

The State of Open Research Software Infrastructure

Katherine Skinner

[0000-0003-0139-7524](#)

Jennifer Kemp

[0000-0003-4086-3196](#)

Kaitlin Thaney

[0000-0002-7217-4494](#)

Emmy Tsang

[0000-0002-9248-1280](#)

Contact: research@investinopen.org

Table of contents

[Executive summary](#)

[Introduction](#)

[Scope and Methods](#)

[Findings](#)

[Definitions and Concepts](#)

[The early evolution of the research software field](#)

[The spectrum of activity in research software](#)

[Recommendations](#)

[1. Surface hidden information](#)

[2. Strengthen the scaffolding](#)

[3. Grow the market](#)

[4. Invest in coordination](#)

[Additional Observations](#)

[Reinforcement vs overcrowding](#)

[The human infrastructure](#)

[Intentional vs accidental infrastructure](#)

[The role of institutions](#)

[The long tail](#)

[Conclusion](#)

[Acknowledgments](#)

[Appendix A: Bibliography](#)

[Appendix B: Interviews brief](#)

Executive summary

“Research Software” serves as a functional cornerstone of many research disciplines in and beyond academic contexts, from the life sciences and physics to the social sciences and humanities. It is integral to the use, modeling, and analysis of data; it also undergirds critical research functions and goals, such as openness, reproducibility, and equity.¹ Even so, research software historically has not been considered a research output in its own right. Long-standing gaps in recognition of and support for research software have narrowed recently, in large part due to the active work of stakeholders in the field.

Over the last two decades, research software developers, users, and stakeholders have founded a range of organizations, initiatives, and communities of practice, both within and

¹ <https://doi.org/10.1038/s43588-021-00048-5> (Last accessed 02-17-2025).

across specific disciplinary and national contexts. A burgeoning field² has emerged with repositories and registries (e.g., GitHub, Zenodo, Astrophysics Source Code Library, OS Monitor, Research Software Directory, Software Heritage);³ metadata standards (CodeMeta, Citation File Format);⁴ training opportunities, societies, and communities of practice (e.g., Pelagios, US Research Software Engineer Association [US RSE], Society of Research Software Engineers, Research Software Alliance [ReSA], Software Sustainability Institute [SSI], Talarify);⁵ reviewing and curation efforts (e.g., Journal of Open Source Software [JOSS], rOpenSci, pyOpenSci);⁶ and even new departments and job titles (e.g., Open Source Program Offices [OSPOs], Research Software Engineers [RSEs]). The support and activities of these groups are quickly elevating the visibility of research software, as seen through its recent inclusion in recommendations and policies (e.g., UNESCO Open Science Recommendations, the San Francisco Declaration on Research Assessment [DORA] and FAIR Principles for Research Software).⁷

This landscape review provides a system-level assessment of the research software infrastructure field. Over a four-month period (October 2024-January 2025), Invest in Open Infrastructure (IOI)⁸ conducted interviews and worked with practitioners, users, and funders of research software to identify gaps, vulnerabilities, and duplicative efforts in the field, as well as opportunities to better coordinate and align their efforts. Our remit has been to identify ways that research software's varied stakeholders can level up, maturing from an emergent field with many independent efforts into an ecosystem that can provide resilient, interdependent, and ongoing support structures for research software.

Herein, we identify concrete ways to help the field evolve into a thriving open environment with community publishing, review, and maintenance infrastructure that is supported by strong business models and collaborative arrangements. Our recommendations include:

1. **Surface hidden data:** Promoting a steadily growing, iterative stream of coordinated information about the field's activities will enable actors to situate their work in relation to each other. Better data collection, synthesis, and use could strengthen the ability to see and reward participation in research software; it could also help to demonstrate the impact of investments (time and funding) on the growth and

² We define a "field" as an economic, cultural, or political arena where knowledge, services, status, or goods are produced, distributed, and used.

³ <https://github.com/>, <https://zenodo.org/>, <https://ascl.net/>, <https://frenchopensciencemonitor.esr.gouv.fr/>, <https://research-software-directory.org/>, <https://www.softwareheritage.org/> (Last accessed 02-17-2025).

⁴ <https://codemeta.github.io/>, <https://citation-file-format.github.io/> (Last accessed 02-17-2025).

⁵ <https://pelagios.org/>, <https://us-rse.org/>, <https://society-rse.org/>, <https://www.researchsoft.org/>, <https://www.software.ac.uk/>, <https://www.talarify.co.za/> (Last accessed 02-17-2025).

⁶ <https://joss.theoj.org/>, <https://ropensci.org/>, <https://www.pyopensci.org/> (Last accessed 02-17-2025).

⁷ <https://doi.org/10.54677/MNMFH8546>, <https://sfedora.org/read/>, <https://doi.org/10.15497/RDA00068> (Last accessed 02-17-2025).

⁸ <https://investinopen.org/> (Last accessed 02-17-2025).

health of the field. Data collection is both a technical task and, importantly, a social and political task, one that would benefit from funded, facilitated coordination between the field's actors.

2. **Strengthen the scaffolding:** As the research software field matures, actors within it need increasingly stronger scaffolding infrastructure to support norms and activities. The scaffolding includes publishing and dissemination apparatuses, review practices, coordination channels, and training opportunities that enable predictability and encourage diverse people to join into the production, review, and support for those outputs. Many of these elements will, over time, become crucial infrastructure elements supporting the field. Investing in stabilizing and securing a core set of emerging scaffolding now is crucial and should be prioritized over, or at least alongside, funding for “new” work.
3. **Grow the market:** The vast majority of the actors in research software do not have an established market, or a set of known and structured interactions, to depend on, and this undercuts the sustainability (resilience) of their work. We recommend that research software actors (including funders) in this maturing field explicitly advocate for a different use of available philanthropic and government funding right now: to make both collective (on behalf of cohorts of institutions) and organizational investments that provide operational support, concrete training opportunities and/or strategic support, and that enable them to establish a viable longer-term, diversified plan for funding their work. This includes research into product-market fit and procurement channels that can support the infrastructure elements that are built.
4. **Invest in coordination:** We recommend this as a prime moment for philanthropic funders to make concentrated investments in the architecture and actualization of the field itself. Such investment could magnify the impact of other investments made in projects, programs, and organizations in research software by providing space and time to undertake well structured, facilitated engagement between some of the identified “key players” from each developing channel or function of the field (as discussed above, e.g., training, packaging, hosting, socialization, advocacy, etc.) and from across boundaries of discipline and geography.

Our findings and recommendations are situated within the larger funding and academic landscape, which poses significant challenges and opportunities of its own, particularly as international political and funding trends inhibit research support. Our recommendations focus on streamlining activities and increasing network-based engagements in ways we believe might best benefit this set of well-intentioned, often deeply engaged players who will be facing a much more challenging fiscal and attention cycle in the near future.

