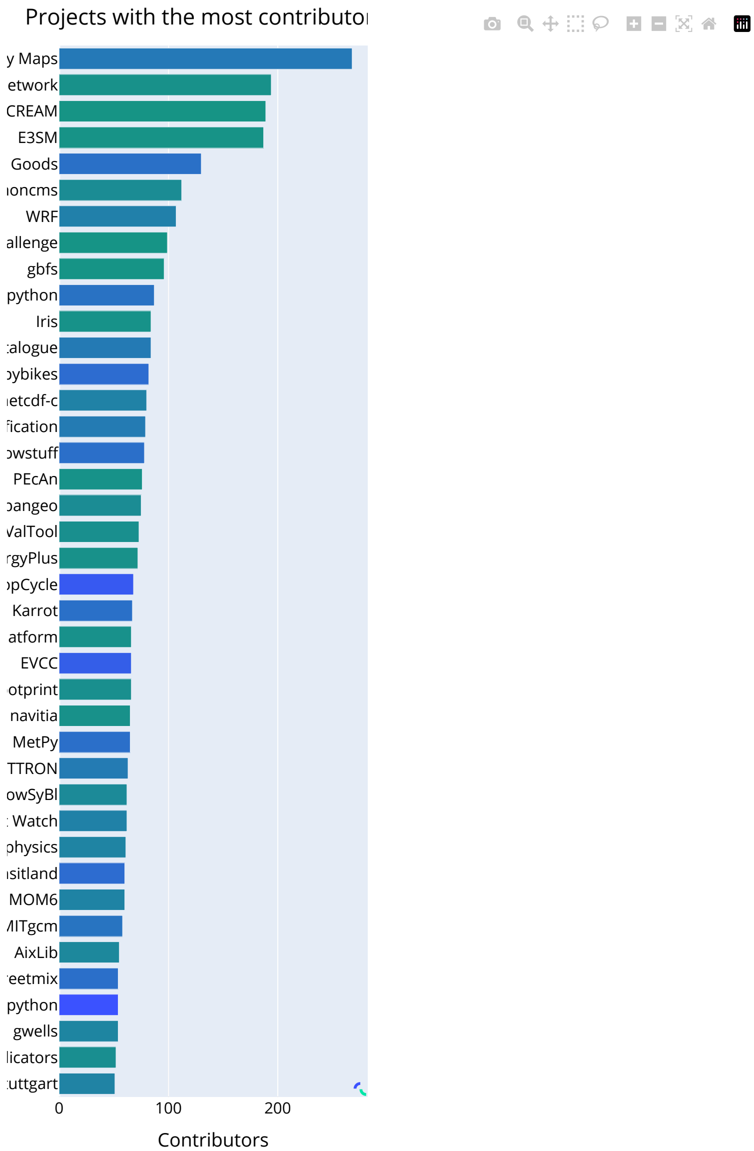


Fig. 32 - Projects with the most contributors

The high DDS among the projects with the most contributors shows the importance of sharing work among a variety of people 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 mobilise most developers around the globe to integrate local grid data and renewable energy share on a single platform.

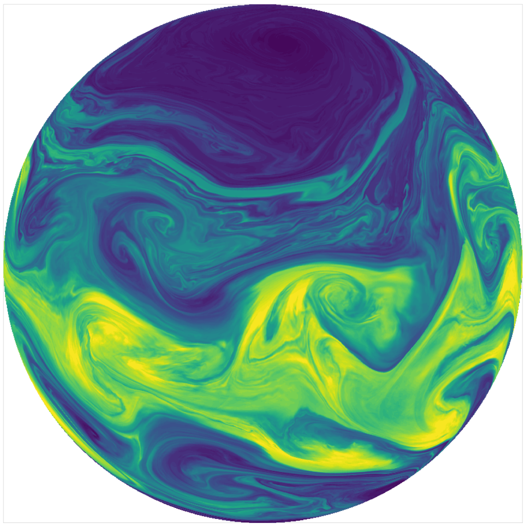


Fig. 33 - Turbulent eddies in water vapour from a 3 km horizontal resolution SCREAM simulation of baroclinic instability using simple physics. License: BSD

Contribution Guidance

Only 30.6% of the repositories have a contribution guide, thus making it clear how new contributors can participate in the projects. There could be numerous reasons for this. Either developers are unaware of the critical importance of such a guide for the overall project, or integrating contributions from external developers into the project is not prioritised.

Organisational Diversity

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

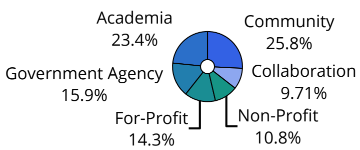


Fig. 34 - The distribution of organisational forms

- **Community-driven**: projects which are without institutional affiliation and are led by individual contributors. The majority of OSS projects within environmental sustainability can be found in this category, with a total share of **25.8%**. These projects are also the oldest we have identified. The organisations leading these projects are characterised 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**: projects which are hosted, managed, or developed by academic labs and research institutions; **23.4%** of all projects are hosted by university departments. 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**: projects which are hosted, led or developed by national and/or subnational governments; governments are responsible for **15.9%** of projects. 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**: projects which are initiated by private sector entities; industry, startups and other private organisations are responsible for 14.3% of projects. 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 harbour 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 for for-profit organisations within open source environmental sustainability like Electricity Maps, Breakthrough Energy, Vizzuality or Ladybug Tools.

- **Non-profit**: projects which are initiated by organisations that do not primarily aim to generate profits for the shareholders but rather pursue charitable goals. For open source development, such an organisation can provide the structure for long-term project sustainability if sufficient resources are provided. Several community-oriented organisations 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. This organisational form is strongly underrepresented, with a total share of 10.8%. Strong non-profits that take precedence here, are rOpenSci, Reiner Lemoine Institute, OpenClimateFix or Drawdown.

- **Collaborations**: initiatives where projects are hosted, led or developed by a consortium of different actors and institutions. This form of organisation is particularly suitable for generating knowledge transfer between different partners and accounts for the smallest share of 9.71%. The diversity of different organisations 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.

