



---


# A TOOLKIT FOR MEASURING THE IMPACTS OF PUBLIC FUNDING ON OPEN SOURCE SOFTWARE DEVELOPMENT

---

**Cailean Osborne\***   
University of Oxford  
Oxford, UK

**Paul Sharratt**   
Sovereign Tech Agency  
Berlin, Germany

**Dawn Foster**   
The CHAOSS Project  
London, UK

**Mirko Boehm**   
The Linux Foundation  
Berlin, Germany

## ABSTRACT

Governments are increasingly employing funding for open source software (OSS) development as a policy lever to support the security of software supply chains, digital sovereignty, economic growth, and national competitiveness in science and innovation, among others. However, the impacts of public funding on OSS development remain poorly understood, with a lack of consensus on how to meaningfully measure them. This gap hampers assessments of the return on public investment and impedes the optimisation of public-interest funding strategies. We address this gap with a toolkit of methodological considerations that may inform such measurements, drawing on prior work on OSS valuations and community health metrics by the Community Health Analytics Open Source Software (CHAOSS) project as well as our first-hand learnings as practitioners tasked with evaluating funding programmes by the Next Generation Internet initiative and the Sovereign Tech Agency. We discuss salient considerations, including the importance of accounting for funding objectives, project life stage and social structure, and regional and organisational cost factors. Next, we present a taxonomy of potential social, economic, and technological impacts that can be both positive and negative, direct and indirect, internal (i.e. within a project) and external (i.e. among a project’s ecosystem of dependents and users), and manifest over various time horizons. Furthermore, we discuss the merits and limitations of qualitative, quantitative, and mixed-methods approaches, as well as options for and hazards of estimating multiplier effects. With this toolkit, we contribute the multi-stakeholder conversation about the value and impacts of funding on OSS developers and society at large.

**Keywords** Open source software · funding · sustainability · impact measurement · maintenance · digital commons

## 1 Introduction

Open source software (OSS)<sup>2</sup> are digital public goods that are increasingly recognised as digital infrastructure [2, 3], which are used in around 96% of codebases [4] and constitute up to 90% of commercial software stacks [5]. The ubiquitous use of OSS, which are often developed and maintained by volunteer communities, have drawn attention to the question of funding as a mechanism to support the sustainability of OSS projects [6, 7]. In particular, the discovery of major security vulnerabilities in widely used OSS projects, such as the Log4Shell vulnerability in November

---

\*Corresponding author: Cailean Osborne, cailean.osborne@oii.ox.ac.uk

<sup>2</sup>OSS is software whose source code is distributed under a license that permits the use, study, modification, and redistribution of the software source code [1]. Please note that we use OSS rather than FOSS, which stands for “free and open source software”.

2021, highlighted the need to fund security and maintenance work in critical OSS projects [8], and contributed to an understanding of the role and responsibility of the public sector as a funder of OSS development [9, 10]. Concurrently, policymakers increasingly recognise OSS funding as a policy lever to support digital sovereignty [11, 12], the growth of domestic software markets [13], and national competitiveness in science and innovation [14, 15], among others.

While governmental interest and involvement in funding OSS funding is increasing, its impacts on OSS development is poorly understood, with no consensus on how to measure them in a meaningful way. Not only is the measurement of the impacts of OSS funding methodologically challenging, but it is complicated by the fact that introducing funding into OSS projects may change contributor incentives and the balance of voluntary and paid participation [16, 17]. This gap hinders assessments of return on public investment and optimising public interest funding strategies. We address this problem with a toolkit of methodological considerations for measuring the impacts of public funding on OSS development. It is informed by methodologies for OSS valuations and community health measurements, in particular metrics developed by the Community Health Analytics Open Source Software (CHA OSS) project, as well as our first-hand insights from developing impact measurement frameworks for the Next Generation Internet (NGI) initiative at the European Commission (EC) and the Sovereign Tech Agency (STA) in Germany.

The toolkit begins with a discussion of key considerations, including funding objectives, project life stages and social structures, and regional and organisational cost factors. Then, we present a taxonomy of potential social, economic, and technological impacts that can be both positive and negative, direct and indirect, internal (i.e. within a project) and external (i.e. among a project’s ecosystem of dependents and users), and manifest over various time horizons. Next, we discuss the merits and limitations of qualitative, quantitative, and mixed-methods approaches for measuring such impacts, and options for and hazards of estimating multiplier effects. This toolkit is not exhaustive or prescriptive, nor does it aim to be; rather, we seek to provide a toolkit that can inform the multi-stakeholder debate about the value of public funding for OSS development and how to meaningfully measure its impacts.

This paper has the following structure. Section 2 reviews relevant literature and practice, providing a background on prior scholarship on OSS funding, OSS valuations, and OSS community health measurements, as well as real-world examples of public funding by the NGI and STA. Section 3 presents the toolkit for measuring the impacts of public funding on OSS development. It includes a discussion of pertinent considerations, such as funding objectives (3.1.1); salary structures (3.1.3); a taxonomy of social, economic, and technological impacts (3.2); multiplier effect estimations (3.5); and an evaluation of qualitative, quantitative, and mixed-methods approaches (Section 3.4). Section 4 discusses the overarching considerations and future research directions. Finally, Section 5 concludes with a call to action for both diverse OSS stakeholders to engage in the debate about the value and impacts of public funding on OSS development.

## 2 Review of Literature & Practice

### 2.1 Review of Prior Work

#### 2.1.1 Funding OSS Development: History, Funders, and Impacts

OSS are digital public goods [18] that are increasingly recognised as digital infrastructure [2, 3], owing to their use in around 96% of codebases [4] and constitution of up to 90% of commercial software stacks [5]. In light of their ubiquitous use, the topic of funding for development, maintenance, and security of OSS has become increasingly important in developer, industry, and policy circles alike. In particular, high-profile security vulnerabilities like the Heartbleed bug in 2014, the Log4Shell vulnerability in 2021, or the recent XZ Utils backdoor in 2024 have underscored the unsustainability of relying on volunteer communities that maintain a large portion of critical OSS projects [8].

Historically, money has been a contentious issue among OSS developers, many of whom take pride in the origins of OSS as a social movement against the privatisation of software or as a “programmers’ paradise” [19], and who do not contribute for financial gain [17], but out of their intrinsic satisfaction of learning and solving technical problems [20], their altruism [21], or their political ideals [22], among others. Some even argue that “money ruins everything” [23]. In addition, evidence from the Rust community illustrates that volunteers have some prejudices against paid developers, such as that they do “do boring work”, “rarely care [for] documentation”, and “lack personal attachment” [24].<sup>3</sup> However, growing concerns about maintainer burnout [6], security vulnerabilities [8], and widespread commercial freeriding on the labour of volunteers [25, 7] have changed minds in OSS developer communities [14, 26]. As Benjamin Birkinbine contends, there is an urgent need “not just [for] investment in institutions, organisations, technologies, or innovations, but long-term and sustainable investment in the true source of their value: people” [7].

---

<sup>3</sup>This study also finds differences in the activity between volunteer and paid developers; for example, paid developers tend to contribute more frequently and implement more features than volunteers [24].

### ***Micro-donations: Overview and Impacts***

Various actors have been involved in funding OSS development, including individuals, companies, philanthropic organisations, governments, and public research agencies. Individuals, typically users and/or developers of OSS, contribute through micro-donations via platforms like NumFOCUS and GitHub Sponsors [27]. Motivations for micro-donations include individuals' desire to show their appreciation for an OSS project and to encourage further development [28]. While the relative impact of micro-donations appears limited [29], they have been shown to shorten maintainer response times to issue reports [30, 31], increase maintenance-related activities [32], and facilitate project spending on community events and summer internships for developers from under-represented groups [14].

### ***Commercial Funding: Overview and Impacts***

The private sector has been the largest direct and indirect funder of OSS development to date, which is unusual among public goods. It is common for companies to sponsor OSS developers and to let their employees contribute to OSS projects during work hours, either as a job responsibility or as part of a voluntary initiative [33, 34, 35]. Companies also sponsor OSS foundations and consortia, paying membership fees to the organisations that host projects [36]. The impact of such funding is contested: while only 12% of a Tidelift survey respondents found this type of funding very useful [37], a case study on the *scikit-learn* project found that commercial sponsorship of its consortium was useful for balancing the influence of companies and provided resources to employ full-time maintainers [14]. Some companies run OSS contribution programmes, such as the Google Summer of Code programme [38], and some have established “FOSS Contributor Funds” that allow employees to nominate OSS projects that they would like to sponsor [39, 40]. However, these funds typically offer limited amounts for short periods of time (e.g. \$10,000 for one year). In addition, companies spin-out proprietary software into OSS projects and invest in their ecosystems as a strategy to increase adoption of their software, benefit from external contributions, or reduce a competitor's market share [41, 42].

The economic value of commercial funding and investments in OSS development are significant by many measures. It has been estimated that €1 billion invested in OSS by companies in the European Union (EU) in 2018 generated up to €95 billion for the EU's gross domestic product (GDP) that year [43], and that companies in the USA invested \$37.8 billion in OSS in 2019 [44]. Furthermore, country-level OSS development activity and the founding of start-ups are positively correlated, suggesting that OSS development activity is a catalyst for innovation [45]. Nevertheless, we face a stark imbalance between the supply-side value (\$4.15 billion) of OSS development and its demand-side value (\$8.8 trillion)[46]. These numbers underline the “risk of underproduction” stemming from the misalignment of the supply of OSS development labour and the demand for OSS [47]. Beyond funding, the impacts of commercial participation on collaborative practices and norms in OSS developer communities is an active area of research [48, 49, 50].

### ***Public Funding: Overview and Impacts***

While the public sector is often overlooked as a major funder, it has funded OSS directly and indirectly since the inception of the world wide web at CERN through research grants [14, 15]. The intensity of public funding for OSS development has changed in recent years, with the recognition of funding as a policy instrument to enhance the security of software supply chains and digital infrastructure [9], digital sovereignty of nations and blocs [11, 12], and national competitiveness in science and innovation [14].<sup>4</sup> Examples of public funding include research grants for innovation [14]; award and prize programmes for OSS developers, such as the BlueHat Prizes by France's Free Software Unit and NLNet [51]; bespoke OSS funding bodies, such as the Open Technology Fund in the USA, the NGI initiative in the EU, and the STA and Prototype Fund in Germany [10]; and bug bounty programmes for security-focused maintenance of OSS projects [52]. However, the most common way in which governments have indirectly supported OSS to date has been through procurement policies that favour OSS over proprietary solutions [11, 53]. Such policies have led to increased domestic OSS development activity and the growth of domestic software markets [13].