Introduction

We see this as an urgent moment for planning towards potential scenarios in which funding and institutional support for “open” initiatives, including volunteer time and bandwidth, are likely to quickly decline.

Science is a highly international endeavour, but support for it is closely tied to national policies and grant funding. Recently, significant cuts to “open science”⁹ research funding have been made in a number of countries, including Argentina,¹⁰ France¹¹ and even in the Netherlands,¹² which has long been a leader in this space. These, along with efforts to dismantle research funding and its administration in the US, are significant destabilizing factors for emerging and established fields in the international research landscape.¹³ We see this as an urgent moment for planning towards potential scenarios in which funding and institutional support for “open” initiatives, including volunteer time and bandwidth, are likely to decline. To that end, we give special attention throughout this report to building more concentrated, nimble, interconnected, and resilient structures for this evolving field, particularly those who are working not just to produce research software, but also to develop infrastructure to support it.

Our research seeks to understand factors that might help the research software ecosystem to thrive and sustainably develop. In this work, we identify:

- **Existing connections** in the research software infrastructure ecosystem, and where new connections can be built to advance sustainability and resilience
- Areas in the ecosystem **where there is overcrowding, and where there are gaps, blockers, vulnerabilities, and/or critical dependencies** that need to be addressed
- **Indications of what helps infrastructure services and initiatives succeed** in adding value to the research software ecosystem

⁹ Open science, like research software, has many definitions. A common definition comes from the UNESCO Recommendation on Open Science: <https://doi.org/10.54677/MH8546> (Last accessed 02-17-2025).

¹⁰ <https://doi.org/10.1126/science.zjcq1nt> (Last accessed 02-17-2025).

¹¹ <https://doi.org/10.1126/science.z6dcmg8> (Last accessed 02-17-2025).

¹² <https://www.openscience.nl/en/news/open-science-nl-budget-cut-by-half-the-transition-to-open-science-is-as-important-as-ever> (Last accessed 02-17-2025).

¹³ At the time of writing in February 2025, this story is still actively developing through 1) a slew of executive orders and policies under the Trump administration and 2) legal challenges made to those orders and policies.

This report opens with an overview of our scope and methods; it then moves into definitions and concepts and provides an overview of some of today's major actors and initiatives. This is in no way meant to be exhaustive, but rather indicative of who and what seems to be “top of mind” in the current climate for those we interviewed.

We then assess the state of the field today. We use a spectrum approach to measure where research software seems to be against common maturity markers. Based on that assessment, we share a set of targeted recommendations, describing some of the first steps that might be taken to test or implement them. We also note specific players and initiatives who seem to be ready for more coordinated action and we look at what barriers stand in their way.

Scope and Methods

This work was conducted between October 2024 and January 2025. By scope, this research largely describes the evolving and partly interconnected research software environments in North America, Australia, and Western Europe. Where possible, we have included additional perspectives, particularly from interview participants active in research software and related infrastructure in Africa and Latin America, to shed light on the different paces and ways research software is evolving in different geopolitical contexts today.

We followed a qualitative approach combining desk research with community engagement. To inform the project direction and approach, the team first prepared an extensive collection of desk research sources and notes, including identifying:

- Relevant organizations, services, and initiatives, categorized by type — for example, documentation and workflow tools, software publishers and repositories, and communities of practice
- Related topics:
 - A detailed discussion of open source software is beyond the scope of this report
 - Similarly, although hardware, including laptops and smartphones, are part of the overall landscape, they are distinct from the research software environment as defined for this report
- A core, annotated bibliography of more than 50 sources, to which additional sources and notes were added by project team members during the project period.¹⁴

¹⁴ <https://doi.org/10.5281/zenodo.14178126> (Last accessed 02-17-2025).

Based on this research, we identified prospective interview participants representing a wide range of activity in research software around the world and invited 23 organizations to participate in interviews with us. We ultimately conducted 19 interviews with key stakeholders: founders, leaders, and participants in research software's social and technical infrastructure, including academic, nonprofit, and private environments. Interview participants included research software engineers, trainers, repository managers, publishers, directors, a funder, and leaders from the affiliated research data community, representing a variety of disciplines and software functions on four continents. Interviews served to validate, interrogate, correct, and/or round out information gathered from desk research. They were also used to surface things that people may be thinking about but not sharing openly in writing or elsewhere. The interview instrument is available on Zenodo and names of the interviewees who wanted to be listed are included at the end of this report.¹⁵

Resources used in our research are available in IOI's public Zotero library.¹⁶

Findings

Based on our desk research and interviews, research software has been emerging and actively evolving as a field¹⁷ for at least two decades. We use the sociological construct of "fields," and more specifically, a field theory lens (Bourdieu, 1993) to help us hone in on the ways that networks of organizations and individuals have been engaging in and actively developing this field during this formation period.¹⁸ Doing so, we draw attention to the patterns of interaction between players, including how they are negotiating for the visibility, meaning, and structural predictability of the field itself during its emergence, as well as for what their own positions will be within this field as it grows. We surface and describe characteristics of the field as seen from a number of vantage points, including those of leaders and practitioners working in research software, as well as those working in related and overlapping fields like research data.

We focus especially on the infrastructure (social, technical, and administrative) supporting research software and mark how this infrastructure both relies on and reinforces the evolving field's definition and boundaries. Using lifecycle development and

¹⁵ <https://doi.org/10.5281/zenodo.14141210> (Last accessed 02-17-2025).

¹⁶ https://www.zotero.org/groups/4377072/invest_in_open/collections/JX8XY6G4 (Last accessed 02-17-2025).

¹⁷ We define a "field" as an economic, cultural, or political arena where knowledge, services, status, or goods are produced, distributed, and used.

¹⁸ Our research uses field theory as a lens to study and understand the ways in which the people and organizations engaged in the support of research software are constructing a "field," or environment to provide structures for the production and exchange of a particular type of knowledge and activity. See Pierre Bourdieu, *The Field of Cultural Production*. (Cambridge: 1993).

health assessment tools, we highlight gaps, overlaps, and opportunities that, if addressed, may help this field to optimize its health and its growth.¹⁹

Definitions and Concepts

What is research software?

Several of our interview participants cited their involvement in writing formal definitions of “research software”²⁰ and many participants provided pointers to these. Most also shared brief personal or shorthand definitions that largely aligned. These informal definitions can be summarised as:

Software specific to research.

A few participants included more detail and some variation in their definitions, e.g.:

- using “software or code” or using “tools” in place of either
- adding that it is used in or to advance knowledge or science, progress research, support scholarship, etc.

Many participants noted that their definitions exclude software that is used in research but is not created by or for researchers or strictly for research purposes (e.g., Microsoft Excel).

What is research software infrastructure?