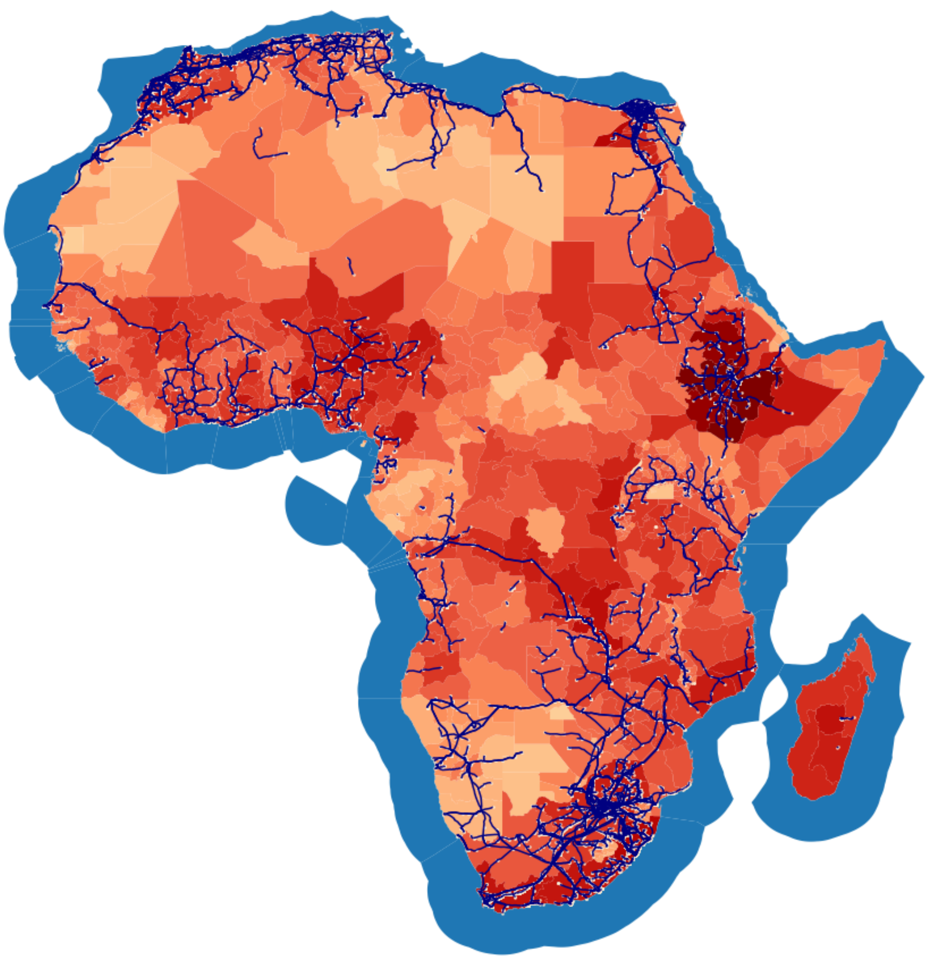


Fig. 35 - PyPSA-Earth is the first open source global energy system model with data in high spatial and temporal resolution. It enables large-scale collaboration by providing a tool that can model the world energy system or any subset of it. License: GPL-3.0

Most Listed Projects

Ranking organisations by the number of active projects in the GitHub namespace shows another important indicator of where strong developments are evident. National Renewable Energy Laboratory and rOpenSci are standouts here. Again, the vast contributions from the US Enviromental Protection Agency is clear. It is worth noting, however, that organisations that adopt frameworks that rely on strong modularity, such as Open Energy Modelling Framework, Wind-Plant Integrated System Design and Engineering Model or the Climate Modeling Alliance may be overrepresented compared to projects using mono repositories (a version-controlled code repository that holds many projects, also known as a monorepo).

Namespace	c
rOpenSci	28
National Renewable Energy Laboratory	24
U.S. Environmental Protection Agency	10
Province of British Columbia	8
oemof community	7
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
NREL-SIIP	4
Joint Global Change Research Institute	4
Open Climate Fix	4

Fig. 36 - Organisations with the most listed projects

Geography of Organisations

Analysis of the geographic distribution of organisations behind OSS in sustainability shows an overwhelming majority (64%) is placed in Europe and North America. 28.7% of the projects are considered global, as no geographical affiliation could be identified. Economic complexity – whereby the more diverse knowledge accumulated in a population, the greater the productive capabilities – may partly explain the distribution of open source communities between continents.

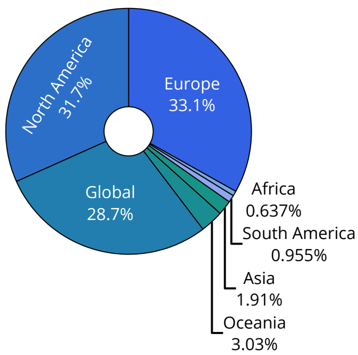


Fig. 37 - Distributions of organisations between continents

However, if one compares the ratios with open source developer statistics, clear differences in origin become apparent. Here, baseline data from "[The State of the Octoverse]" is used, a study that provides the geographic distribution of millions of active GitHub users.

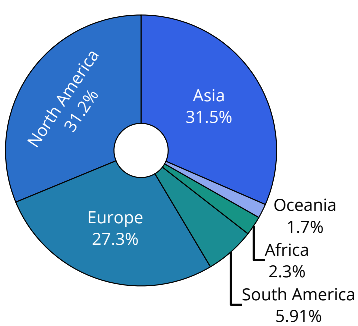


Fig. 38 - Distributions of all GitHub users between continents

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 organisations working in OSS for sustainability. Moreover, the absence of Indian communities is notable, with no large organisation or project identified, despite a high number of open source developers present. Similarly, very few organisations 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 organisations and global projects with no geographical affiliation, this deviation requires further investigation.