Prior work provides the following insights on the impact of public funding on OSS development. Few public research grants provide dedicated funding for scientific OSS development and those that do tend to prioritise innovation over maintenance [54]. A case study on a €32 million grant awarded by the French government to the *scikit-learn* project highlights both challenges and benefits of research grants, from its initial focus on innovation over maintenance and misalignments between policy objectives and the needs and know-how of the maintainers, to benefits for the project's long-term stability, ecosystem development, and maintenance capacity [14]. Similarly, an evaluation of the “Software

---

<sup>4</sup>We note that beyond OSS direct funding for R&D is recognised as one of the most effective, short-term policy instruments for promoting innovation [15].

Sprint” funding programme by the Prototype Fund<sup>5</sup> shows mixed results. On the one hand, about two-thirds of funded projects successfully completed a functional prototype during their funding period and roughly 60% continued development after funding ended [55]. In addition, the funding fostered the professional growth of individuals, with over a quarter developing new career opportunities. On the other hand, the programme faced challenges in supporting the sustainability of projects; for example, over half struggled to secure follow-up funding for scaling [55]. In addition, an evaluation of the NGI initiative shows that it has supported the development of OSS projects that protect EU digital values and rights, enable EU legislation, and provide alternatives to proprietary solutions, among others [56].

In addition, research illustrates the positive impact of funding for security work. There is a moderate positive correlation between general-purpose funding and enhanced security practices among the 1,000 most downloaded OSS packages in the Python and Javascript ecosystems [57]. In addition, the greater number of funding sources in a project corresponds with better security practices [57]. Bug bounty programmes enhance OSS security in mature projects, but should only be implemented after basic security practices are in place, as they can potentially overwhelm under-maintained projects and may disrupt the reciprocal nature of OSS communities unless properly designed [52]. Holistic funding approaches that pair targeted security funding with general funding for maintenance are advised [52].

### 2.1.2 OSS Valuation Models and Community Health Approaches

An obstacle that stands in the way of funding for OSS development is that to make the argument to invest in OSS development, managerial decision-makers demand quantitative evidence of the value of OSS contributions [58]; for example, a return on investment (ROI) for a company or the responsible use of taxpayers’ money for a public sector organisation. However, it is difficult to quantify the value of OSS contributions [58]. Motivated by the motto, “If we can appropriately quantify the value of OSS work, then we can adequately invest in it,” Sophia Vargas reviewed valuation models that can be used to measure the value of OSS [58] (see Table 1). Vargas identifies several key challenges in estimating the value of OSS: the need for consistent measures of demand and usage [59], the difficulty in quantifying non-code contributions [60], and the “force multiplier” effect where collective benefits exceed individual inputs [58]. While economic models like GDP may guide government funding strategies, methodological challenges (e.g. data availability and causality issues) limit the robustness of such measurements and they are less applicable to companies and individuals who are motivated by different types of value creation (e.g. skill development or ROIs).

What is more, Vargas cautions against over-reliance on technical metrics like commits or lines of code (LOC), which flatten socio-technical dynamics that characterise activity in OSS projects and their wider ecosystems of dependents and users. By contrast, socio-technical models that focus on the social structure and health of OSS projects provide a more comprehensive lens on value creation in OSS projects and ecosystem [61]. The health of an OSS project concerns “an OSS project’s capability to stay viable and maintained over time without interruption or weakening” [62]; or in other words, the sustainable availability of maintenance labour that can come from either the maintainers or the wider community [63]. A key ingredient for the health of OSS projects is the provision of “human infrastructure” that can support and secure maintenance labour from both project maintainers and their community of contributors, including the securing of a work-life balance, the management of social pressure, and the diversity of contributors [63]. The authors note funding as an important enabler for this human infrastructure [63].

While every OSS project is unique, Nadia Eghbal identifies four common social structures based on a project’s contributor-to-user ratio: federations, clubs, stadiums, and toys [6]. Federations have both high contributor and user growth, typically with complex governance processes and working groups. Clubs have high contributor growth but lower user growth, characterised by tight-knit communities of enthusiasts. Stadiums have high user growth but low contributor growth, typically maintained by a small group of developers. Toys are personal projects with low contributor and user growth. Beyond developers, Julia Ferraioli advocates for a social model of OSS that considers “the people who consume, contribute to, and maintain it, rather than a technical model of open source, based on the legal rights and responsibilities as defined in the license” [64]. Similarly, “all contributors” models consider both technical and non-technical contributions to OSS projects, providing a more comprehensive representation of contributorship [60].

Measuring the health of an OSS project is complex because there are many sociotechnical aspects to consider, from a project’s social composition (e.g. culture, governance, and contributor diversity) to its long-term stability (e.g. funding and code maintainability) [62]. For this reason, scholarship on community health and the metrics of the CHAOSS project look beyond the code, considering social aspects like community welcomingness and diversity, equity, and inclusion; economic aspects, such as funding and business readiness; as well as technological aspects like security and license compliance [65]. Furthermore, a key principle of the CHAOSS project is that every project is different, and accordingly metrics should always be interpreted with the needs and context of the project taken into account [61].

---

<sup>5</sup>The Prototype Fund is a funding program of the Federal Ministry of Education and Research (BMBF) that is managed and evaluated by the Open Knowledge Foundation Germany. Its “Software Sprint” distributed €12.3 million across 12 funding rounds to 293 projects that aimed to build public-interest OSS that addressed societal challenges.

Table 1: Valuation Models and their Application to OSS Development by Sophia Vargas (2024) [59]

Model	Application to OSS	Key Variables	Challenges
<b>Constructive Cost Model (COCOMO)</b> <ul style="list-style-type: none"> <li>COCOMO estimates the effort, cost, and schedule of software projects [66]</li> <li>COCOMO II includes reused/adapted code &amp; code incorporation effort [67]</li> </ul>	<ul style="list-style-type: none"> <li>LOC as proxy variable for value creation</li> <li>Complexity multipliers for context adjustment</li> </ul>	<ul style="list-style-type: none"> <li>LOC</li> <li>Task complexity</li> </ul>	<ul style="list-style-type: none"> <li>No maintenance tracking</li> <li>Does not account for code history or maintenance</li> </ul>
<b>DevOps Research and Assessment (DORA)</b> <ul style="list-style-type: none"> <li>Evaluates maturity of DevOps performance and processes [68]</li> <li>Focus on operational metrics for DevOps evaluation [69]</li> </ul>	<ul style="list-style-type: none"> <li>Quality versus quantity measurement</li> <li>Compares upstream versus company-specific quality</li> </ul>	<ul style="list-style-type: none"> <li>Deployment frequency</li> <li>Lead time for changes</li> <li>Change failure rate</li> <li>Time to restore service</li> </ul>	<ul style="list-style-type: none"> <li>Assumes team structure</li> <li>Heterogeneity of OSS communities (e.g. varied incentives and time commitments)</li> </ul>
<b>OSS as Socio-technical Model</b> <ul style="list-style-type: none"> <li>Complex software development processes by loosely coordinated developers [70]</li> <li>Social, economic, political network relationships [71]</li> <li>Social model: classify project by creator intent [64]</li> </ul>	<ul style="list-style-type: none"> <li>Recognises OSS projects as complex systems</li> <li>Captures diversity of contributors and incentives</li> <li>Recognises multiple value types for individuals (skill development, learning, etc.)</li> </ul>	<ul style="list-style-type: none"> <li>People (e.g. incentives)</li> <li>Infrastructure (e.g. tools)</li> <li>Content (e.g. documents)</li> <li>Governance (e.g. participation in roadmap decisions)</li> </ul>	<ul style="list-style-type: none"> <li>The more complexity one recognises, the more difficult it becomes to quantify value</li> <li>What kind of value, by whom, and for whom?</li> </ul>
<b>Business Models (TCO/ROI)</b> <ul style="list-style-type: none"> <li>Total cost of ownership (TCO): a financial estimate of direct and indirect costs</li> <li>Return on investment (ROI): ratio between net income (over time) and investment</li> </ul>	<ul style="list-style-type: none"> <li>Business benefits of OSS [72], some more quantifiable than others (e.g. cost savings vs. open standards creation)</li> <li>TCO includes infrastructure, maintenance, and integration</li> <li>ROI based on percentage of codebase that is OSS</li> </ul>	<ul style="list-style-type: none"> <li>Infrastructure costs</li> <li>Maintenance expenses</li> <li>Community funding</li> <li>Integration effort</li> </ul>	<ul style="list-style-type: none"> <li>Organisationally subjective</li> <li>Assumption of “OSS is free”</li> <li>Difficult to measure % of income associated with a OSS package or project</li> </ul>
<b>Economic Models (e.g. GDP)</b> <ul style="list-style-type: none"> <li>GDP measures market value of goods and services in a time period in a country</li> </ul>	<ul style="list-style-type: none"> <li>Tracks economic impact</li> <li>Estimated: €1B invested in OSS in 2018 generated €65-95B for EU GDP [43]</li> </ul>	<ul style="list-style-type: none"> <li>OSS contributions</li> <li>Investments in OSS</li> <li>Economic indicators</li> </ul>	<ul style="list-style-type: none"> <li>Limited data availability</li> <li>Attribution difficulties</li> </ul>
<b>Supply and Demand Models</b> <ul style="list-style-type: none"> <li>Market equilibrium analysis</li> </ul>	<ul style="list-style-type: none"> <li>OSS creation versus replacement cost estimation [46]</li> <li>Estimated: Supply-side value = \$4.15 billion, demand-side value = \$8.8 trillion [46].</li> </ul>	<ul style="list-style-type: none"> <li>Value: Cost of software development labour (LOC)</li> </ul>	<ul style="list-style-type: none"> <li>Difficult to measure due to non-pecuniary nature and lack of usage tracking [46]</li> <li>LOC focus ignores maintenance, non-technical labour, and different types of value</li> </ul>
<b>Underproduction Risk Model</b> <ul style="list-style-type: none"> <li>Adapted supply and demand model for OSS context [47]</li> </ul>	<ul style="list-style-type: none"> <li>Maintenance-focused</li> <li>Recognition of collaboration dynamics in OSS projects</li> </ul>	<ul style="list-style-type: none"> <li>Maintenance supply</li> <li>Usage demand</li> </ul>	<ul style="list-style-type: none"> <li>Correlational limitations</li> <li>Quality measurement issues</li> <li>Lack of usage data</li> </ul>

## 2.2 Review of Practice

In this section, we review two public funding initiatives that we are familiar with as practitioners: the NGI by the EC and the STA in Germany. The aim is to demonstrate the diversity of designs and implementations of public funding for OSS, demonstrating that a one-size-fits-all approach to measuring the impact of funding is not feasible. We review the NGI initiative and the STA as examples of public funding because two authors are members of the NGI Commons consortium tasked with evaluating the NGI initiative<sup>6</sup> and one author is an employee of the STA tasked with evaluating the impact of the STA's funding, thus providing a practitioner perspective on the impact of these funding programmes.