In this report, we are concerned with “research software infrastructure,”²¹ by which we mean the functions, tools, structures, and human labour on which research software relies. We acknowledge that open-source code (and even “research software” itself) is

¹⁹ These tools help to describe the health and evolution of projects, communities, organizations, and other entities, making clear not only what the health or status of an entity is at a moment in time, but also shedding light on what that entity can do to continue growing and maturing. See e.g., It Takes a Village (LYRASIS, 2018) <https://itav.lyrasis.org/guidebook/> (Last accessed 02-17-2025), Community Cultivation Field Guide (Educopia, 2018) <https://doi.org/10.5281/zenodo.1491172> (Last accessed 02-17-2025), FOREST Framework for Values-Driven Scholarly Communication (Educopia, 2022) <https://doi.org/10.5281/zenodo.6557301> (Last accessed 02-17-2025).

²⁰ As discussed in Gruenpeter et al. (2021), defining “research software” can be controversial. A commonly used definition is one that the FAIR for Research Software (FAIR4RS) Working Group discussed: <https://doi.org/10.1038/s41597-022-01710-x> (Last accessed 02-17-2025). Sochat et al. (2022) discuss how they created The Research Software Encyclopedia in order to help a broad community to define the term in various contexts: <https://doi.org/10.5334/jors.359> (Last accessed 02-17-2025).

²¹ DFG (the German Research Foundation) has defined “research software infrastructure in the context of their “Research Software Implementation Programme” as “technically and organisationally networked services and products, e.g. for creating, processing and using as well as accessing and maintaining research software.” <https://www.dfg.de/en/research-funding/funding-opportunities/programmes/infrastructure/lis/funding-opportunities/research-software-infrastructures> (Last accessed 02-17-2025).

sometimes described as “infrastructure,”²² but in these cases, the code-as-infrastructure is supporting something else (e.g., research). Here, our definition intentionally focuses on those infrastructure elements that support the development, use, publishing, reviewing, discovery, and maintenance of research software.

Not all infrastructures meant to become infrastructure or originally thought of their work as “infrastructure,” as we describe later in this report. Shifting from a landscape peppered with a lot of “accidental infrastructure” elements (e.g., social, technical, or administrative scaffolding that grew organically and without cultivation towards a shared future) to one in which some thoughtful compromises and combinations are made could significantly improve the FAIR-ness²³ and resilience of our research ecosystem.

The early evolution of the research software field

Use of what is now termed research software dates back to the middle of the 20th century. As a field, research software and its supporting infrastructures began to form more than two decades ago, when people doing similar things found each other and began collaborating and competing with each other. Researchers working in a variety of lab environments and disciplinary structures whose research depended on specialized software to collect, parse and analyze data (for example) found colleagues doing similar things in other lab environments and disciplinary structures. These peers grew their own informal networks over time, beginning with many disconnected, small, topical clusters. They used existing social and technical infrastructure based in the fields within which they were embedded (often both discipline and geography based), including publishing venues, conferences, and other existing forums, to share their work, convene formally and informally, and identify and recruit others with similar interests and identities to the field. As is typical of emerging fields, these small networks expanded, came into contact with each other, and began to overlap. After recognizing their similarities across boundaries, these networks worked actively to enable and encourage new intersections.

In 2025, the field is still navigating its evolution towards formalization; it remains dominated by discipline-based and geographical pockets of activity, many of which are overlapping and trying to sort out their potentially competing missions, goals, and funding options.

²² See e.g., Nadia Eghbal, Roads and Bridges: The Unseen Labor Behind Our Digital Infrastructure (Ford Foundation: 2016) <https://www.fordfoundation.org/work/learning/research-reports/roads-and-bridges-the-unseen-labor-behind-our-digital-infrastructure/> (Last accessed 02-17-2025). ; Rob Van Nieuwpoort and Dan Katz (March 14, 2023) “Defining the roles of research software.” Upstream blog, <https://doi.org/10.54900/9akm9y5-5ject5y> (Last accessed 02-17-2025).

²³ <https://www.go-fair.org/fair-principles/> (Last accessed 02-17-2025).

As these practitioners worked together to validate their work and efforts in research software, the “research software” identity began to coalesce, attracting more practitioners and adjacent stakeholders (including funders) to the work. Throughout the 2010s, the visibility and viability of this new field was established, as evidenced by the spawning of new events, training opportunities, grant-funded research, and community groups, as well as the development of initial vocabulary, definitions, and even an emerging professional role, Research Software Engineer (RSE).²⁴

By the early 2020s, the research software field established relatively solid footing, if uneven acceptance, in locations including Australia, Asia, Canada, the EU, UK, and US; by this time additional communities and activities began to ramp up in Latin America, South America, India, and Africa as well. Signals that research software was maturing strengthened, with a growing slate of training and education opportunities, publishing/deposit standards and environments, community support mechanisms (including new programs and organizations), and citations in geopolitical policies and recommendations.

In 2025, the field is still navigating its evolution towards formalization; it remains dominated by discipline-based and geographical pockets of activity, many of which are overlapping and trying to sort out their potentially competing missions, goals, and funding options. This is relatively typical of the institutionalisation process: practitioners and other stakeholders in an emerging environment find their “home” in small communities, experiment to express their voices and form identity within these units, and then begin to network across those units to form a meta layer of interdisciplinary and international support infrastructures.

In a stable political and financial environment, we could expect to see that meta layer strengthen and the set of social, organizational, and technical infrastructure elements narrow in number and focus over the next decade or so. Eventually, synergies would be found, gaps would be filled, duplication would be reduced, and the field of practice would settle into a more formalized rhythm, much as we are seeing now happen with adjacent fields, such as research data.

The political and fiscal climate makes this a moment in which we need to champion the early efforts to build across initiatives and increase collaborative planning and action as much as possible.

²⁴ See <https://www.software.ac.uk/blog/not-so-brief-history-research-software-engineers-0> (Last accessed 02-17-2025).

Unfortunately, the research enterprise is not in a stable or supportive political or financial environment today, and we believe this directly impacts the forms of attention and activity this field now needs from its advocates, adopters, and investors. We have entered into a period in which many geopolitical systems are cutting support and funding to “open science” and academic endeavours. This has ripple effects across the broad research ecosystem in which research software is just one of many players. The political and fiscal climate makes this a moment in which we need to champion the early efforts to build across initiatives and increase collaborative planning and action as much as possible. Reaching consensus among stakeholders on the most important outcomes the research software community(ies) is aiming to achieve and establishing a neutral (not-directly-invested) facilitation-oriented “hub” or “backbone” to help spotlight, track, and celebrate progress towards those shared outcomes/goals may help to advance this field’s maturation and make it more resilient in the face of change. Our interviews have surfaced important viewpoints and perspectives regarding what such outcomes/goals might include, including how this work might relate to and help to forward similar work in adjacent fields like research data.