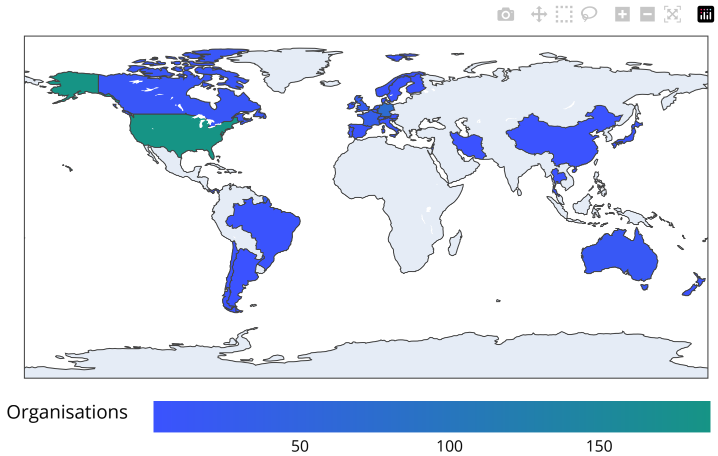


Fig. 39 - Global distribution of organisations

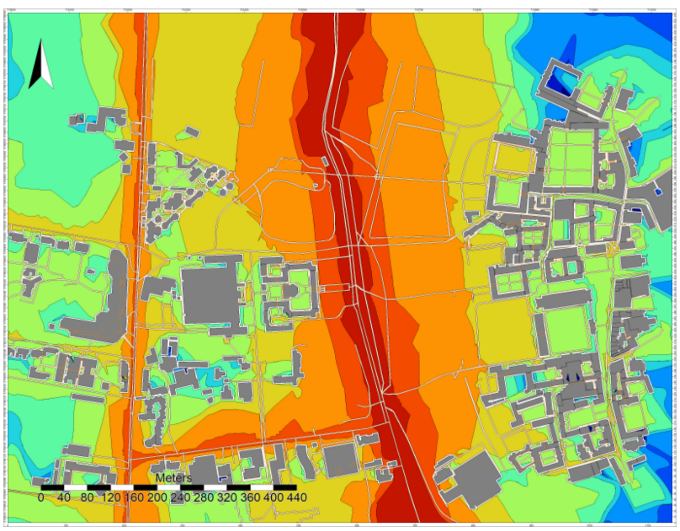


Fig. 40 - NoiseModelling is a free, open-source tool designed to produce environmental noise maps for very large urban areas. GPL-3.0 license

Users and Usage

The demand for OSS is everywhere, though its users and usage is not always apparent. Recent reports estimate that 95-97%[1,2] of mainstream IT organisations 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 realised. 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. Users act as evangelists by sharing OSS with other users or organisations that can benefit from it, and in return provide valuable feedback and expertise.

Project usage data from public software development and version control platforms remains scarce. GitHub, unfortunately, offers little support in compiling accurate statistics. Additionally, statistics on package manager downloads are not universally available and must be obtained through the various platforms and their APIs. While this is technically possible, it was not feasible given the study’s limited resources and timeframe. 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 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 modelling, 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.

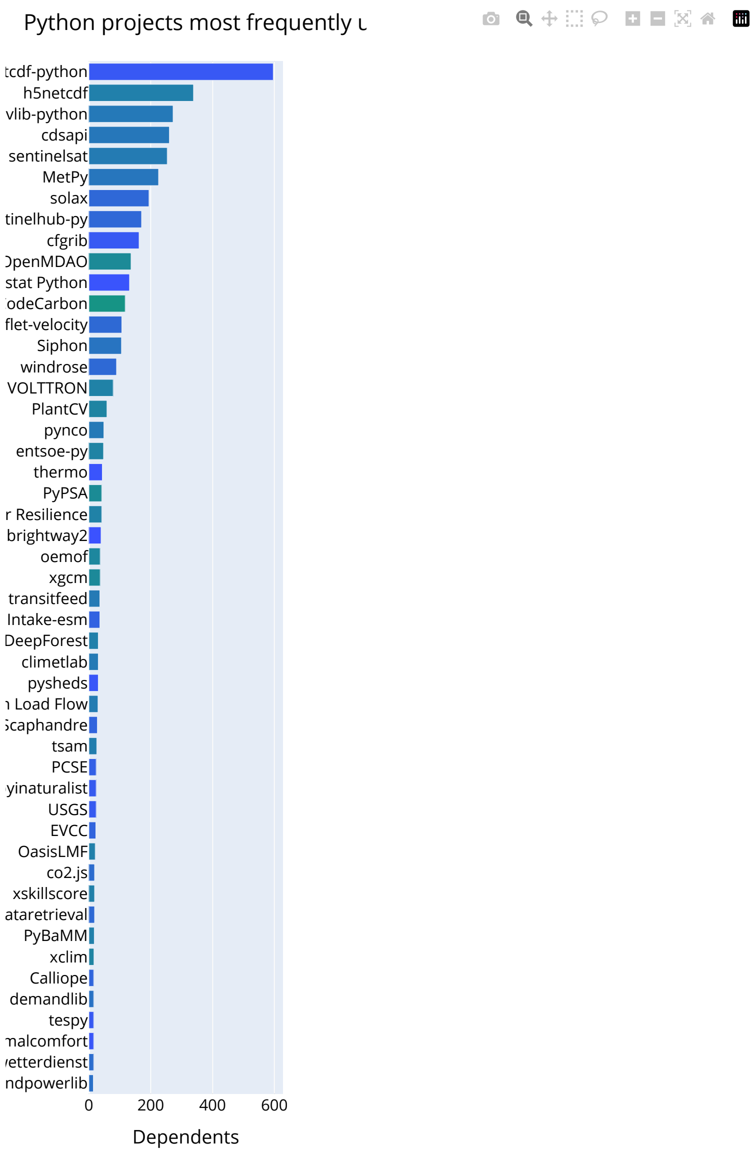


Fig. 41 - Python Projects most frequently used as dependencies

End Users

Capturing the end users and ulitmate 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 significate capacity towards, and demostrate allignment 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 higlight 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, as highlighted in the chapter on growth, 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. OpenLCA is one of the leading lifecycle assessment tool trusted by governments, consultants and researchers to quantify the impacts of goods, services and processes. The FIWARE Foundation's Smart Data Models are used by thousands of smart cities globally, improving the interoperability between smart devices, sectors and organisations and enhancing decision-making related to sustainable development. 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 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."[3] While not exclusive to sustainability, this policy development acknowledges the potential for impact in this area.