### 2.2.1 Next Generation Internet

The NGI initiative by the EC is a public funding initiative that funds research and development (R&D) for open internet technologies in the name of an "Internet of Trust". As part of Horizon Europe, the NGI initiative has provided €140 million for over 1,200 projects between 2019 and 2024, with an additional €32 million allocated for 2024-2027 [73]. The majority of funded projects have contributed to OSS projects and/or released solutions in OSS repositories [56]. The NGI initiative employs two main funding mechanisms: Research and Innovation Actions (RIAs), which focus on funding the development of NGI technologies, and Coordination and Support Actions (CSAs), which facilitate scaling efforts through outreach, community building, and international collaboration.

The NGI initiative's RIAs employ cascade funding to distribute funds, allocating 20% to intermediary coordinators and 80% to third-party individuals and organisations.<sup>7</sup> The RIAs have employed different funding approaches. First, over €50 million has been distributed in small-to-medium, equity-free grants, ranging from €5,000 to €50,000, to hundreds of recipients, including individuals and start-ups (approximately 800/>1,200 projects by NLNet). These grants are completed by delivery milestones within 12 months, and recipients are paid per milestone completion. According to a NGI coordinator, this funding approach "evaporates the BS factor" and results in a "if they fail, they fail fast" model, which reduces the waste of public funds. Several RIAs have awarded medium-to-large grants, but they have differed in their approach. For example, NGI Trust (€7 million) funds the development of privacy and trust-enhancing technologies, using a three-tier funding model ranging from €75,000 for viability studies to €200,000 for commercialisation, with varying matching fund requirements. Meanwhile the European Self-Sovereign Identity Framework Lab (eSSIF-lab) distributed €5.6 million to 56 projects, offering between €15,000 and €106,000 depending on project stage and an additional €155,000 for developing open source self-sovereign identity components. In addition to funding, all NGI0 projects, as well as the NGI0 Review Coordination and Support Action (CSA), have conducted security audits since 2019 on the entire NGI portfolio as a systematic measure to improve security and resilience.

A survey of 291 NGI-funded projects by Gartner Consulting offers insights on the impact of the NGI initiative [56]. 63% of projects have shared their solution in a public OSS repository, and 41% said their project was part of a larger OSS community effort. 76% of the funded projects have external communities (other than the core developers and maintainers) contributing to their projects, with 40% reporting small communities (fewer than 10 people) and 15% reporting large communities (over 50 people). Based on the survey results, the authors estimate that each NGI-funded contributor has supported a community of roughly 50 contributors, resulting in a 1:50 multiplier effect, and around 80,000 individuals actively contribute to NGI-funded OSS projects through code contributions, testing, and bug reporting, among others. However, this multiplier effect estimation follows a pragmatically simple calculation.<sup>8</sup>

Despite these promising indicators, many challenges hinder the ability to assess the impacts of the NGI initiative systematically or at scale. First, the aforementioned estimates are based on a survey of a third of funded projects, leaving us with a partial understanding of the impacts of NGI funding. Furthermore, the survey focused on qualitative impacts (e.g. relevance to EU legislation and participation standards bodies) and lacks data required for quantitative analysis (e.g. funding amounts, funding dates, and repository URLs). Beyond this survey, key methodological obstacles include the heterogeneity of funding models and projects, as well as limited data availability. For instance, the NGI initiative does not maintain a registry of project-level funding data, such as amounts, dates, and objectives, creating a significant data gap which is difficult to resolve due to the decentralised nature of NGI's funding approach and a lack of systematic record-keeping across NGI RIAs. Many grants are given to individuals, which creates difficulties in publishing names and funding due to the General Data Protection Regulation. One NGI coordinator described obtaining precise funding data as difficult as "looking for the hen with the golden eggs."<sup>9</sup>

<sup>6</sup>N.B. While the authors are funded by the EU's Horizon Europe programme under grant agreement number 101135279 (NGI Commons), the EU does not endorse the research outputs of the consortium.

<sup>7</sup>Cascade funding is also referred to as Financial Support for Third Parties (FSTP).

<sup>8</sup>The report simply notes "The multiplier effect regarding the community takes into account an estimation provided by DG CONNECT of the number of grantees per project: 1.5."

<sup>9</sup>N.B. They used this expression to convey the difficulty of obtaining funding data, rather than to convey greed, as per the idiom.

### 2.2.2 Sovereign Tech Agency

The Sovereign Tech Fund (STF) was established in 2022 by the German Federal Ministry for Economic Affairs and Climate Action (BMWK) and hosted by SPRIND GmbH, the Federal Agency for Disruptive Innovation, to strengthen the resilience, security, and sustainability of critical open-source digital infrastructure. The STF provides OSS projects with contracts for maintenance and development work starting from €50,000, with no upper funding limit, and supports project durations ranging from 6 to 24 months. In November 2024, the STF was reorganised under a newly formed umbrella entity: the Sovereign Tech Agency (STA). This transition allowed the STA to expand its scope beyond the STF’s original focus, establishing a broader array of initiatives and programs. The STF continues as a core component within the STA, which now operates with an annual budget of approximately €19 million.

Since its inception, the STF has provided €23.5 million in funding to OSS developers and projects for various tasks, including code contributions, security improvements, and infrastructure overhauls. In its pilot phase, the STF allocated around €1.3 million to set up the programme and funded nine projects. The pilot included OpenMLS, curl, WireGuard, OpenSSH, Bundler/RubyGems, OpenBGPd, Sequoia-PGP, Fortran Package Manager, and OpenPGP.js/openPGP. A notable example is the investment in the PHP ecosystem, with €205,000 allocated between 2023 and 2024 to support the PHP Foundation. This funding is aimed at addressing outdated infrastructure, enhancing security measures, and improving documentation. Given PHP’s foundational role in web development—powering approximately 75% of websites globally—these efforts were critical for ensuring the sustainability and security of a widely used programming language. Another key investment was in Log4j, a widely-used Java logging library, which received €596,160 in 2023 to improve its security and stability following the critical vulnerabilities exposed during the Log4Shell incident in 2021. This investment supported key enhancements such as fuzz testing, improvements to the project’s release pipeline, and the development of a Software Bill of Materials (SBOM) to better track dependencies and security risks.

In addition to the project-based funding of the STF, the STA offers a bug resilience programme (Sovereign Tech Resilience), which was launched in early 2023 to enhance the security and resilience of OSS. The resilience programme focuses on reducing technical debt, conducting code security audits, and managing a bug bounty programme to address undiscovered software vulnerabilities proactively. With a budget of €2.5 million over two years, the programme is designed to offer a holistic approach to vulnerability management, ensuring that OSS projects are better prepared to prevent and mitigate security risks. The Sovereign Tech Agency is also piloting a fellowship programme (Sovereign Tech Fellowship) aimed at directly supporting the maintainers of critical OSS projects. This programme aims to pay for the typical work of maintainers up to €78,000 per year per individual. Fellows work across multiple technologies and perform essential tasks such as technical reviews, community management, and security triage. The fellowship provides both freelance contracting options and a full-time “maintainer-in-residence” position, offering maintainers the financial stability to continue their contributions. By investing in the economic stability and security of individuals behind the code, the Fellowships seeks to address structural issues within the OSS community at large, mitigate maintainer burnout, and ensure the long-term sustainability of critical OSS projects.

## 3 A Toolkit for Measuring the Impacts of Public Funding on OSS Development

It is difficult to measure the impacts of funding on OSS development due a number of reasons, including the heterogeneity of OSS projects and ecosystems, the heterogeneity of funding approaches and objectives, and a range of methodological challenges. As a guide through this complexity, we provide a toolkit that may help researchers and practitioners conceptualise and empirically measure the multifaceted impacts of public funding on OSS development. It is informed by methodologies for OSS valuation and community health measurement as well as our first-hand experiences as practitioners tasked with evaluating the impacts of public funding, as discussed in Section 2. It is guided by the following motto: If we can meaningfully quantify the impacts of public funding on OSS development, then we can inform funding strategies that are beneficial both to OSS developers and society at large. We underscore that this toolkit is neither prescriptive, nor exhaustive; rather, it is a systematic organisation of key learnings and considerations derived from prior scholarship and our first-hand insights at the NGI initiative, the STA, and CHAOSS project.

### 3.1 Start with Contextual Considerations

#### 3.1.1 Start with the Funding Objectives

To state the obvious: not all funding is the same. Different funding instruments—whether milestone-based contracts with deliverables, bug bounty programmes, or general project support—have varying purposes, from innovation to maintenance, and as such varying impacts. Furthermore, one should separate funding at the project and ecosystem levels, as well as funding for project-specific goals and systemic goals, such as infrastructure and developer tools. For

example, funding aimed at software development might focus on technological outcomes, such as bug fixes or feature enhancements, while community-oriented funding may focus on contributor diversity or governance changes.

Understanding the specific objectives of the funding helps to align the impact measurements with the expected outcomes. For example, as described above, both NGI and the STA fund projects based on a set of milestones that are mutually agreed upon by the funder and the project. As a result, each project works toward different milestones with different goals and objectives, so there is no one-size-fits-all measurement to determine the impacts of these funding programmes. Therefore, one approach is to assess impacts against each project's agreed milestones. However, there is also a case to be made against such an approach, as examining impacts according to how well the work conforms to specified goals may just be a way to measure compliance rather than impact. In Section 3.2, we discuss a range of possible economic, social, and technological impacts, which can be both positive and negative, direct and indirect, and manifest over varying time horizons, which provide a broader perspective beyond funding objectives.

### 3.1.2 Consider Project Life Stages and Social Structures

The impacts of funding vary depending on both a project's life stage and its social structure. For example, a new prototype project will have very different needs and potential outcomes compared to a mature project with an established community. While the former might need funding to build initial features and attract contributors, the latter might require support for security improvements or maintenance work.

Similarly, considering the social structure of a project is crucial for understanding funding impacts. For instance, using the project types identified by Nadia Eghbal<sup>10</sup>, funding for a federation-type project might need to account for complex governance processes and multiple working groups, while funding for a stadium-type project might focus on supporting its small core of maintainers who serve a large user base. These structural differences affect not only how funding should be allocated but also how its impact should be measured.

### 3.1.3 Account for Salary Structures and Cost Factors

One should consider salary structures and cost factors across regions and organisations. When similar budgets are allocated to different organisations or regions, they can support different numbers of developers due to varying salary levels. Furthermore, even with access to salary and costs data, its relevance is problematised by the fact that remuneration and the quality of contributions do not always correlate. For example, many highly skilled and experienced OSS developers work in organisations that pay more modest salaries compared to major technology companies, meaning a senior developer in the former may earn less than a junior developer in the latter.