The spectrum of activity in research software

Throughout the interviews, a number of themes arose that pertain to the development and formalization of the field. Below, we surface three of these themes, and identify where interview participants place different aspects of the field on a spectrum of activity. For example, many of our interview participants commented on the degree of visibility they see “research software” having today. To capture this, we have roughly defined a spectrum from hidden to visible, and we try to fairly represent both the dominant view(s) of our interview pool, as well as the specific impressions of individual interview participants. The categories we are using for these spectra are not meant to be binaries, and there is no “ideal state” that the field needs to strive for or reach; we aim to provide a more descriptive, less judgmental framework for considering what various stakeholders are looking for and what these might mean for the development of a shared agenda in the near future.

The categories we are using in these spectra are not meant to be binaries, and there is no “ideal state” that the field needs to strive for or reach.

Our research shows that, while significant progress has been made towards the stabilization, formalization, and recognition of research software as a field, this progress is experienced unevenly by its stakeholder groups. Stakeholders generally agree that awareness and use of research software appear to be on the rise, although some note that the terms and definitions still differ significantly in different geographical contexts. They also agree that funding sources and amounts for various parts of research software,

including community infrastructure, have grown over time, but many point out the ways that funding gets trapped within geopolitical boundaries, making sustained collaborations at multi-national levels challenging at best.²⁵ Sustainability of research software tools, projects, and organizations remains elusive and, on the whole, the ecosystem is still largely siloed (or niche, depending on one's perspective). No sectors or functions we have studied have emerged as fully mature or formalized, which is not at all unreasonable at this juncture but, as per our statement above, leaves this field at risk in this political and fiscal climate.

Below, we use three key categorizations to place and understand "research software" today according to the range of perspectives we gathered in our interviews, beginning with "hidden/visible," and moving through "fragmented/coherent," and "fragile/stable." Again, note that these are not meant to represent scales of activity, with the second term somehow "better" than the first. In many cases, healthy engagement will require a mix of both qualities. For example, some siloed work enables new approaches and encourages the field to change and grow over time, but too much siloed work leads to expensive duplication and sows confusion. Likewise some interdependence can be an asset, setting up formalized and recognizable paths to shared aims, while too much interdependence can indicate that structural control is defined and guarded by a small number of players, limiting diversity and challenging the health and growth of the field. Our intention is not to label the field or judge its development, but rather to understand how our interview participants seem to be experiencing its maturation progress through lenses that we can then use to identify key threats, challenges, or opportunities that might accompany particular maturation moments.

The landscape as a whole is at an "in between" stage.

We use language from our interview participants wherever we can below, although we have taken care to anonymize it rather than quoting specific organizations or individuals. That choice hopefully increased openness and frankness within our small and recognizable interview participant pool.

²⁵ Notably, this characteristic is shared by many other fields in this space; the funders and assessors of research and academic ecosystems often hail primarily from government sources and/or aim towards specific national or regional goals, and even philanthropic funding sources face significant challenges regarding the rules and regulations governing the movement of funding across national boundaries. This may be in part due to the degree to which the "market" is not-for-profit and not structured for the type of engagements that often happen in commercial industries across for-profit players.

1. Hidden ↔ Visible

Hidden



Visible

As described by one interview participant, the research software landscape is “like an iceberg - some is visible, most is not.” Almost all interview participants added caveats to their overall assessments, calling out areas with which they are unfamiliar and/or where they see other members of the extended “research software community” landscape missing. Synthesizing these views, most of our interview participants placed research software squarely in the middle of this spectrum, citing that some areas are visible and others very niche. Interestingly, the *areas* they referenced as “visible” vs. “hidden” varied, with some citing the need for stronger bridges and better information exchange between different types of stakeholder communities in order to increase awareness of existing resources and services.

[...] the research software landscape is “like an iceberg - some is visible, most is not.”

Our participants shared that:

- Research software is more and less visible depending on geopolitical location; it also looks different in various contexts. For example, while RSE has emerged as an established professional role in the UK, EU, and US, people performing this professional role use different terminology and different understandings to indicate this work in other locations including Africa, Canada, and Latin America.
- There is a need for canonical resources that the field can turn to and trust, though participants cited different elements. Some noted the need for resources to better identify and cite research software; others talked about the need for better ways to track grants/funding for software projects (who gives, who receives, how much), and still others wanted to see concrete and trusted ways to track citations and mentions.
- A nexus of authority is not yet established; there is no coherent “center” or set of players that document and corral and make more visible and known the evolving players and resources, though a few organizations have tried to provide such a hub in specific areas (e.g., Software Sustainability Institute with leadership on

professionalizing Research Software Engineering, or ReSA with its recent documentation efforts and work to align funders).

- There seems to be a major chasm between academia (i.e., students and faculty) and some communities; academics are often unaware that these resources and social infrastructure supports exist, or how contribution/participation in these infrastructures can be beneficial to their work/careers. For example, academics and recent graduates report feeling alone figuring out how to package code; they are unaware of resources like the pyOpenSci and rOpenSci communities.
- There is a lot of “real world” data that the field does not have or know how to gather, at least within current resource constraints. Some of this is information that, in commercial or production software, would be tracked and evaluated by a paid, product management role that is typically lacking in research software.
 - How much software is produced/used in research environments? This information is not generally tracked by institutions or funders, let alone shared.
 - How much of research software activity is funded privately vs publicly?
 - How much is open-source vs proprietary?
 - What features or aspects are popular/used? What percentage of those who go to training never use a given tool?
- Research software is now being noted in some policies, including the UNESCO Recommendations on Open Science and in documentation in related fields, but often it is mentioned only in passing and without specifics, i.e., “and software.”²⁶
- Visibility of research software also translates to funding for research software infrastructure. What role(s) might different groups in academic and research institutions play (e.g., Vice Provosts of Research, Chief Information Managers, Open Source Program Officers) in tracking need and use of research software infrastructure locally? How can practitioners and advocates build better awareness of (and established procurement channels for) research software infrastructure in labs, universities, and research centers? These bridges do not yet exist.
- There is a need for ways to clearly identify and reward practitioners. Interview participants noted existing national and international awards that are recently arising, including the French National Open Science Award for Research Software, and also Australian Research Data Commons (ARDC) and Helmholtz Association awards. Having opportunities for recognition may help increase visibility for both individual winners and for the field.

²⁶ In some policies, research software is curiously absent, for example, the US “NSPM-33” Memo from the National Science and Technology Council (NSTC) on research security (2021) and related Guidance (2022): <https://bidenwhitehouse.archives.gov/wp-content/uploads/2022/01/010422-NSPM-33-Implementation-Guidance.pdf> (Last accessed 02-17-2025).