Monetary Value

There are few reports available that attempt to assess the economic impact of open source software at national or higher scales. 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 Foward

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 exists 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.

> **Note**
>
> Want to share a case study and/or evidence of open source impact within sustainability? Get in touch.

Ecosystem Collaborations

The OSS ecosystem can not be sustained unless projects, supporting organisations, and businesses collaborate to develop and foster a shared digital infrastructure. Collaboration within and across sectors reduces duplication, concentrates efforts to tackle major issues, and builds social capital. Bridging sub-communities also creates connections between projects, which can affect the sustainability of the projects. For example, research has shown that social network ties between developers influence the survival of OSS projects and that GitHub contributors are more likely to join projects with which they have prior social connections.[1] For these reasons, combined with the urgency to decarbonise various industries, we consider areas where more collaboration is needed. For general recommendations, see Collaboration.

Across and Within Fields

One promising area for more collaboration is between the **physics-based modelling and data science communities and machine learning**. For example, in modelling physical systems for Earth system science or other physics-based models in battery and energy applications. Furthermore, many recent technological advances in machine learning exist, and several sustainable applications could benefit from those advancements. However, such collaborations are hampered by an underlying methodological gap. While new efforts are already being made to integrate machine learning and physics-based models, these distinct communities could join forces to combine the best of both methods: highly accurate forecasts based on machine learning models, long-term knowledge building, and a deep understanding of causal relationships based on physical models. This reduces the risk of relying too much on purely statistical models without understanding the causal relationships. Models thus remain "open source" in the sense that they are still traceable for humans. Decision-making based on these models remains traceable to humans in the final instance.

Competition in both industry and academia can hamper cooperation. This is particularly challenging in the energy systems community, where dozens of similar OSS modelling frameworks exist, and there is little to no incentive from different academic or industry groups to coordinate. Only very few attempts to combine open source projects can be identified. OS-Climate is one of the few efforts to combine different open datasets to provide an integrated environmental rating for sustainable investment. Such approaches are known as Integrated Assessment Modelling. In total, three active projects from this field were found that attempt to build this bridge: [Dynamic Integrated Model of Climate and the Economy], [Global Change Analysis Model] and [Integrated Assessment Modeling Consortium]. The following figure shows a typical result of this modelling, where various economic and social scenarios change the release of carbon dioxide.

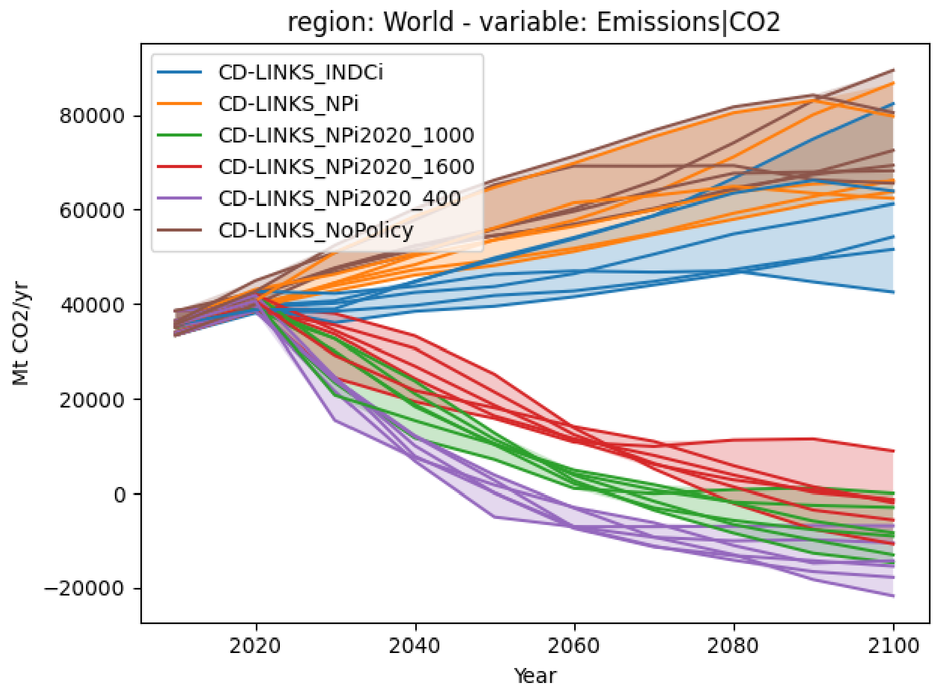


Fig. 42 - Integrated Assessment Modeling with pyam enables the calculation of scenarios for our society by combining ecological and economic models.

The coupling of different Earth systems is another important development in modelling that allows a better understanding and prediction of complex interrelationships, such as climate feedback effects. Notable examples are The Community Earth System Model, the Energy Exascale Earth System Model and Pangeo, which each have their own open source ecosystems and large communities.

In the field of data processing of Earth observation data, the application of Google Earth Engine is ubiquitous. The platform offers open interfaces but is a closed platform and therefore bears the risk, like other large closed-source online platforms, of developing against the users' interest. With their Planetary Computer, Microsoft is taking a different approach. Instead of a proprietary "Earth Engine," it is possible that a free and open community will form around the platform due to its open licence. Moreover, various open source tools and datasets offer an easy entry into this field. Awesome Earth Engine Apps and Awesome Gee Community Datasets provide an overview of the healthy software ecosystem that has emerged in recent years.