## 3.2 A Taxonomy of Economic, Social, and Technological Impacts

When assessing the impacts of OSS funding, there may be a tendency to focus on technological impacts, which may in part be due to the technological nature of OSS development or the relative ease of measuring technological impacts due to data availability from OSS repositories. However, as discussed in Section 2.1.2, the potential impacts of funding extend beyond the code itself, and it is crucial to consider economic and social impacts on OSS projects and their wider ecosystems of dependents and users. Furthermore, impacts can be direct or indirect, internal (i.e. affecting the project and its community of contributors) or external (i.e. affecting its ecosystem of dependents or users). It is also important to note that impacts are not linear or unidirectional; funding can lead to both improvements and degradations across different metrics. This multidirectional nature of impacts means that measurement frameworks need to be flexible to capture both positive and negative changes, rather than assuming funding inherently leads to improvements or that metrics only move in one direction. We provide illustrative examples of various impacts in Table 2, as well as a longer list of question prompts about economic, social, and technological impacts in Appendix A. These questions are not exhaustive; rather, they intend to provoke reflection by practitioners about the possible impacts of funding.

---

<sup>10</sup>As mentioned in Section 2.1.2), Nadia Eghbal identifies four distinct models of OSS projects based on their contributor and user ratios: federations, clubs, stadiums, and toys [6].



Table 2: Examples of Social, Economic, Technological Impact Areas of OSS Funding

	Internal Impacts (Project-level)	External Impacts (Ecosystem-level)
<b>Direct Impacts</b>	<ul style="list-style-type: none"> <li>• <i>Social</i>: Contributor retention, community engagement, community events, contributor diversity, work-life balance, reduced burnout, mentorship</li> <li>• <i>Economic</i>: Paid developer time / support roles, infrastructure coverage, conference sponsorship, project-related revenue via various channels</li> <li>• <i>Technological</i>: Maintainer responsiveness, commit velocity, code security, dependency management, documentation quality, consistent releases</li> </ul>	<ul style="list-style-type: none"> <li>• <i>Social</i>: User trust, ecosystem community, ecosystem events</li> <li>• <i>Economic</i>: Cost savings for adopters, integration/support costs, shared maintenance burden</li> <li>• <i>Technological</i>: Stability of APIs, ecosystem-wide security updates, interoperability</li> </ul>
<b>Indirect Impacts</b>	<ul style="list-style-type: none"> <li>• <i>Social</i>: Leadership development, governance and decision-making processes, knowledge preservation, conflict resolution/prevention mechanisms</li> <li>• <i>Economic</i>: Job market value for developers, partnership opportunities, academic collaborations, consulting opportunities, funding diversity</li> <li>• <i>Technological</i>: Standardisation, interoperability</li> </ul>	<ul style="list-style-type: none"> <li>• <i>Social</i>: Cross-project collaboration, training &amp; education resources, ecosystem engagement</li> <li>• <i>Economic</i>: Market growth, job creation, industry cost reduction, start-up creation</li> <li>• <i>Technological</i>: Standardisation, research papers, patents, ecosystem-wide security improvements</li> </ul>

### 3.3 Temporal Considerations

When measuring the impact of funding on OSS development, it is crucial to consider the effects across different time horizons. The influence of funding can manifest differently in the short, medium, and long term, and these temporal distinctions are essential for a comprehensive understanding of funding impacts.

1. **Short term (<1 year)**: These are the immediate effects of funding, which may be easier to observe and attribute to the funding received. Short-term impacts might include immediate increases in development activity, improvements in code quality, or rapid growth in community size.
2. **Medium term (1-3 years)**: Medium-term impacts may reveal how the project adapts and grows with the resources provided. This period might show the development of new features, expansion of the project's reach, or evolution of the project's governance structures.
3. **Long term (>3 years)**: Long-term impacts are the enduring effects of funding that manifest over an extended period, which can be the most challenging to measure and attribute directly to funding.

### 3.4 Methodological Approaches and Challenges for Impact Measurements

We can employ a range of qualitative, quantitative, or mixed-method approaches, which offer relative strengths and weaknesses, to measure the impacts of public funding on OSS development. We provide a summary of the strengths and weaknesses of these methodological approaches in Table 3. This review draws on prior work on empirical methodologies in software engineering research [74, 75] and community health methodologies discussed in Section 2.1.2.

#### 3.4.1 Quantitative Approaches

Quantitative methods offer a structured, scalable approach for measuring relationships between funding and outcomes in OSS projects and their ecosystems. We can differentiate between three main types of quantitative analysis. First, correlational analysis, which is currently the most common approach, helps establish broad patterns and magnitudes of relationships between funding and various outcome measures. Second, quasi-experimental methods attempt to identify causal relationships by isolating quasi-random variation in observational funding data to estimate multiplier effects. Third, experimental approaches, such as randomised controlled trials, aim to establish direct causal relationships through controlled experiments with random assignment of funding; however, such approaches have not been used in OSS funding research to date, in large part due to their impracticality in this context. While each approach offers different levels of causal inference, most existing research relies on correlational analysis, highlighting an opportunity for more rigorous causal investigation through quasi-experimental and experimental methods.

Table 3: Strengths and Weaknesses of Methodological Approaches to OSS Funding Impact Measurements

Methods & Data	Strengths	Weaknesses / Challenges	Relevance to OSS Funding
<b>Qualitative</b> <ul style="list-style-type: none"> <li>• Interviews</li> <li>• Case studies</li> <li>• Participant observation</li> <li>• Focus groups</li> <li>• Content analysis</li> </ul>	<ul style="list-style-type: none"> <li>• Rich, contextual insights into funding impacts</li> <li>• Uncovers unexpected or indirect impacts</li> <li>• Explores complex interactions between funding and project dynamics</li> <li>• Captures human stories and motivations behind funding impacts</li> </ul>	<ul style="list-style-type: none"> <li>• Limited scalability for large funding initiatives</li> <li>• Potential for researcher and participant bias</li> <li>• Time-consuming and resource-intensive</li> <li>• Difficulty in generalising findings across diverse OSS projects</li> </ul>	<ul style="list-style-type: none"> <li>• Identifying and accessing key stakeholders from OSS projects</li> <li>• Ensuring representative sampling in global, often anonymous contributor bases</li> <li>• Capturing long-term impacts within typical research timeframes</li> <li>• Distinguishing funding impacts from other influencing factors</li> </ul>
<b>Quantitative</b> <ul style="list-style-type: none"> <li>• Repository mining</li> <li>• Surveys</li> <li>• Econometric modelling</li> <li>• Time series analysis</li> <li>• Network analysis</li> </ul>	<ul style="list-style-type: none"> <li>• Scalable analysis for large numbers of funded projects</li> <li>• Identifies trends and patterns in funding impacts</li> <li>• Provides quantifiable evidence for policymakers</li> <li>• Leverages digital trace data from repositories</li> <li>• Enables comparative analysis across projects and time periods</li> </ul>	<ul style="list-style-type: none"> <li>• Oversimplifies complex social and technical dynamics</li> <li>• Identifying random variation in funding</li> <li>• Isolating funding effects from other variables</li> <li>• Data availability/quality</li> <li>• Project workflows affect metric comparability</li> <li>• Difficulty in quantifying intangible outcomes (e.g. security, innovation)</li> <li>• Risk of creating perverse incentives through metric selection/optimisation</li> </ul>	<ul style="list-style-type: none"> <li>• Lack of standardised financial data across OSS projects</li> <li>• Complexity in defining control groups</li> <li>• Accounting for the non-linear and often delayed impacts of funding</li> <li>• Measuring indirect or ecosystem-wide effects of funding</li> <li>• Balancing the need for standardised metrics with project diversity</li> </ul>
<b>Mixed-Methods</b> <ul style="list-style-type: none"> <li>• Sequential explanatory design (first quantitative, then qualitative)</li> <li>• Sequential exploratory design (first qualitative, then quantitative)</li> <li>• Concurrent triangulation (qualitative &amp; quantitative simultaneously)</li> </ul>	<ul style="list-style-type: none"> <li>• Balances depth of insight with breadth of analysis</li> <li>• Triangulates findings for increased validity</li> <li>• Adaptable to diverse OSS project contexts</li> <li>• Captures both quantifiable outcomes and complex social dynamics</li> <li>• Allows for iterative refinement of research questions</li> </ul>	<ul style="list-style-type: none"> <li>• Resource-intensive, requiring expertise in multiple methodologies</li> <li>• Complexity in integrating different data types</li> <li>• Potential for conflicting results between methods</li> <li>• Challenges in synthesising diverse data into coherent narratives</li> <li>• Risk of compromising depth or breadth in pursuit of integration</li> </ul>	<ul style="list-style-type: none"> <li>• Aligning qualitative and quantitative data collection in decentralised OSS projects or ecosystems</li> <li>• Balancing the needs of various stakeholders (e.g. developers, policymakers) in research design</li> <li>• Addressing varying timescales of funding impacts across methods</li> </ul>

A critical question to consider is which variables to include in the quantitative analysis. Rather than taking a “big net approach” to fish for impacts in the wild, we recommend taking a cautious approach to variable selection. As discussed in Section 3.1.1, we recommend beginning by accounting for funding objectives. If one seeks to analyse many OSS projects, we suggest considering high-level objectives of a funding initiative rather than the specific milestones agreed between the funder and individual projects. Critical input variables concern funding information, such as allocated funding amounts, the number of funded developers, funding period, and specific objectives (e.g. security audits, feature development, or community building), as they help you know what kinds of relationships to look for and when.

With regards to output variables, we refer to the taxonomy of potential direct and indirect, internal and external economic, social, and technological impacts in Section 3.2, as well as a list of relevant variables in Appendix B). Platforms like GitHub or GitLab provide access to digital trace data about development activity in OSS projects that can be turned into metrics, such as commits, number of contributors, and forks. Metrics like forks have also been shown to be reliable proxies for usage [59]. Additional data about OSS usage or investment, such as downloads, patents, publications, and dependent projects, may be tracked through third-party sources. Furthermore, variables about various technical and non-technical contributors in a project or its ecosystem [60, 64] and community governance, diversity, and health [61, 63] can provide a more comprehensive picture of social dynamics in OSS projects.

While quantitative methods offer scalable approaches, they are fraught with methodological challenges and limitations. The first challenge concerns data availability. OSS repositories do not publicly report detailed information about funding sources or amounts, making it difficult to assess the exact relationship between funding inputs and project outputs by relying simply on repository data. Beyond funding data, data about OSS usage is infamously scarce [59], limiting analysis of impact by adoption. Second, a major challenge stems from the non-random nature of funding and the difficulty to isolate funding effects from other influences on project activity (i.e. causal attribution). While an increase in the number of commits or contributors may be observed during or after a funding period, it is difficult to attribute these changes directly to the funding itself. The non-randomness of funding complicates efforts to create clear causal narratives from quantitative data alone, highlighting the need for mixed-method approaches that can provide both depth and scale in measuring the impacts of funding on OSS projects. Furthermore, significant confounding factors limit our ability to make causal claims about funding and specific outcomes. For example, variables such as the maturity of the project, its pre-existing level of community or contributor activity, or external factors such as varying socio-economic contexts of contributors or market dynamics can all affect project growth and development independent of the funding received or the work contracted. Without accounting for these confounding variables, statistical models run the risk of overestimating or underestimating the impacts of funding.