2. Fragmented ↔ Coherent



The majority of interview participants noted that a major challenge for the field is balancing concurrent and conflicting needs: 1) for independent, sometimes siloed work on specific topics and in specific locations to serve a specific target community and 2) the need for coordination among organizations and initiatives to propel the field forward. Interview participants demonstrated alignment around several well-cited “visions” for the field, including elevating research software to a first-class research object, making research software more discoverable and reusable, and recognizing contributors and functions across the research software lifecycle. However, attention to specific actions and goals (when, where, and what needs to be done by whom) and expectations regarding how to measure progress towards these goals across collaborating groups is still lacking. Most of the individuals in our interview participant pool reported they are focused either on specific projects and/or the success of their own organisation. Unified or collaborative efforts are still challenging to undertake, not least due to visibility constraints and funding limitations. Key sub-elements in this spectrum include the degree to which the field is currently siloed and/or interdependent, and how much work is discipline-specific and dependent on regional and/or interdisciplinary and international lenses. Overall, interview participants placed “research software” more towards the fragmentation end of the spectrum and agreed that while some level of independent action and discovery work is appropriate, more care currently is needed for encouraging collaboration and coordination.

Interview participants shared that:

- There is general agreement among most interview participants on a number of potential common goals and approaches,²⁷ including:
 - Elevating research software to a first-class research object
 - Making software easier to discover, circulate, cite, and reuse
 - Recognition for functions and contributors across the research software lifecycle
 - Providing support for human infrastructure to grow/deepen skills for research software developers, maintainers, and practitioners across the landscape

²⁷ Notably, these parallel some of the common goals and approaches that are often named in the preprint arena and the research data field. We discuss this further in the fragile/robust section below.

- There is a significant degree of siloed work in the landscape. Interview participant opinions were split on whether this is a feature or a bug. Silos and the lack of overall visibility means there is duplication of work for some interview participants; others cited that there are so many drivers and contexts that maybe “silos” are really sometimes “cylinders of excellence”. Some advocated for protecting these boundaries, while others encouraged bridging across them. Some talked more about disciplinary boundaries, some were attuned to national boundaries, and still others discussed functional boundaries.
- Discipline-specific work and the associated “long tail” of research software was often called out as niche (siloed) work; however, several interview participants stressed the importance of this long tail and the need for recognition and support for these sometimes highly influential or important, but overwhelmingly one-off, developments.
- Legal barriers to cross-border funding in different national contexts complicate the ability of the landscape to reflect and support the international nature of research. Several interview participants cited the need to attend to this ongoing dilemma, perhaps through new mechanisms built to move funding between jurisdictions.
- Broad geographic representation is lacking in the ecosystem and disincentives to improving this are significant. Interview participants shared their concerns about what this means for furthering equitable access to the creation, use, and dissemination of research software and research software infrastructure.
- There are challenges as work moves towards needing more cross-, multi- and interdisciplinary collaboration and synthesis; some players/mechanisms (e.g., GitHub) are making it easier for practitioners to find each other and make their own work more visible across these disciplinary boundaries, but interview participants noted we are still far away from envisioning or realizing opportunities to create or adapt research software so that it attends to or accounts for the methodological specificity of different disciplines.

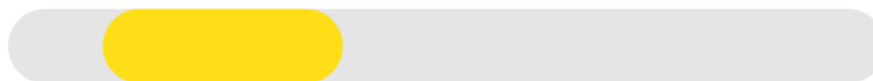
“[...] when competing efforts are started because someone doesn’t know about another, it fractures the space, becomes a matter of who you know, not what’s best.”

- Individuals and small networks of players have a tendency to want to do their own thing rather than joining up with existing efforts, often starting similar (if not duplicative) efforts that then also contribute to the strain felt by other initiatives on resources, attention, and funding. As one interview participant stated, “Everyone

wants to start something new rather than working on things that ... already exist ... when competing efforts are started because someone doesn't know about another, it fractures the space, becomes a matter of who you know, not what's best." This especially arose in our interviews around two issues: 1) competing citation standards and 2) around competing coordination efforts, for example, multiple infrastructure entities funded to do similar things, sometimes even by the same funding agency, versus exploring a more networked approach.

3. Fragile ↔ Robust

Fragile



Robust

Interview participants spoke at length with us about their concerns regarding the fragility of the current research software field and their ideas on what might propel it towards more sustainability, including more transparency around both needs and costs. Some also shared their worries about what compromises might be necessary in order for research software to reach a more stable state. In particular, we heard from some interview participants that because research software is not fully institutionalized,²⁸ its practitioners have more freedom to define appropriate structures and flexibly encompass a wider range of practice. We also heard a range of opinions regarding both the value and risks implicit in several fiscal hallmarks of the field today, including the prominence of volunteer work, the reliance on grant funding, and the lack of additional revenue sources to cover costs. While interview participants leaned towards defining research software using the "fragile" end of the spectrum, they noted ways that the field has thus far kept its costs low, in part by building on and adopting the practices of adjacent fields, especially research data. Some interview participants are wrestling with the multiplicity of social infrastructure elements like community organizations (e.g., the US Research Software Sustainability Institute [URSSI] and ReSA, or the regional/national RSE associations and the Society of Research Software Engineers, and also components of other longer-standing players with slightly different remits like the Software Sustainability Institute [SSI] and Software Heritage); others wonder if research software could sidestep costs and duplication of effort by being accepted as or promoted to an equal status with research data and in effect, build on that existing infrastructure.

Interview participants shared that:

²⁸ By "not fully institutionalized," we mean it has permeable structures rather than the standard systems of production, exchange, and appropriation that are instantiated as fields mature and formalize.

- Research software is not yet seen as a peer with other areas of published research output (e.g., articles, books, or even data) though it has made significant progress towards this goal. Some interview participants spoke about it as being seen as the “ugly stepsister” to research data and aspired to change that perception. By doing so, they hoped advocacy and other work could serve to promote both research data and research software simultaneously, lending greater sustainability to both efforts.
- The landscape appears to rely heavily on volunteer work, and in many ways this matches existing hidden and unpaid labour practices within both publishing and open source communities. Some interview participants indicated that “money complicates things,” including sustainability, and shared their view that unpaid contributions were a sign of health. Others noted that dependence on those who are not paid significantly (even egregiously) narrows the potential participant pool to those who can afford to volunteer and prioritizes those with existing power, income, and reputation. It also risks volunteer burnout, perhaps especially with behind-the-scenes maintainers. Notably, it also reduces the prospective “next generation” of professionals who are not gaining the requisite opportunities for exposure, training, and mentorship in this field.
- Funding is challenging to measure and understand because so much detail is missing, including even the most basic information about who has been funded for what activities to develop and support the field’s necessary infrastructure. Transparency of cost information would be enormously useful in the development of paid roles and in estimating costs of meeting a growing range of funder policies by providing public access to research software outputs.
- “Resilience” of research software and RSE activities are both tied to a number of vectors that do not inspire high confidence:
 - the challenges of national/international and disciplinary/interdisciplinary boundaries around available funding;
 - the tightness of the funding landscape in a moment where defunding of Open Science is rampant and expected to increase;
 - new institutional forms (e.g. OSPOs) that have potential but are not explicitly tied to research software aims, and that are being planted via start-up grant funding that draws from existing funding pools and will be available to OSPOs for a limited timeframe. These models also are designed based on examples in the commercial sector, leading to another adjacent institutional initiative around open source and/or software that is not linked to existing efforts on campus to serve researchers.
- There are too many overlapping efforts; it requires time and resources to follow developments, let alone to understand them all. Interview participants disagreed

about the source of this challenge, with some pointing to funders for distributing funding in duplicative ways, others commenting on particular players who have built too many separate-but-repetitive projects and programs over time, and still others pointing to the complexities inherent in the also-overlapping field boundaries that leave many efforts competing for attention in the “open research,” “open access,” “open source,” and “open science” arenas.