Many of the identified projects and communities use similar programming languages and software tools, but few visible links exist between the models and communities — resulting in the "reinvention of the wheel" often occurring within each community. To close the silo of Earth observation archives and to simplify data access, the SpatioTemporal Asset Catalogs API (STAC) was developed. This standard is becoming increasingly widespread with its own toolchains and software ecosystem. It enables developers to access almost all open satellite data from around the world and combine data from various sources in a standardised format. It is an excellent example of how generic applications can be built from OSS components, and different isolated projects can be bridged using standard APIs. Another prime example is the FIWARE Smart Data Models, which seek to improve interoperability between smart devices, sectors and organisations, including the consolidation and dissemination of environmental data.

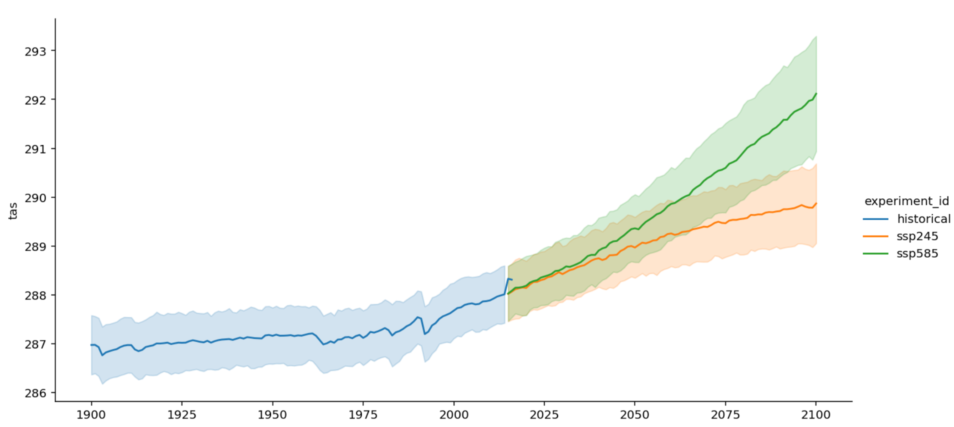


Fig. 43 - Global Mean Surface Temperature prediction with Pangeo, an open source development environment for the earth science. By providing simple interfaces, examples, and data access to state-of-the-art climate simulations like Coupled Model Intercomparison Project (CMIP6), anyone with little Python knowledge can see for themselves the science behind topics like climate change.

Between Different Actors

There is a reasonable level of industry-academia collaboration in domains dominated by the private sector, such as energy modelling and optimisation, battery modelling, wind turbine modelling, PVs, and transportation. Collaboration, however, does not occur at a sectoral or systematic level. In particular, sector-wide coordination in data collection and analysis lags behind. Opening up data requires global collaboration between industry and academia, and new incentives must be created for companies to open up relevant data safely and effectively. Similarly, in sustainable investment and ESG, there is a lack of open data and open source frameworks to evaluate corporate sustainability in a scientific and verifiable way.

Collaboration between state and non-state actors is lacking in many places. State and non-state actors can accelerate their digital and sustainability transformations by collaborating more closely towards shared objectives. This often requires transition brokers to guide and facilitate multi-stakeholder alliances – which bring together governments, industry, academia and civic society – from a politically neutral standpoint, to help create the necessary preconditions for change to emerge across scales and system boundaries. By adopting an open source approach, multi-sector and cross-sector alliances are strengthened through shared knowledge and resources, further strengthing the OSS ecosystem. Very few OSS collaboration efforts of this kind can be identified outside of Europe.

Collaboration between government agencies is uncommon. While several governments promote the sharing and reuse of OSS solutions between agencies,[2,3] very few OSS projects within environmental sustainability can be identified as a result of multi-agency collaboration efforts; neither horizontally across, nor vertically between many levels of government. The interoperability of open knowledge and information is critical to navigating increasingly complex environmental and societal challenges. Given the importance of multi-level governance in tackling climate change, the lack of open collaboration and innovation here is concerning.

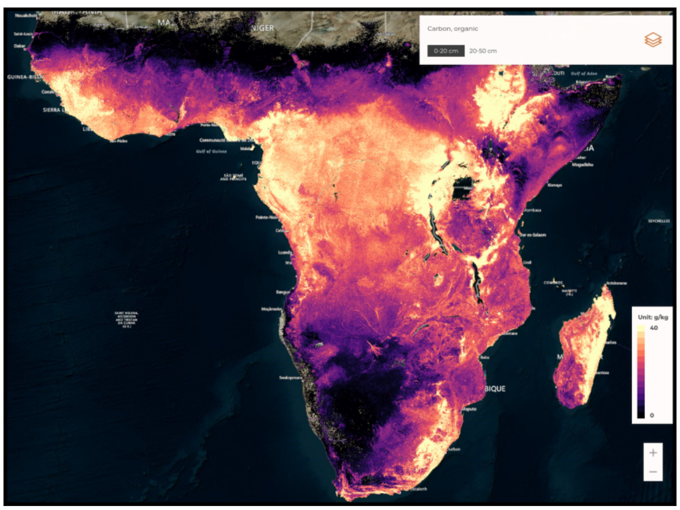


Fig. 44 - iSDA built the first field-level soil map for Africa, with 20+ soil properties predicted at 30m resolution for the entire continent. License: CC BY 3.0

Funding Models

The developers and organisations 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.

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 seeking financial support.

Public sector funding from research councils, the European Union, academic labs and national labs play a transformative role in sustaining OSS projects in sustainability. Universities are not only knowledge providers but can also bring stability, independence and access to developer resources via their Research Software Engineering teams.

However, there are very few dedicated funds for OSS development and maintenance in sustainability. Large academic projects prioritise 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. There is no dedicated fund or consideration around OSS maintenance; Developers and Research software engineers, in academic labs, are trying to squeeze in person-months in multiple concurrent projects and make this open research infrastructure reusable across many projects. And since features are used across many projects, assigning maintenance or user support resources to any particular OSS development is complex.