Additional challenges stem from the different collaboration tools of platforms (“platform affordances”) and development practices across OSS projects, from commit strategies (e.g. micro-commits versus squashed-commits) and code review processes (e.g. direct merge to main branch or code review) to issue management approaches. The same metrics may reflect different underlying activities across projects, complicating cross-project comparisons and large-scale analysis of funding impacts. This necessitates the establishment of project-specific baselines against which to measure changes, rather than comparing absolute metrics across projects. What is more, the measurement of intangible outcomes is difficult. These outcomes are often not captured directly in public data but may instead be inferred from proxies, such as the number of security vulnerabilities fixed, the diversity of contributors, or the creation of new OSS projects. Quantifying these outcomes requires careful selection of metrics and can lead to reliance on external pre-defined measures (e.g. OpenSSF Scorecards). However, these aggregate measures come with the risk of oversimplification, making it challenging to interpret what is being measured (i.e. the estimand of a statistical analysis).

### 3.4.2 Qualitative Approaches

Qualitative methods, such as case studies, semi-structured interviews, and participant observation, are useful strategies for collecting contextual data about the history, role, and impacts of funding on OSS projects. A major strength of interviews lies in their ability to provide rich, contextual understandings of a phenomenon by gaining insights into the perspective of the funding recipients, revealing why funding was needed and how it has influenced project dynamics, individual motivations, and community structures [76, 75, 77]. Furthermore, interviews can uncover the specific ways in which funding has altered a project’s trajectory, such as enabling full-time work on previously volunteer-driven projects or facilitating mentorship programmes that were previously out of reach due to limited funds or capacity. One may also learn about unforeseen consequences, such as changes in community dynamics or project governance. Moreover, qualitative methods are useful for capturing intangible outcomes and hard-to-quantify impacts, such as changes in project culture or enhancements in the overall sustainability of the project, which often play a crucial role in the long-term success of OSS projects but can be challenging to measure through quantitative means alone.

However, qualitative methods face significant limitations for measuring OSS funding impacts. One major challenge is stakeholder identification and access. The decentralised nature of many OSS projects can make it difficult to identify

and engage key stakeholders for interviews or observations. This is particularly true for projects with large, globally distributed contributor bases, where key individuals may be geographically dispersed or prefer to remain anonymous. Ensuring representativeness is another challenge. It can be difficult to ensure a representative sample in qualitative studies. This limitation can potentially skew findings and make it challenging to draw broader conclusions about funding impacts. Scalability presents another limitation. The time-intensive nature of qualitative research restricts its applicability to large-scale funding initiatives with numerous beneficiaries, such as the NGI initiative ( $n > 1200$ ). This constraint makes it challenging to draw broad conclusions about funding impacts across diverse projects. Finally, temporal limitations also pose a significant challenge. Capturing long-term impacts of funding may be difficult within typical research time frames, particularly for projects where the effects of funding may take years to fully manifest. This limitation can result in an incomplete picture of funding impacts, potentially missing important long-term effects.

### 3.4.3 Mixed-Methods Approaches

Mixed-method approaches that combine qualitative and quantitative techniques, in theory, offer the best of both worlds: scalability and contextual depth. Mixed-methods approaches are particularly useful for examining complex OSS ecosystems, where development spans multiple channels and goes beyond repositories. For example, Casari et al. (2023) recommend that researchers collaborate with OSS communities to better contextualise quantitative data [78]. Below we discuss three different approaches to sequencing the mixed-methods analysis of OSS funding impacts.

#### *Sequential explanatory design (first quantitative, then qualitative)*

The sequential explanatory design begins with quantitative analysis, followed by qualitative methods to explore the relationship between the funding and these patterns in detail. Such an approach may be particularly useful if one seeks to understand the impact of funding for a large number of projects; for example, it could be applied to assess the impact of the NGI initiative's funding for over 1,200 R&D projects, many of which contributed to OSS projects and/or released new software in open source repositories. Given the sample size, we require a scalable and systematic approach for collecting and analysing data about these projects, which a quantitative approach is well-suited for.

However, this approach faces challenges. First, we would need to gather repository URLs for NGI-funded projects, and identify which platforms they are hosted on and available options for programmatic data collection (e.g. via APIs). One could use existing survey data, such as the Gartner Consulting survey ( $n = 291$ ), then manually fill data holes via internet searches for these repositories. An alternative approach would be to distribute a new survey to all NGI-funded projects, which could collect additional missing data, such as funding amounts, objectives, time period, and project maturity, but it would require coordination with decentralised RIA coordinators and the responsiveness of projects is not guaranteed. Regardless of the approach taken, once we have repository URLs and funding information for some percentage of the projects, we must select the variables for the analysis. Given the diversity of projects in the NGI portfolio, it could be sensible to divide projects into buckets based on some criteria and to select variables per bucket. For example, given the range of objectives and project life stages, from proof-of-concepts to feature development in mature projects, one could divide projects into maturity-based buckets. Then, once one has collected data, one may conduct time-series analysis or econometric modelling to identify general trends of changes to activity across all or perhaps among clusters of projects. This analysis may also highlight outliers or unexpected patterns.

However, given the heterogeneity of the more than 1,200 NGI-funded projects, one should be cautious about using a broad-stroke, quantitative analysis to the many projects, which are diverse by their technology domain, funding approach and objectives, and project life stage, among others. This is where a qualitative analysis of a smaller sample of projects would be useful; for example, by conducting qualitative case studies on a smaller sample of projects (e.g.  $n = 10$ ). Through the analysis of publicly available information (e.g. documentation, blogs, mailing lists, etc) and interviews with the funding recipients, one can gain a more nuanced understanding of how funding was used by and impacted a project. This qualitative phase also allows us to test the usefulness of our quantitative findings or lack thereof, and uncover impacts that were not captured by the quantitative analysis.

#### *Sequential exploratory design (first qualitative, then quantitative)*

Sequential exploratory design begins with qualitative analysis, followed by quantitative analysis. Consider a scenario where we want to evaluate the impacts of the STA's pilot phase funding on a smaller number of OSS projects ( $n = 9$ ). We may begin with a thorough review of funding agreements for all projects in the STA pilot phase, providing a contextual understanding of what the funding was intended to achieve. Next, we conduct semi-structured, exploratory interviews with key stakeholders in each funded project. These interviews allow us to gather rich, contextual data

about the funding experience. Questions might include: “How did you utilise X amount of funding to achieve Y objective?”, “What specific activities or outputs were facilitated by this funding?”, “Are there digital traces of these activities that could be programmatically scraped and quantitatively analysed, and if so on which platforms?”, and “In your opinion, what variables should we control for in our analysis?”

This qualitative phase offers several advantages. It allows for the discovery of unexpected impacts that may not have been captured in quantitative metrics. It also provides insight into the human experiences and decision-making processes behind the numbers. However, it is time-intensive and may be subject to individual biases or selective recall. Using these qualitative insights, we then inform our subsequent quantitative analysis. We can identify relevant data sources, mine data from repositories with a focus on areas highlighted in interviews, and conduct targeted quantitative analyses that account for project-specific contexts and potential confounding variables. This approach ensures that our quantitative analyses are grounded in the realities and nuances of each project’s funding objective and experience. It allows for a more targeted and meaningful quantitative analysis, potentially uncovering patterns or relationships that might have been overlooked in a purely quantitative approach. However, this method may be less suitable for large-scale evaluations due to the time-intensive nature of the qualitative phase. It also requires careful consideration to ensure that insights from the qualitative phase are appropriately translated into quantitative measures.

### *Concurrent triangulation approaches (e.g. maturity models)*

Maturity models are a concurrent triangulation approach that combine qualitative assessments with quantitative metrics at the same time to create a holistic view of the development of an OSS project [79, 80]. A maturity model approach might involve defining relevant dimensions of project maturity (e.g. code quality, community size, documentation completeness), establishing qualitative criteria and quantitative metrics for each dimension, collecting data through surveys, interviews, and automated repository analysis, and analysing changes in maturity levels in relation to funding inputs. For example, maturity models have been proven to be an effective approach to understand the health of young and data-scarce OSS ecosystem health [81]. This approach offers a standardised yet flexible method for comparing diverse projects and allows for tracking changes over time, potentially correlating these with funding inputs. However, developing a comprehensive and relevant maturity model is time-consuming, and there may be disagreement among stakeholders about what constitutes “maturity” in different dimensions.

## **3.5 Options for and Hazards of Multiplier Effect Estimations**

With this knowledge, let us now consider how we might estimate multiplier effects of OSS funding as one form of quantitative evidence of funding impacts.

### **3.5.1 Conceptualising Multiplier Effects of OSS Funding**

In the context of R&D, the economic multiplier is the measurement of how much additional economic value is generated for each euro spent on R&D [82]. It is also relevant to consider knowledge spillovers (i.e. when R&D conducted by one entity creates value for other entities without compensation to the innovator) and social rates of return (i.e. the total benefits to society from R&D investment, including both private returns to the innovator and spillover benefits to others). Multiplier effects are often measured through econometric analyses that quantify the relationship between an investment and direct and indirect outcomes over time [83, 84]. Relevant variables include input measures (e.g. expenditure on salaries and equipment), output measures (e.g. patents, publications, or products), time lags (i.e. accounting for delays between investment and observable outcomes), and spillover effects (i.e. impacts on related industries).

The multiplier effect of OSS funding may be understood as the generation of value that exceeds the initial funding amount or objective. This may include community contributions, technological innovations, community or ecosystem growth, or downstream economic benefits. This calculation requires the adaptation of methodologies to the collaborative nature of OSS development, where value creation and capture operate differently than in the typical R&D processes and structures found in labs. Specifically, contributors have the freedom to join or leave projects at will, development paths can fork into various directions, and improvements made by one contributor are instantly accessible to everyone. These dynamics can amplify positive multiplier effects through increased collaboration and knowledge sharing, but they can also lead to negative multiplier effects, such as by creating unsustainable dependencies.