- Interview participants recognized, significant opportunities to explore convergence between key groups such as:
 - pyOpenSci, rOpenSci, and other packaging tools and community support structures
 - persistent identifiers (PIDs), standards and metadata players
 - the national and regional RSE groups and associations
- Mergers and shared service environments were named as approaches that might help alleviate some of the fiscal challenges ahead, particularly if they happen before funding cuts are too far underway.
- There is a strong need cited by many interview participants for trusted intermediaries and facilitators (“community building agents” as one interview participant said) who can develop registries and other tools that help groups to find each other and individuals to locate the communities they need. Strong community facilitation could also help bridge and translate perspectives across different players, helping to prevent real or perceived top-down agendas or prescribed approaches.
- Many viewed fundraising for their initiatives/organizations as not necessary, or as someone else's responsibility. Sometimes this response was based on an aversion to taking on a pay structure for labour that may tie them to grant-based fundraising and rising operational costs. Other times, interview participants noted explicitly that the history of open source projects was to be contributor / volunteer based (and shared that they themselves dedicate time to these efforts, sometimes from their paid positions).
- There are different expectations of academic work in different countries. In parts of Europe for example, academics are civil servants and the atmosphere is different than in the US or UK where academics might do consulting or have other, additional paid work.
- Several paid, professional roles that are missing and sorely needed:
 - There is an overall lack of product management, which means that work on user experience and accessibility, for example, is also generally lacking in the space
 - The need for community managers came up frequently

- Professional marketing and branding was also frequently cited as a need

Recommendations

Below, we offer a set of recommendations based on our research into this evolving field. Please note that a few of our recommendations draw on and reinforce existing advice, including those from the Amsterdam Declaration on Funding Research Software Sustainability (ADORE),²⁹ Science Europe,³⁰ and others.³¹

1. Surface hidden information

So much of what is happening in research software - and in the development of infrastructure to support it - remains invisible due to a paucity of available, standardized, and meaningful data. This includes data about funding (who funds what actors to do what activities with how much funding), about production (who creates what tools and objects and what do these do and where can they be found and with what licensing), about use/reuse (what is being used, by whom, and how much) and about engagement (where are people coming together via a broad range of forums including listservs, working groups, associations, events).

Maturation and formalization of the field will depend in part on what data can be captured and used to assess the field's development, players, resources, and products/services. Current research software incentives and structures overwhelmingly reward independent "leadership" rather than collective action. Each actor can build their own plans, definitions, tools, and guidance documentation, using these to negotiate for their own power and funding rather than for growing the health and status of the field

As a result, where we need a collective with many team roles, we have multiple projects, programs, and organizations, each trying to claim centrality through broadcasting what they produce and define. In place of a recognized (and ideally decentralized) "canon" or "core" to guide a "commons" for research software (including definitions, metadata and quality standards and best practices, policy information, events/forums, and other components that require broad buy-in to be successful), we have a free-for-all that most players can dodge participation in because it is not producing visible information.

Establishing a kind of "core" via data collection is largely not a technical task, it is a social and political task, one that would greatly benefit from funded, facilitated coordination between the field's actors. The "center" does not need to (nor do we recommend that it

²⁹ <https://doi.org/10.5281/zenodo.13735888> (Last accessed 02-17-2025).

³⁰ <https://doi.org/10.5281/zenodo.13740998> (Last accessed 02-17-2025).

³¹ See, for example, Strasser, C., Hertweck, K., Greenberg, J., Taraborelli, D., & Vu, E. (2022). Ten simple rules for funding scientific open source software. *PLOS Computational Biology*, 18(11), 1010627. <https://doi.org/10.1371/journal.pcbi.1010627> (Last accessed 02-17-2025).

should) be or become a single association or voice. Instead, it could be a temporary sub-structure that supports ongoing efforts by helping them to gather and synthesize what current associations, projects, and programmatic areas in research software are doing.

Specifically, we recommend:

- **Bringing together a small group of incentivized players**, beginning with representatives from major points of activity in the growing field. Facilitate them through a time-bound process to establish 1) a clear set of standard data points for reporting purposes, and 2) a clear way to make that initial set of data visible and usable. This work will require funding, as does all collaborative activity, and for the sake of equity and diversity, this should be funded centrally through external funders (philanthropic, government, or institutional) and be structured to fund participation from underrepresented regions/groups.
- **Making the initial data from this group visible, even if it is imperfect.** This should include providing an initial demonstration that is led by the group devising the reporting mechanism and responding to feedback transparently and publicly. All participants should also use their networks to broadcast their agreed upon starting place.
- **Fostering understanding of, and compliance with, new data reporting standards by quickly broadening the set of participants.** Incentives for participation might include funding and/or concrete activities and support for the first few rounds of actors. As per above, this also helps to ground the reporting in as much equity and diversity of participation as possible.
- **Recognizing that any standardization will start off with lots of missing players**, and that addressing those gaps is a necessary component of the building process; developing standards will be refined (both for collection and use) as challenges come to light.

Promoting a steadily growing, iterative stream of coordinated information about the field's activities will enable actors to situate their work in relation to each other, shifting the focus from individual "leaders" in the field to the commons that provides adequate structure to open participation and demonstrate the impact of investments through the growth and health of the field. Timing for such activity is auspicious, particularly given the increasing attention given to "open data" for citation, ranking, and assessment, e.g., OpenAlex,³² OpenCitations,³³ the Barcelona Declaration,³⁴ dashboarding efforts by groups like Curtin

³² <https://openalex.org> (Last accessed 02-17-2025).

³³ <https://opencitations.net/> (Last accessed 02-17-2025).

³⁴ <https://barcelona-declaration.org/> (Last accessed 02-17-2025).