> *"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

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. The developers we spoke with had prior experience with hybrid business models such as Open Core, where the backbone of the codebase is open source. At the same time, the project maintainers want to support open source without commercialisation being at the expense of the communities. Maintainers are aware that their users and contributors may be upset if they feel that open source work is being used by companies for profit. Choosing the right open source business model, developer involvement and organisational 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 centre 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, it is critical to demonstrate and quantify the impact of OSS sustainability projects across industry and civic society. Government departments fund academic labs that, in turn, collaborate with industry partners to develop new functionality. In addition, research funders and government agencies, such as the US Department of Energy, are interested in encouraging innovation in specific strategic sectors, such as renewable energy. When academic projects can demonstrate the value of open source work outside academia and the co-benefits and return on investment to public funders, it becomes easier for these projects to raise funding.

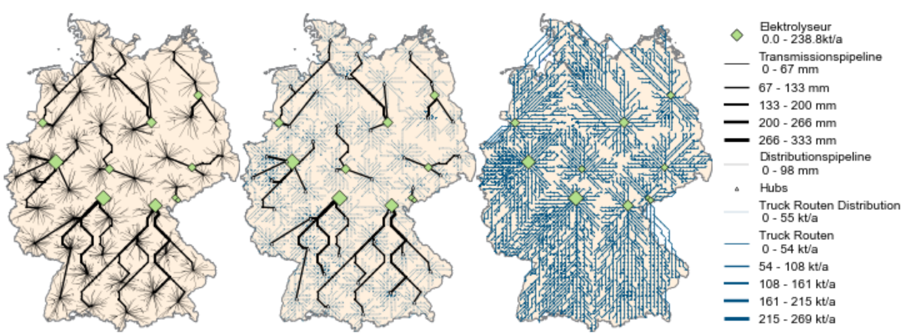


Fig. 45 - Hydrogen Infrastructure Model offers the functionality to calculate predefined hydrogen supply chain architectures with respect to spatial resolution for the analysis of explicit nationwide infrastructures.

License: MIT

Data Wishlist

This is a list of datasets whose availability could accelerate progress or remove bottlenecks in different areas examined in this study. Since many open source projects subscribe to the concept of, and employ, other forms of openness, including open data, quantifying the available data stock based on this study's identified projects is possible in the future. This can help to create datasets on thematic priorities with temporal and geographical coverage.

This "wishlist" was obtained via interviews with project developers. Comparable wishlists have previously been created by ClimateChangeAI. We classify potential datasets according to the topics selected within our study and the following types of data availability:

1. Public, needs structure
2. Private, needs to be released
3. Scattered, needs collation
4. Proxies, needs to be inferred
5. Scarce, needs collection

Dataset	Description	Availability	Next Steps
Multicriteria data about bare-metal server manufacturing and their components	Measuring the power consumption of servers accounts for ~70% of the IT emissions. 30% of the footprint comes from server manufacturing. We cannot create comprehensive impact assessments without measuring resource depletion, the use of metals, the impact of resource extraction on biodiversity and people and the end of life of the product.	1 and 4	
Global detailed demographic data and Census breakdown at the city and neighbourhood level for transport planning interventions	Analysis-ready census data for public transport simulations do not exist publicly. When planners simulate traffic, edit streets and intersections and plan bike networks, they need to investigate equity questions about any kind of intervention i.e. who is going to benefit from a low-traffic neighbourhood or see increased traffic. While census data sets are available in the US and the UK, they are cumbersome to work with i.e. parsing, merging two different census tracks or slicing them. Where they exist, they are expensive to licence.	2 and 3	Aggregate, contextualise, and standardise data from publicly available datasets
Open UK map of curbside space—from parking meters, loading zones, curb cuts, and ramps to make the necessary changes for parking, ride-hail, and micro-mobility access for their community	Curbs and their regulations have a significant impact on citizens and businesses daily. Curbs, for example, dictate where, when, and how long you can park, the dedicated use of a parking spot, such as loading, pick-up/drop-off, or accessible parking, and the type of vehicle a parking spot is reserved for (i.e. motorcycles, taxis, trucks), and so on. With so many different curb uses dispersed throughout a city, cities must have up-to-date information on curb assets that can then be incorporated into broader urban plans to make the most effective decisions to maximise a streetscape. Currently, the Ordnance survey holds curb data in the UK, which is not public. If the datasets were made public, the public sector would have the information required to make strategic decisions, advance policies, and respond to changing market dynamics with the same agility as their private sector counterparts.	2 and 3	Those data are available at a cost through various providers.
Travel demand	Traffic planning software requires detailed data on the trips that people take daily i.e. where they come from and where they go to. These can be built from aggregate census data, but they are hard to parse for transport simulations.	2 and 3	
Grid	We need good data to tackle climate change. This data is available; it is published by the International Energy Agency (IEA). Despite being a largely public-funded organisation, a lot of IEA data is behind paywalls. Without it, subsectors and nations cannot accurately model their energy consumption or develop projections for usage in a range of scenarios, and they will ultimately be unable to manage their transition to zero-carbon economies. (OpenMod letter)	2	Demand data: In certain parts of the United States, aggregated demand data may be publicly available through Independent System Operators or Regional Transmission Operators (ISOs/RTOs) and can simply be pulled from their websites. In the case where more granular data is needed, or in regions (within and outside the US) where such data is not public, it may be necessary to coordinate with relevant system operators.
Consumer gas consumption	Data on how much gas consumers use in their homes or industry doesn't exist, and researchers are using heuristics to generate the data.	4 and 2	
Gas grid infrastructure data (with a focus on the North Sea)	To create more accurate IAMs and assess where wind turbines can be installed, researchers require spatially resolved demand datasets to be examined alongside data on shipping lanes and nature reserves, land restrictions, birdlife restrictions so that they can exclude certain areas from the models.	2 and 3	Environmental data for these kinds of assessments appear to be public sources for many regions, though they appear to be somewhat inaccessible, and the granularity of the measurements is unclear.
Battery operation, electrical and mechanical response under many conditions	There is a pressing need for more open battery data, including current, voltage, and capacity curves, operation cases and degradation information. Obtaining data for the entire lifetime of a battery with many batteries and many different conditions is time-consuming and expensive. The conditions that affect the lifecycle of batteries are complex and grow exponentially in size. While academic labs are willing to publish openly, industry (e.g car manufacturers) don't want to share their proprietary data. Thus, academic labs working on battery modelling research are often constrained by not being able to compare their data with that of industry, and validating models becomes more difficult. This means researchers could be overfitting the battery models to the scant data available.	3 and 2	These datasets are available in various publications, so some degree of standardisation is required in addition to collection efforts.
Wind Turbine Controllers (fixed and floating offshore wind turbines).	The wind turbine systems research community uses reference controllers to compare and evaluate modern and advanced control algorithms and enable dynamic simulations in other wind turbine studies. Design constraints, such as blade tip deflection, have become increasingly important as wind turbines have grown and modelling tools have improved and increased in fidelity. Without a representative controller that performs consistently across turbine designs, dynamic simulations cannot provide reliable turbine design results. Reference wind turbine controllers are often difficult to modify for other turbines, making it difficult for researchers to run representative, fully dynamic simulations of other wind turbine designs. This is because the control software algorithms companies keep these data close. Any open commercial controllers or their data would be helpful in this line of research. To our knowledge, a modern, open source controller with specific logic for floating offshore wind turbines (FOWTs) is not available	2	Wind turbine modellers must collaborate with wind turbine companies and industrial manufacturers to create and/or access these datasets.