### **3.5.2 Measuring Multiplier Effects of OSS Funding**

There is a serious risk of oversimplifying complex dynamics by reducing them to single metrics, and as such it is important to measure multiplier effects with caution and with an awareness of the limitations of the data and methods

used in such measurements. Due to the quantitative nature of such measurements, the concerns are similar to those elaborated in Section 3.4.1. Above all, it is challenging to account for the non-random nature of funding and to isolate funding effects from other influences on project activity, as projects that receive funding often differ systematically from those that do not in ways that affect their outcomes. For instance, more mature or popular projects might be more likely to receive funding and also more likely to grow independently of funding. This selection bias makes it difficult to establish causal relationships between funding and observed outcomes. What is more, it is difficult to quantify intangible benefits, and time lags between funding periods and observable impacts can exceed typical measurement periods, making it difficult to capture the full scope of impacts. There are also important normative or behavioural considerations: the choice of multiplier effect variables itself might create perverse incentives, leading projects to optimise for measured outcomes at the expense of factors that are important to the project and its community.

For these reasons, measuring multiplier effects of OSS funding requires careful consideration of the following factors:

- The bidirectional nature of funding impacts, including both positive and negative multiplier effects
- The complex network effects and externalities in OSS ecosystems
- The role of community dynamics and governance in mediating funding impacts
- The potential for funding to affect project independence and sustainability
- The distribution of benefits across different stakeholder groups (e.g. developers, dependents, and users)

Given these considerations, estimating the multiplier effect of OSS funding should be approached with care and ideally combine quantitative and qualitative metrics that provide a comprehensive, contextual understanding of funding in OSS projects. To aid such measurements, we provide a list of relevant input, output, and control variables in Appendix B. Input variables include the funding amount and the number of funded developers; while output variables may be technological (e.g. maintainer responsiveness or growth in dependent projects), economic (e.g. project revenue or cost savings for adopters), or social (e.g. contributor diversity or community events). As noted in Section 3.2, these economic, social, and technological outputs can be direct or indirect as well as internal or external to the project. Relevant control variables include project life stage, its pre-existing community size, and concurrent market trends.

## 4 Discussion

### 4.1 Key Components of the Toolkit

The provision of public funding is increasingly considered important as a lever for supporting the maintenance, security, and long-term sustainability of OSS projects that form a critical part of our digital infrastructure, yet we still have a limited understanding of the impact of funding and a lack of consensus on how to meaningfully assess or measure impact. We have sought to address this gap by reviewing prior scholarship and documenting our first-hand learnings in the form of a practical toolkit that may inform attempts to measure the impacts of public funding on OSS development.

Core to this discussion is the following question: How do we capture the myriad ways in which public funding can shape—positively or negatively, directly or indirectly—the development, sustainability, and evolution of OSS projects and their wider ecosystems? While the diversity of OSS projects and ecosystems complicates efforts to develop standardised approaches to impact measurement, our toolkit seeks to offer some practical ways forward. For example, it is important to start with the objectives of the funding and consider OSS projects as socio-technical systems. If one is the funder, one should have a record of the funding objectives, and one should design one’s impact measurement with those objectives in mind. It is important to look beyond the code; that is, to consider the various social and economic impacts of funding OSS development. The toolkit provides guidance on the various potential impacts (see Section 3.2) as well as methodological options for and hazards of estimating multiplier effects of OSS funding see Section 3.5).

We discussed the merits and limitations of qualitative and quantitative methods. While quantitative measurements of ROI or multiplier effects may be desired by decision-makers to justify public spending and optimise funding strategies, we caution against hasty calculations that do not consider the methodological challenges and hazards involved in quantitative measurements. Qualitative methods, in turn, may yield rich and contextual insights into the role and impact of funding within OSS projects, but do not scale and do not produce desired quantitative evidence. Mixed-method strategies may balance these trade-offs, though our examples highlighted that they face challenges, too.

## 4.2 Broader Questions and Challenges

We note that measuring OSS project impacts through metrics carries inherent risks and normative implications. Indicators often become interventions themselves, potentially steering project development in unintended directions [85].<sup>11</sup> The diverse values and priorities of OSS developers must be carefully translated into meaningful metrics for funders. While multiplier effects provide valuable insights, they risk oversimplifying complexity and creating misaligned incentives. A balanced approach combining qualitative and quantitative measurements may help preserve nuance while providing necessary accountability.

The impacts of funding cessation also warrant careful consideration. Research suggests that funded projects often face sustainability challenges post-funding, as evidenced by the Prototype Fund’s evaluation finding that over half of their funded projects struggled to secure follow-up funding [55]. This raises critical questions about project sustainability: Does initial funding create dependencies that make projects more vulnerable when funding ends? How can we better prepare projects for post-funding sustainability? More research is needed to understand these transition periods and develop strategies to support projects through funding changes.

A further question to consider is how responsive this toolkit is to various funders in the public sector and beyond, from companies to philanthropic foundations. While the objectives of a company may differ from a government (e.g. focus on specific business needs), the toolkit’s recommendations to account for funding objectives, cost factors, and various economic, social, and technological impacts should remain relevant and responsive to companies. We seek to test this hypothesis with diverse funders through future workshops, research, and iterations of this toolkit.

## 4.3 Future Work

Current understanding of OSS funding impacts remains incomplete in several key areas. Significant knowledge gaps exist in our understanding of long-term effects, the effectiveness of different funding models, and potential unintended consequences, such as impacts on volunteer motivations or creating dependencies. These gaps are particularly concerning given the importance of OSS as digital infrastructure and the challenges projects face, including maintainer burnout and security vulnerabilities. Further research is needed to examine the long-term sustainability of funded projects, compare funding models, and develop adaptable metrics for assessing impact across diverse contexts. Understanding how funding influences the broader OSS ecosystem, including its effects on innovation and collaboration, is also crucial. Addressing these gaps requires collaboration among the OSS developer community, funders, policy-makers, and researchers to develop a deeper understanding of the multifaceted impacts of various funding approaches. Such efforts are vital to ensure that funding initiatives support, rather than undermine, OSS developers.

## 5 Conclusion

This paper presented a toolkit for measuring the impacts of public funding on OSS development. The toolkit is not prescriptive, nor is it exhaustive; rather, we provide practical recommendations and discuss potential hazards for funding impact measurements. It starts with a discussion of key contextual considerations to inform impact measurements, such as accounting for funding objectives, project life stages and social structures, and regional and organisational cost factors. Then, we present a taxonomy of social, economic, and technological impacts, which can be both positive and negative, direct and indirect, internal (i.e. within a project) and external (i.e. among a project’s ecosystem of dependents and users), and manifest over various time horizons. Then, we provide guidance for the use of quantitative, qualitative, and mixed-methods approaches to measure such impacts, as well as the measurement of multiplier effects of funding. While multiplier effect metrics are useful for decision-makers who require quantitative evidence about returns on investment to justify public spending and optimise funding strategies, we highlight a number of methodological challenges, from data availability to the non-randomness of funding, that complicate making meaningful quantitative estimates. This toolkit is not exhaustive and we still have major gaps in our understanding of the impacts of funding and optimal impact measurements. We, therefore, call for diverse stakeholders—OSS contributors (technical and non-technical), policymakers, funders, researchers, and others—to join this effort to develop frameworks for meaningful, evidence-based evaluations of the multifaceted impacts of public funding on OSS development.

---

<sup>11</sup>We recommend reading “Indicators as Interventions: Pitfalls and Prospects in Supporting Development Initiatives” by Davis and Kingsbury (2012), especially the “Indicators: Pathologies and Pitfalls” table on page 3, for a thorough discussion of the normative implications and potential risks of creating quantitative indicators [85].

## References

- [1] OSI, “The Open Source Definition (v1.9),” Mar. 2007.
- [2] N. Eghbal, “Roads and Bridges: The Unseen Labor Behind Our Digital Infrastructure,” tech. rep., Ford Foundation, June 2016.
- [3] S. Scott, S. A. Brackett, T. Herr, and M. Hamin, “Avoiding the Success Trap: Toward Policy for Open-Source Software as Infrastructure,” tech. rep., Atlantic Council, Washington D.C., USA, Feb. 2023.
- [4] Synopsys, “Open Source Security and Analysis Report,” tech. rep., Mountain View, CA, USA, 2023.
- [5] O. S. S. Foundation, “OSS Mobilization Plan,” tech. rep., Open Source Security Foundation, Linux Foundation, Washington D.C., USA, May 2022.
- [6] N. Eghbal, *Working in Public: The Making and Maintenance of Open Source Software*. San Francisco: Stripe Press, Aug. 2020.
- [7] B. Birkinbine, *Incorporating the Digital Commons: Corporate Involvement in Free and Open Source Software*. University of Westminster Press, Feb. 2020.
- [8] S. J. Vaughan-Nichols, “Log4Shell: We Are in So Much Trouble,” Dec. 2021.
- [9] S. Herpig, “Fostering Open Source Software Security – Blueprint for a Government Cybersecurity Open Source Program Office,” tech. rep., Stiftung Neue Verantwortung, Berlin, Germany, May 2023.
- [10] P. Keller, “European Public Digital Infrastructure Fund White Paper,” *Open Future*, Dec. 2022. Publisher: Open Future Foundation.
- [11] C. Osborne, M. Boehm, and A. Jimenez Santamaria, “The European Public Sector Open Source Opportunity: Challenges and Recommendations for Europe’s Open Source Future,” tech. rep., The Linux Foundation, Brussels, Belgium, Sept. 2023.
- [12] F. Burwell and K. Propp, “Digital sovereignty in practice: The EU’s push to shape the new global economy,” tech. rep., Nov. 2022.
- [13] F. Nagle, “Government Technology Policy, Social Value, and National Competitiveness,” Mar. 2019.
- [14] C. Osborne, “Public-private funding models in open source software development: A case study on scikit-learn,” May 2024. arXiv:2404.06484 [cs].
- [15] N. Bloom, J. Van Reenen, and H. Williams, “A Toolkit of Policies to Promote Innovation,” *Journal of Economic Perspectives*, vol. 33, pp. 163–184, Aug. 2019.
- [16] M. Böhm, “Economics of Open Source,” in *Open Source Law, Policy and Practice* (A. Brock, ed.), p. 0, Oxford University Press, Oct. 2022.
- [17] M. Gerosa, I. Wiese, B. Trinkenreich, G. Link, G. Robles, C. Treude, I. Steinmacher, and A. Sarma, “The Shifting Sands of Motivation: Revisiting What Drives Contributors in Open Source,” in *Proceedings of the 43rd International Conference on Software Engineering, ICSE ’21*, (Madrid, Spain), pp. 1046–1058, IEEE Press, Nov. 2021.
- [18] M. Dulong de Rosnay and F. Stalder, “Digital commons,” *Internet Policy Review*, vol. 9, p. 15 p., Dec. 2020. Publisher: Alexander von Humboldt Institute for Internet and Society.
- [19] E. S. Raymond, *The Cathedral and the Bazaar: Musings on Linux and Open Source by an Accidental Revolutionary*. Sebastopol: O’Reilly Media, Incorporated, 2001.
- [20] C. Loebbecke and A. A. Angehrn, “Open Source Platforms Under Co-opetition: A Comparative Analysis of SourceForge and ‘CodeX’ (Xerox) as Two ‘Co-opetitive Learning and Knowledge Exchange Networks’ (CoLKENS),” in *ECIS 2003 Proceedings*, 2003.
- [21] L. Markus, B. Manville, and C. Agres, “What Makes a Virtual Organization Work: Lessons From the Open-Source World,” *MIT Sloan Management Review*, Oct. 2000.
- [22] C. M. Kelty, *Two bits: the cultural significance of free software*. Experimental futures, Durham, N.C. ; London: Duke University Press, 2008.
- [23] N. Eghbal, “Where money meets open source,” 2017.
- [24] Y. Zhang, M. Qin, K.-J. Stol, M. Zhou, and H. Liu, “How Are Paid and Volunteer Open Source Developers Different? A Study of the Rust Project,” in *Proceedings of the IEEE/ACM 46th International Conference on Software Engineering, ICSE ’24*, (New York, NY, USA), pp. 1–13, Association for Computing Machinery, Apr. 2024.