Open Knowledge Initiative (COKI, Curtin University),³⁵ Scholarly Communications Lab (ScholCommLab), Simon Fraser University and University of Ottawa),³⁶ and the Centre for Science and Technology Studies (CWTS, Leiden University),³⁷ as well as other efforts.

2. Strengthen the scaffolding

New fields emerge through the work of a set of often-visionary individuals, and the projects, programs and organizations they create. These early actors advocate for meaning and negotiate for visibility and status; they also provide space for establishing practices and expectations regarding how the field's players produce, disseminate and build on each other's work. Emergence is concerned primarily with defining and validating the field, and this early work is often led by individuals with the desire and means to volunteer. They do not depend on the new structures for their own stability; they take risks and often intentionally operate on narrow margins.

As the field matures, actors within it need increasingly stronger scaffolding to support norms and activities. In the case of research software, the scaffolding³⁸ includes publishing and dissemination apparatuses, review practices, coordination channels, and training opportunities that enable predictability and encourage diverse people to join into the production, review, and support for those outputs. Many of these elements will, over time, become crucial infrastructure elements supporting the field.

It is time to strengthen this scaffolding for research software. Doing so will require:

- **Identifying key functions and empowering the actors who are best positioned to support those functions to focus and specialize.** Where early actors tend to morph and grow to fill the space, maturing actors need to hone in on a specific service or specialty and offer it in ways that are stable and scalable. The functions we identify here require the ability to guide projects, people and initiatives through their next steps with purpose and care. Functions we see as particularly crucial to develop further generally fall within professional management, operations and community development skill sets and include organizational leadership and change management, professional facilitation/convening, staff and volunteer management, fundraising/business model/partnership development, product management, branding/marketing, policy/advocacy and community engagement.
- **Prioritizing resourcing and training integrators and maintainers.** Predictable environments and services require different skill sets: leaders need to excel not

³⁵ <https://open.coki.ac/> (Last accessed 02-17-2025).

³⁶ <https://www.scholcomm.ca/> (Last accessed 02-17-2025).

³⁷ <https://www.cwts.nl/> (Last accessed 02-17-2025).

³⁸ We use the term “scaffolding” to denote those elements that, with appropriate instantiation, might become backbone infrastructure (social, technical, administrative) supporting the field.

only at vision, but also at building trusted operations, good governance, and their own succession plans. Strong leaders for infrastructures will need to manage paid staff, attract predictable sources of income, and support longer-term commitments. A maturing field will also require product managers and sales teams who can work in coordination with technical actors to build services that conform to defined user needs. To say this a slightly different way, visionary players have identified a lot of the work that needs to be done, but that does not mean they are the right people with the right skills to do the next part. In our interviews we often heard a lot about personality-driven efforts and about the need to encourage leadership transitions to offset the challenges that some of the field's founders inadvertently bring into the work.

- **Encouraging and enabling consolidation, specialization, mergers, and sometimes transitions and hand-offs to different players.** The field is rife with experiments, and organizations and communities are propped up with small amounts of funding only to falter or fail after a short time of activity. As long as new creation is valued and rewarded over stabilization, the field will lack the coherence needed to structurally support the activities of creating, sharing, and reusing research software. If investment prioritizes strengthening key players, a stronger service environment can develop more rapidly.
- **Opening better opportunities for equitable and inclusive participation.** As is also true in many open source communities, volunteer labour and below-cost offerings are valorized in the research software landscape. In our research, we heard many key players talk about building “sustainability” through ensuring they have no financial obligations or ongoing income/revenue expectations. While such energy is important (particularly in early emergence), it is also exclusionary: only those with existing means can play. It undercuts maturity and growth, particularly in the early days when each service is trying to build a “market” of users.
- **Provide actionable guidance for improving practices.** There are several “best practices” available for the field. While best practices can be useful, they can also be intimidating to those who are still getting grounded - and during field emergence, that is likely to be most people and organizations. “Best practices” are often aspirational; it can be more helpful to build guidance that shows how to build from wherever you are, developing good and better practices in stages over time. A framework for community consultation and feedback on such guidance is recommended. To that end, we recommend developing and documenting a spectrum of practices from good to “best,” coupled with steps that any organization or project can take to better their own practices.³⁹

³⁹ Examples include Garijo D, Ménager H, Hwang L, Trisovic A, Hucka M, Morrell T, Allen A, Task Force on Best Practices for Software Registries, SciCodes Consortium. 2022. Nine best practices for research software registries and repositories. PeerJ Computer Science 8:e1023 <https://doi.org/10.7717/peerj-cs.1023> (Last

In our work, we have encountered many who cited their surprise or annoyance at places where their own mission creep may be contributing to their overlap with other players in the field. As long as the field rewards low- and no-cost infrastructures for providing “just enough” to get by, it will be nearly impossible for actors to grow their activities into “critical infrastructure” elements. In turn, the field’s many short-lived activities, often spawned by the same set of players, will be too weak to build upon, paralyzing the field at an early state of formation.

As long as the field rewards low- and no-cost infrastructures for providing “just enough” to get by, it will be nearly impossible for actors to grow their activities into “critical infrastructure” elements.

3. Grow the market

Emerging fields do not spring up in order to meet the needs of a new market.⁴⁰ Instead, they emerge from activities (often funded in short-term bursts) that may indicate that there is enough shared need to warrant *cultivating and building* a market. This is an essential step, and one that stymies many actors who assume they know their users’ needs and wants (present or projected future), and believe that those users voluntarily will financially support their operational costs as they grow.⁴¹ Even if/once they have users and know their needs well, they will need to spend time converting those users to customers. This often includes identifying the appropriate procurement channels and agreement formats to support purchasing as well, especially if the user does not manage a budget. This real, complicated work of creating a product or service and aligning it with the financial norms and other products and services in the field is a separate element that is often overlooked to the detriment of business development. These are necessary steps for any ongoing “business,” regardless of its type or sector (e.g., for profit, not-for-profit, or otherwise).

That is to say, the vast majority of the actors in research software do not have an established market, or a set of known and structured interactions, to depend on as they

accessed 02-17-2025), and US National Institutes of Health (NIH) Best Practices for Sharing Research Software: <https://datascience.nih.gov/tools-and-analytics/best-practices-for-sharing-research-software-faq> (Last accessed 02-17-2025).

⁴⁰ “Market” carries connotations of commercial enterprise and Capitalism; we use this term in a different way, to mark the system of structured interactions between defined players to support a set of known activities.