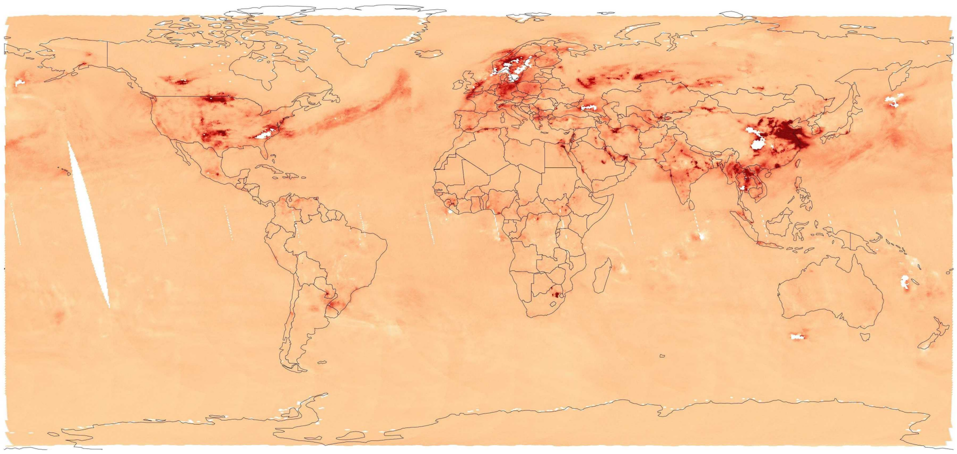


Fig. 46 - Emissions API's mission is to provide easy access to satellite based emissions data without the need of being an expert in satellite data analysis and without having to process terabytes of data. License: MIT

Overall

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

- **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 standardised 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 standardised 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. Projects, networks, and collaborations between independent scientists, regulators and industry that support open science in sustainability can facilitate a deeper understanding of the environmental impacts of entire industries, and provide investors and consumers with verifiable, life cycle assessments of products and services. 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. Meanwhile, such investments are resource efficient and easily transferable worldwide. The Learning 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 technologies, there are few incubators or support programs specialising in open source project development in environmental sustainability. While a handful of individual programs focus on non-profits 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 commercialised 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.

- **Transform financial institutions through transparent and scientific decision-making for sustainable investments.** Banks, rating agencies and investors are at the centre of our economy, 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 lack of openness means decision-making is often opaque and vulnerable to greenwashing. 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 favourable 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 recognised 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 prioritises 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.

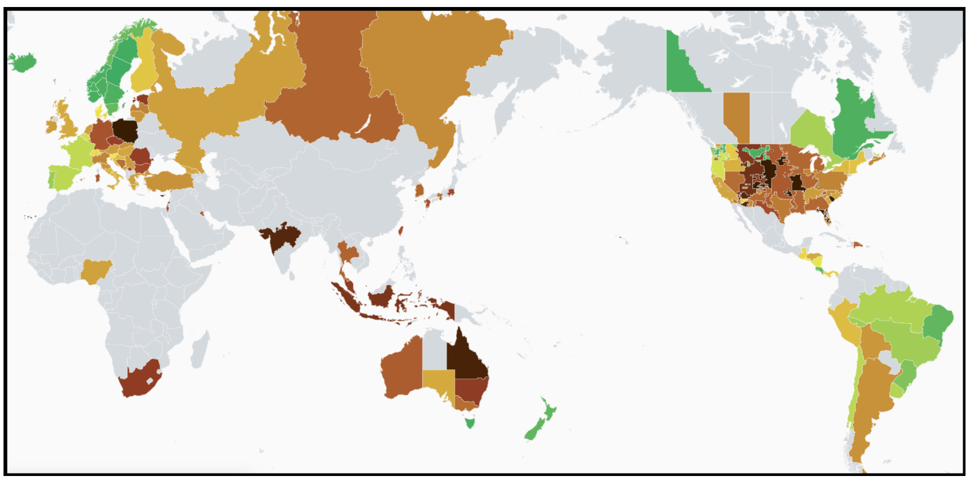


Fig. 47 Electricity Maps provides time-resolved historical data and forecasts about the amount of renewable energy in the electrical grid for countries all around the globe. Energy consumption can be adjusted in such a way that as few emissions as possible are produced. It provides both a solid business model and an open source community behind the project. License: MIT

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 learnt here 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 modelling, 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 revolutionise 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 organisation 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. Organisations 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.