- [25] R. S. Geiger, D. Howard, and L. Irani, “The Labor of Maintaining and Scaling Free and Open-Source Software Projects,” *Proceedings of the ACM on Human-Computer Interaction*, vol. 5, pp. 175:1–175:28, Apr. 2021.
- [26] A. Salkever, “Open Source Maintainers: Exploring the people, practices, and constraints facing the world’s most critical open source software projects,” tech. rep., Linux Foundation, July 2023.
- [27] GitHub, “Sponsoring an open source contributor,” 2023.
- [28] X. Zhang, T. Wang, Y. Yu, Q. Zeng, Z. Li, and H. Wang, “Who, What, Why and How? Towards the Monetary Incentive in Crowd Collaboration: A Case Study of Github’s Sponsor Mechanism,” *arXiv:2111.13323 [cs]*, Nov. 2021. arXiv: 2111.13323.
- [29] C. Overney, J. Meinicke, C. Kästner, and B. Vasilescu, “How to not get rich: an empirical study of donations in open source,” in *Proceedings of the ACM/IEEE 42nd International Conference on Software Engineering, ICSE ’20*, (New York, NY, USA), pp. 1209–1221, Association for Computing Machinery, Oct. 2020.
- [30] K. Nakasai, H. Hata, S. Onoue, and K. Matsumoto, “Analysis of Donations in the Eclipse Project,” in *2017 8th International Workshop on Empirical Software Engineering in Practice (IWESEP)*, pp. 18–22, Mar. 2017.
- [31] K. Nakasai, H. Hata, and K. Matsumoto, “Are Donation Badges Appealing?: A Case Study of Developer Responses to Eclipse Bug Reports,” *IEEE Software*, vol. 36, pp. 22–27, May 2019. Conference Name: IEEE Software.
- [32] P. K. Medappa, M. M. Tunc, and X. Li, “Sponsorship Funding in Open-Source Software: Effort Reallocation and Spillover Effects in Knowledge-Sharing Ecosystems,” June 2023.
- [33] S. Butler, J. Gamalielsson, B. Lundell, P. Jonsson, J. Sjöberg, A. Mattsson, N. Ricko, T. Gustavsson, J. Feist, S. Landemoo, and E. Lonroth, “An Investigation of Work Practices Used by Companies Making Contributions to Established OSS Projects,” in *2018 IEEE/ACM 40th International Conference on Software Engineering: Software Engineering in Practice Track (ICSE-SEIP)*, pp. 201–210, ACM, 2018.
- [34] L. Dahlander and M. W. Wallin, “A man on the inside: Unlocking communities as complementary assets,” *Research Policy*, vol. 35, pp. 1243–1259, Oct. 2006.
- [35] X. Xia, W. Wang, S. Zhao, S. Bian, and R. Wang, “Lessons Learned From the Ant Group Open Source Program Office,” *Computer*, vol. 56, pp. 92–97, Apr. 2023. Conference Name: Computer.
- [36] S. O’Mahony and B. A. Bechky, “Boundary Organizations: Enabling Collaboration among Unexpected Allies,” *Administrative Science Quarterly*, vol. 53, pp. 422–459, Sept. 2008. Publisher: SAGE Publications Inc.
- [37] Tidelift, “2020: The managed open source survey,” tech. rep., Tidelift, 2020.
- [38] L. Hawthorn, “Google Summer of Code,” *The Open Source Business Resource*, pp. 5–10, Dec. 2008. Num Pages: 6 Place: Ottawa, Canada Publisher: Talent First Network.
- [39] D. O’Brien, “The FOSS Contributor Fund: Forming a Community of Adopters,” Nov. 2019.
- [40] D. O’Brien, “Sustaining FOSS Projects By Democratizing The Sponsorship Process,” 2019.
- [41] J. West and S. O’Mahony, “Contrasting Community Building in Sponsored and Community Founded Open Source Projects,” in *Proceedings of the 38th Annual Hawaii International Conference on System Sciences*, (Big Island, HI, USA), pp. 196c–196c, IEEE, 2005.
- [42] C. Osborne, F. Daneshyan, R. He, H. Ye, Y. Zhang, and M. Zhou, “Characterising Open Source Co-opetition in Company-hosted Open Source Software Projects: The Cases of PyTorch, TensorFlow, and Transformers,” Oct. 2024. arXiv:2410.18241 [cs].
- [43] K. Blind, S. Pättsch, S. Muto, M. Böhm, T. Schubert, P. Grzegorzewska, and A. Katz, “The impact of open source software and hardware on technological independence, competitiveness and innovation in the EU economy,” tech. rep., European Commission, Brussels, Belgium, 2021.
- [44] G. Korkmaz, J. B. Santiago Calderón, B. L. Kramer, L. Guci, and C. A. Robbins, “From GitHub to GDP: A framework for measuring open source software innovation,” *Research Policy*, vol. 53, p. 104954, Apr. 2024.
- [45] N. L. Wright, F. Nagle, and S. Greenstein, “Open source software and global entrepreneurship,” *Research Policy*, vol. 52, p. 104846, Nov. 2023.
- [46] M. Hoffmann, F. Nagle, and Y. Zhou, “The Value of Open Source Software,” 2024.
- [47] K. Champion and B. M. Hill, “Underproduction: An Approach for Measuring Risk in Open Source Software,” in *2021 IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER)*, pp. 388–399, Mar. 2021. ISSN: 1534-5351.

- [48] X. Li, Y. Zhang, C. Osborne, M. Zhou, Z. Jin, and H. Liu, “Systematic Literature Review of Commercial Participation in Open Source Software,” *ACM Transactions on Software Engineering and Methodology*, Aug. 2024. Just Accepted.
- [49] M. Germonprez, G. J. Link, K. Lombard, and S. Goggins, “Eight Observations and 24 Research Questions About Open Source Projects: Illuminating New Realities,” *Proceedings of the ACM on Human-Computer Interaction*, vol. 2, pp. 57:1–57:22, Nov. 2018.
- [50] C. Osborne, “Why Companies Democratise Artificial Intelligence: The Case of Open Source Software Donations,” *arXiv preprint arXiv:2409.17876*, 2024.
- [51] B. Guerry, “code.gouv.fr presenting the BlueHats awards,” Sept. 2024.
- [52] R. Ellis and J. Bollampalli, “Opportunities, Risks, and a Path Forward,” tech. rep., Sovereign Tech Fund, Leipzig, Germany, Oct. 2024.
- [53] E. Lostri, G. Wood, and M. Jain, “Government Open Source Software Policies | Resources,” 2022.
- [54] C. Strasser, K. Hertweck, J. Greenberg, D. Taraborelli, and E. Vu, “Ten simple rules for funding scientific open source software,” *PLOS Computational Biology*, vol. 18, p. e1010627, Nov. 2022. Publisher: Public Library of Science.
- [55] N. Galati, L. Rabe, D. Obeth, and J. Biela, “Evaluation der Fördermaßnahme „Software-Sprint“ - Evaluationsbericht,” tech. rep., Bundesministerium für Bildung und Forschung, Berlin, Germany, Sept. 2023.
- [56] C. a. T. European Commission: Directorate-General for Communications Networks and C. Valayer, “Benchmarking the impact of the next generation internet initiative – Final study report,” tech. rep., Publications Office of the European Union, Brussels, Belgium, June 2024.
- [57] J. S. Meyers, S. A. Brackett, and S. Scott, “O\$\$ Security: Does More Money for Open Source Software Mean Better Security? A Proof of Concept,” tech. rep., Atlantic Council, Washington D.C., USA, Apr. 2024.
- [58] S. Vargas, “FOSSY 2024 | Presentation: A review of valuation models and their application to open source models,” Aug. 2024.
- [59] S. Vargas, G. J. P. Link, and J. Lee, “Estimating Usage Of Open Source Projects,” pp. 652–653, *IEEE Computer Society*, Apr. 2024.
- [60] J.-G. Young, A. Casari, K. McLaughlin, M. Z. Trujillo, L. Hébert-Dufresne, and J. P. Bagrow, “Which contributions count? Analysis of attribution in open source,” in *2021 IEEE/ACM 18th International Conference on Mining Software Repositories (MSR)*, pp. 242–253, May 2021. ISSN: 2574-3864.
- [61] S. Goggins, K. Lombard, and M. Germonprez, “Open Source Community Health: Analytical Metrics and Their Corresponding Narratives,” in *2021 IEEE/ACM 4th International Workshop on Software Health in Projects, Ecosystems and Communities (SoHeal)*, pp. 25–33, May 2021.
- [62] J. Linåker, E. Papatheocharous, and T. Olsson, “How to characterize the health of an Open Source Software project? A snowball literature review of an emerging practice,” in *Proceedings of the 18th International Symposium on Open Collaboration, OpenSym ’22*, (New York, NY, USA), pp. 1–12, Association for Computing Machinery, Sept. 2022.
- [63] J. Linåker, G. J. P. Link, and K. Lombard, “Sustaining Maintenance Labor for Healthy Open Source Software Projects through Human Infrastructure: A Maintainer Perspective,” Aug. 2024. arXiv:2408.06723 [cs].
- [64] J. Ferraioli, “Bringing together a social model of open source,” Oct. 2022.
- [65] CHAOSS, “About CHAOSS,” Dec. 2023.
- [66] B. W. Boehm, *Software engineering economics*. Englewood Cliffs ; London: Prentice-Hall, 1981.
- [67] B. Boehm, B. Clark, E. Horowitz, C. Westland, R. Madachy, and R. Selby, “Cost models for future software life cycle processes: COCOMO 2.0,” *Annals of software engineering*, vol. 1, pp. 57–94, 1995. Publisher: Springer.
- [68] N. Forsgren, J. Humble, and G. Kim, *Accelerate: The science of lean software and devops: Building and scaling high performing technology organizations*. IT Revolution, 2018.
- [69] Puppet and CircleCI, “2020 State of DevOps Report,” tech. rep., Puppet, United States, 2020.
- [70] W. Scacchi and M. Ackerman, “ITR: Understanding Open Software Communities, Processes and Practices: A Socio-Technical Approach,” 2005. Publisher: Citeseer.
- [71] R. Kling, G. McKim, and A. King, “A bit more to it: scholarly communication forums as socio-technical interaction networks,” *Journal of the American society for Information Science and Technology*, vol. 54, no. 1, pp. 47–67, 2003. Publisher: Wiley Online Library.