- **Standardise environmental data and models using open source across different levels of government.** The standardisation 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 delivers 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 standardised 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 realised 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 organisations. 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 labour 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 optimised 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 decentralised 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.

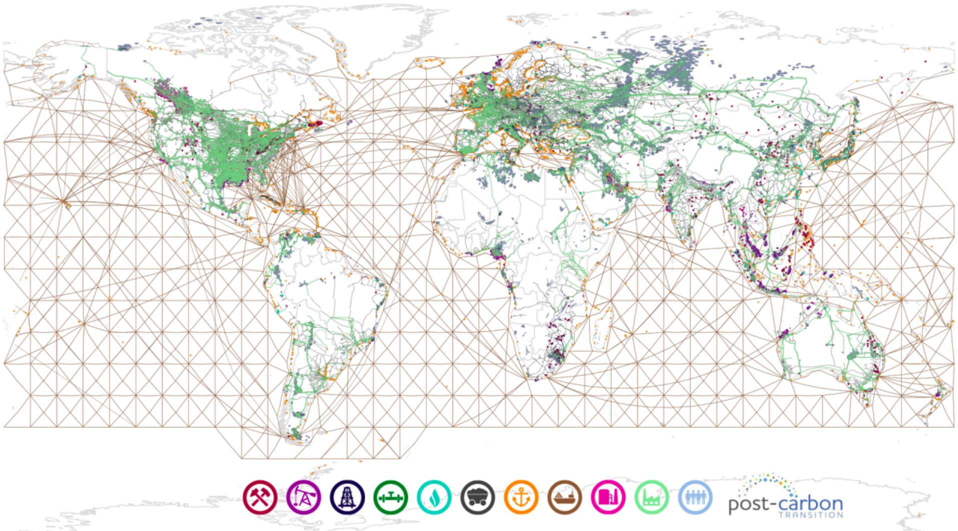


Fig. 48 Asset-level Transition Risk in the Global Coal, Oil, and Gas Supply Chains. License: MIT

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 organisations, 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.

- **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 organisation.

- **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. As a high degree of technical expertise is often required, it is necessary for OSS developers to integrate such technologies within user-friendly interfaces. 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 standardised interface, while researchers can share digital 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. The reasons for this are vast, complex, and vary across regions, but is often the result of market forces and public policy, with universities increasingly adopting a for-profit orientation. While the protection of intellectual property remains common practice, research constantly demonstrates that regulatory and legal obstacles actively hampers innovation[1,2]. 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 government.** 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.

- **Create open communities for monitoring greenhouse gas emissions through remote sensing.** Although emissions are central to climate change, and Earth observation provides an unprecedented wealth of new data and information about our planet, our analysis highlights a lack of collaboration in this emerging field. 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 licences 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.

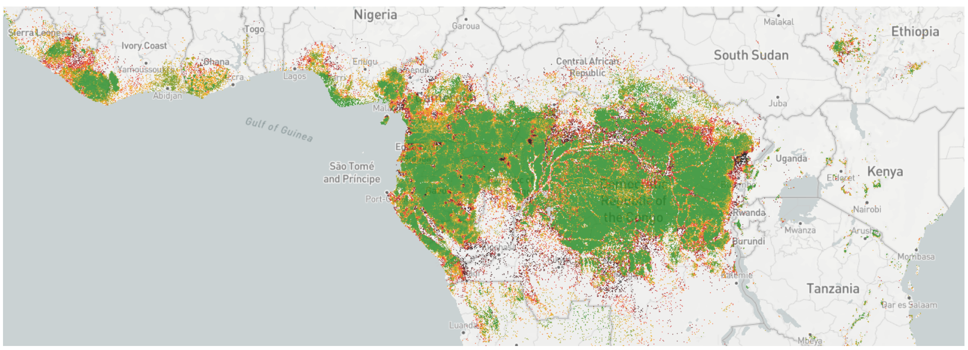


Fig. 49 The forestatrisk Python package can be used to model the tropical deforestation spatially, predict the spatial risk of deforestation, and forecast the future forest cover in the tropics. License: GPL-3.0

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**. Organisations across sectors are increasingly recognising their own reliance on open source, with over 97% of modern digital infrastructure leveraging open source code.[1] There is also a realisation of the material benefits organisation 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 recognise 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 maximise 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 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. There are many direct and indirect benefits of taking this approach, which are highlighted throughout this report. 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,[1,2] labour productivity,[1] and the formation of start-ups.[3] Increased economic complexity and resource efficiency gains are also evident.[1,2] According to research from the European Commission, an increase in contributions to open source software (OSS) of 10% per year results in a GDP increase of 0.4% to 0.6%, as well as the creation of more than 600 new technology start-ups in the EU.[3] 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.

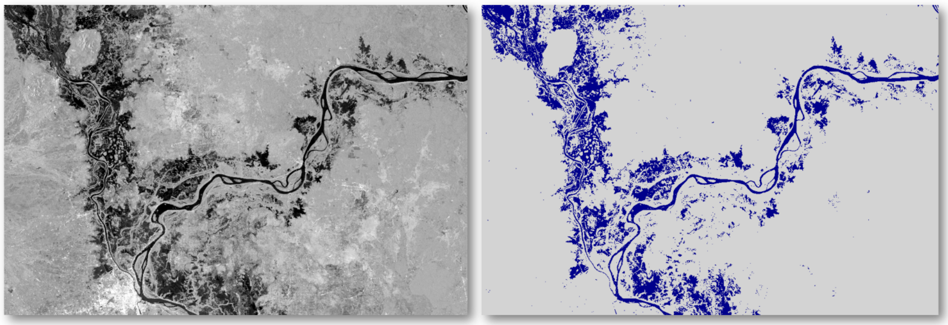


Fig. 50 HYDRAFloods is an open source Python application for downloading, processing, and delivering flood maps derived from remote sensing data. License: GPL-3.0

By Tobias Augspurger, Eirini Malliaraki and Josh Hopkins