- [72] H. Chesbrough, “Measuring the Economic Value of Open Source,” tech. rep., Linux Foundation, San Francisco, CA, USA, Mar. 2023.
- [73] NGI, “Discover NGI Innovations,” Nov. 2023.
- [74] S. Easterbrook, J. Singer, M.-A. Storey, and D. Damian, “Selecting Empirical Methods for Software Engineering Research,” in *Guide to Advanced Empirical Software Engineering* (F. Shull, J. Singer, and D. I. K. Sjøberg, eds.), pp. 285–311, London: Springer, 2008.
- [75] P. Runeson, M. Höst, A. Rainer, and B. Regnell, *Case Study Research in Software Engineering: Guidelines and Examples*. John Wiley & Sons, Mar. 2012. Google-Books-ID: T7rXoaxqPIAC.
- [76] N. King, *Interviews in qualitative research*. London: SAGE, 2009. Book Title: Interviews in qualitative research.
- [77] J. W. Creswell and C. N. Poth, *Qualitative inquiry and research design: Choosing among five approaches*. Sage publications, 2016.
- [78] A. Casari, J. Ferraioli, and J. Lovato, “Beyond the Repository: Best practices for open source ecosystems researchers,” *Queue*, vol. 21, pp. Pages 30:14–Pages 30:34, May 2023.
- [79] J. Becker, R. Knackstedt, and J. Pöppelbuß, “Developing Maturity Models for IT Management,” *Business & Information Systems Engineering*, vol. 1, pp. 213–222, June 2009.
- [80] J. Pöppelbuß and M. Röglinger, “What makes a useful maturity model? A framework of general design principles for maturity models and its demonstration in business process management,” *ECIS 2011 Proceedings*, Oct. 2011.
- [81] P. van Vulpen, A. Menkveld, and S. Jansen, “Health Measurement of Data-Scarce Software Ecosystems: A Case Study of Apple’s ResearchKit,” in *Software Business* (A. Ojala, H. Holmström Olsson, and K. Werder, eds.), (Cham), pp. 131–145, Springer International Publishing, 2017.
- [82] Z. Griliches, “Issues in assessing the contribution of research and development to productivity growth,” *The bell journal of economics*, pp. 92–116, 1979. Publisher: JSTOR.
- [83] Z. Griliches, “The Search for R&D Spillovers,” *The Scandinavian Journal of Economics*, vol. 94, pp. S29–S47, 1992. Publisher: [Wiley, The Scandinavian Journal of Economics].
- [84] B. H. Hall, J. Mairesse, and P. Mohnen, “Measuring the Returns to R&D,” in *Handbook of the Economics of Innovation*, vol. 2, pp. 1033–1082, Elsevier, 2010.
- [85] K. Davis and B. Kingsbury, “Indicators as Interventions,” *SI: New York: Rockefeller Foundation*, 2012.

## 6 Acknowledgements

We would like to thank the following individuals for their advice and/or feedback: Marco Berlinguer, Michael Brennan, Jean-Luc Dorel, Ryan Ellis, Nathalia Foditsch, Matthew Gaughan, Jeff Gortmaker, Bastien Guerry, Karolina Gyurovszka, Valérian Guillier, Anna Hermensen, Benedict Kingsbury, Fiona Krakenbürger, Michiel Leenaars, Johan Linåker, Georg Link, Katharina Meyer, Samstag Prakke, Mirko Presser, Fabien Rezac, Lorenzo Sciandra, Powen Shiah, Katherine Skinner, Paul van Vulpen, Chrys Wu, as well as the members of the NGI Commons consortium.

## 7 Funding Declaration

Cailean Osborne and Mirko Boehm are funded by the EU’s Horizon Europe programme under grant agreement number 101135279 (NGI Commons). Dawn Foster is funded by an Alfred P. Sloan Foundation grant for the CHAOSS project. Paul Sharratt is employed and funded by the Sovereign Tech Agency. The views and opinions expressed are those of the authors only and do not necessarily reflect those of the funders.

## A Prompt Questions about Potential Technological, Economic, and Social Impacts

### A.1 Technological Impacts

#### Direct impacts:

- How many critical bugs or security vulnerabilities were resolved due to the funding?
- Has the funding improved maintainer responsiveness to issues and pull requests? And how has this affected the project's development velocity and community engagement?
- What measurable improvements in code quality metrics (e.g. test coverage, cyclomatic complexity, documentation completeness) have been achieved post-funding?
- How have user experience (UX) metrics and user feedback evolved since receiving funding?
- What improvements in documentation, tutorials, and user support resources have been achieved with the funding?
- How has the project's technical debt been reduced as a result of the funding, and what impact has this had on development activity?
- What new technologies or development methodologies has the funding enabled the project to adopt, and how have these improved the project's capabilities?
- Has the funding led to rushed development of new features, potentially introducing new security vulnerabilities or leading to the neglect of other areas of the project?

#### Indirect impacts:

- How has the funded project contributed to the establishment or evolution of industry standards or best practices, and what is the scope of their adoption?
- How has the funding influenced collaborations with other OSS projects or organisations?

### A.2 Economic Impacts

#### Direct impacts:

- How many full-time equivalent positions has the funding created or sustained, and how does this compare to pre-funding levels?
- What percentage increase in project-related revenue can be directly attributed to development enabled by the funding, and through which channels (e.g. donations, commercial contracts, partnerships, services, etc.)?
- What quantifiable cost efficiencies (e.g. in development time, infrastructure costs) have been realised due to the funding, and how do these translate to financial savings?
- Has the funding created an over-reliance on a single funding source, potentially compromising the project's long-term sustainability?

#### Indirect impacts:

- How has the funding influenced the project's contribution to innovation in its sector, as measured by research publications, standards, patents, new features, or other?
- How has the funding impacted the project's ability to compete with proprietary alternatives, and what are the associated economic implications?
- How has funding impacted trust from enterprise users, and how has this affected their security auditing costs?
- What critical business operations or infrastructures now depend on the funded project?
- How has funding contributed to skill development and career advancement in the broader tech workforce?
- What impact has the project had on reducing software licensing costs at a societal level?
- What spillover effects has the funding had on the broader OSS ecosystem, such as the creation of complementary OSS projects or commercial products or services?
- What is the estimated economic impact of productivity gains for businesses using the funded OSS project?
- Has the project's growth due to the funding led to increased maintenance costs that may not be sustainable in the long term?

### A.3 Social Impacts

#### Direct impacts:

- How has the funding affected the diversity (e.g. gender, geographical, skills) of new contributors, and how does this compare to pre-funding demographics?
- What measurable impact has the funding had on the project's mentorship programs or onboarding processes for new contributors?
- How has the absolute number and retention rate of contributors changed post-funding?
- What impact has funding had on the project's ability to support and retain existing contributors?
- How has the funding influenced the project's ability to organise or participate in community events and conferences, and what has been the reach of these activities?
- Has the introduction of funding created tensions or conflicts within the community, particularly between funded and unfunded contributors?
- Has the funding led to a shift in project governance that excludes or marginalises certain community members?

#### Indirect impacts:

- What changes in community engagement metrics (e.g. mailing list activity, issue response times, pull request review times) can be observed post-funding?
- Has the project's growth due to funding led to a loss of community culture or a shift in culture that alienates contributors?
- What societal benefits have emerged from increased accessibility to the software (e.g., digital inclusion, reduced costs for public institutions)?
- How has funding influenced the project's impact on digital sovereignty and technological independence?
- Has the increased professionalisation of the project due to funding created barriers to entry for new or casual contributors?

## **B Variables for Quantitative Measurements of OSS Funding Impacts**

### **B.1 Dependent Variables**

#### **B.1.1 Technological Impact**

- # pull requests, issues, and commits
- # lines of code added or modified
- Time to first response by maintainer
- Time to close issue
- # releases or new features
- # users across installations
- # downloads or installations
- # forks and clones
- Code quality metrics (e.g. maintainability, test coverage)
- # patents filed
- # new projects initiated
- # standards organisations collaborated with

#### **B.1.2 Economic Impact**

- # jobs supported (e.g. # full-time employees)
- Revenue generated by the project or company (in €)
- Amount of follow-on funding or investment (in €)
- Cost savings for organisations using the project
- # commercial contracts or partnerships
- Market value of the project or company (if applicable)
- Economic value of volunteer contributions
- # spin-off projects or companies

#### **B.1.3 Social Impact**

- Growth in # contributors
- # active maintainers
- Contributor absence factor
- Bus factor before and after funding period
- Growth in # users
- Metrics for diversity, equity, and inclusion (e.g. CHAOSS metrics)
- Growth in stars, forks, and watchers on GitHub
- Geographic distribution of contributors
- # new contributors mentored
- # educational initiatives or workshops conducted
- Community engagement metrics (e.g. forum activity, meetups)
- # academic collaborations or citations
- Impact on skills development in the community
- # speaking engagements or conference presentations by project members

**B.2 Independent Variables**

- Funding objective(s)
- Funding amount
- Funding duration or distribution schedule

**B.3 Control Variables****B.3.1 Project Characteristics**

- Project age at time of funding
- Project life stage (e.g. inception, maturity)
- Experience level of project leaders
- Changes in project leadership
- Software category (e.g. library, protocol, devtools, application, decentralised platform)
- License type
- Governance model (e.g. BDFL, committee, foundation)
- Previous funding history and/or other funding sources

**B.3.2 Community Factors**

- Initial community size before funding
- Contributor absence factor
- Bus factor before and after funding period
- Presence of corporate backers
- Geographic concentration of core contributors
- Proportion of volunteer vs. paid contributors

**B.3.3 Technological Factors**

- Programming language
- # outdated dependencies
- Architectural complexity
- Dependency on other projects or technologies

**B.3.4 Market and Industry Factors**

- Market size for the software category
- Level of competition (both open source and proprietary)
- Industry trends and technological shifts
- Regulatory environment

**B.3.5 External Factors**

- Economic conditions (e.g. recession)
- Major world events (e.g. COVID-19 pandemic)
- Major technological events (e.g. the emergence of AI co-pilots)
- Changes in related technologies or standards