⁴¹ In a fiscal climate where the kind of “good-will generosity” from academic and research players that once funded some “free” offerings through membership fees and other formats is becoming more fraught and less popular, the voluntary contribution model is failing. Fundraising based on this “goodwill” expectation is still popular today, but is also increasingly unrealistic. It is also not yet replaced by other models, leaving a sizable hole in many academic/research open infrastructures today.

conduct their work. The emerging products in research software may have an established user base. They may even have a few paying “customers,” usually in the form of research labs or universities. But most of their funding stems from several key investors (most prominently, Sloan Foundation, Chan Zuckerberg Initiative [CZI], and NASA [US] as well as government and philanthropic funders more broadly), each of whom makes significant investments, typically on limited-term cycles, largely centered on innovations or creative enterprises, and often funding the same small group of actors as they pursue a range of new initiatives. This combination of factors is not conducive to supporting the full software lifecycle span, or to developing or embedding the culture changes needed to mature a field.

Philanthropic and government funders are “usual” stakeholders and investors in research domains and are tremendous support structures that have sometimes provided funding for a decade or more of a project’s development. Their funding waxes and wanes based on trends and politics and internal priorities that are not controlled by their program officers. It is a critical market player in research software today, but it is not a dependable, long-term source of funding.⁴² Particularly in the current, quickly changing funding landscape of 2025 in the US, EU, and UK (among other locations), philanthropic and government funding cannot be considered “stable” for most initiatives, even over shorter periods of time.

We recommend that research software actors (including funders) in this maturing field explicitly advocate for a different use of available philanthropic and government funding right now: to make both collective (on behalf of cohorts of institutions) and organizational investments that provide operational support, concrete training opportunities and/or strategic support from consultants, and that enable them to establish a viable longer-term, diversified plan for funding their work.

This should include:

- **Laying early groundwork to level-set expectations that a service or product is not “free” to create even if it is made freely available to users.** Someone, somewhere, always has to pay for the work. Because so much research software work is initially funded by subsidies coming from grants, contracts, or home-institution investments, its creators and managers currently tend to offer it for “free” without concretely acknowledging the need for future funding sources to support ongoing maintenance work. We note that it becomes harder over time to charge prospective users/customers for what they are used to receiving for free, especially if the cost of creation and maintenance are opaque or hidden.

⁴² Though the focus of this report is on open infrastructure, industry and commercial players in the broader software space and collaborations between academia and industry may provide some opportunities for partnerships and funding, such as GitHub Sponsors.

- **Investing time and funding in key “scaffolding” actors that enable them to explore, develop, test, and implement market strategies.** While there are increasing numbers of “sustainability” grants given to entities that are trying to transition, these are rarely enough to cover both the operational costs *and* the new costs of creating concrete development workflows, product-market fit studies that explicitly define their customer segments and value propositions, and initial, well-defined pricing and service structures. These also rarely include coaching and training or the hiring of new staff who can provide skills that are largely missing from the toolkits of visionary founders and the set of researchers, students, and users that might initially work for or with a new entity.
- **Investing in studies with prospective “customers” to understand the procurement and purchasing dynamics⁴³ within which they must fit.** Often, this will first require defining and understanding the user, and then identifying with that user 1) what other users might be in their organization, and 2) who within their organization might provide financial support for the research software infrastructure elements they need to use in their work.
- **Addressing the needs around training, recognition, community building** and other research software functions is already well underway but there is a “chicken and egg” aspect to providing such services and any underlying financial support for them. The variety of initiatives developed in response to needs identified by practitioners does not equate to or operate like a marketplace of goods and services with, for example, product management, marketing and a diverse set of “investors.”
- **Negotiating for more consistency and co-construction between social and technical infrastructure players regarding definitions, policies, goals, costs and workflows.** The current preponderance of slightly different terms, definitions, and documentation created by both individuals and organizations in the evolving research software field is distracting and confusing, compromising advocacy work and making it harder, rather than easier, for newcomers and outsiders to make sense of this space. A concentrated effort to streamline existing approaches would give the field more coherence; it would help to identify real knowledge gaps (e.g., cost/funding information) and save time wasted recreating resources.
- **Building cohorts that enable entities to tackle this in a collective enterprise rather than as completely separate ventures.** Previous work by the Mellon

⁴³ For example, “research software” right now falls into a gap between several groups in a research lab or university setting - namely, the library, IT, and departments/labs, none of whom have predictable or recognizable procurement channels for research software support structures. Advocating to Vice-Provosts of Research or another high-level group for funding to support this function may be an important step; pairing up across institutions, or even with another field like research data that fights some of the same battle, could also advance this work.

Foundation⁴⁴ with the Nonprofit Finance Fund⁴⁵ is an example of how targeted interventions designed for a particular moment of development can provide both the mechanisms and the community to support change at the market level.

- **Bringing to bear the lessons from adjacent markets and fields (especially beyond “research” arenas) on the market-building process.** This might include looking carefully at some of the successful experiments underway in the broader open source ecosystem and in other “mixed markets” that include different sectors (private, nonprofit, academic, government) and a range of business forms (e.g., sole proprietorships, partnerships, corporations, LLCs, and cooperatives).

4. Invest in coordination

No player in research software (and its adjacent and related fields) can afford to stand alone, particularly in the fraught liminal space between field emergence and formalization. Research software practices, scaffolding, and market structures are not well established, and the field is still in relatively early stages of formation during a troubled fiscal and political moment for “open science” and “open research” more generally.

We recommend this as a prime moment for philanthropic funders to make concentrated investments in the architecture and actualization of the field itself. Such investment could magnify the impact of other investments made in projects, programs, and organizations in research software. It would provide the necessary space and working time to undertake well structured, facilitated engagement between some of the identified “key players” from each developing channel or function of the field (as discussed above, e.g., training, packaging, hosting, socialization, advocacy, etc.) and from across boundaries of discipline and geography. These groups could be provided with space, time, tools and structure, potentially aided by in-kind support from institutions, to help them to build trust, buy-in and momentum to establish a formal agenda to guide the field’s development over the next five years.⁴⁶

We posit that strong, consistent facilitation needs to be provided by an entity(ies) without “skin in the game,” ideally an organization that is not already embedded in the maturing field. This group could serve a temporary, time limited “backbone” function that assists

⁴⁴ <https://www.mellon.org/> (Last accessed 02-17-2025).

⁴⁵ <https://nff.org/> (Last accessed 02-17-2025).

⁴⁶ While an increasing number of forums and events already enable various subgroups to gather to share information about topics of shared interest, these lead only to slow discovery of and planning toward potential collaborative activities that often die on the vine due to lack of funding.

the key players in first co-defining a vision⁴⁷ (not a platitude, but a measurable, achievable future state) and then explicitly pursuing and tracking progress toward shared goals that lead to vision fulfilment. A shared agenda would not demand altruism and additional volunteer work from research software actors; rather, it would help these entities hone in on points of alignment between their individual work (as organizations, projects, and/or programs) and the vision and goals set in the agenda. Concretizing these with formal commitments and supporting these with regular check-ins and measurements of progress by the “backbone” would promote visibility of the work and an ongoing building of trust in the agenda and in each other.

To that end, our recommendations include:

- **Use funding to incentivize the key players to build trust, explore working together, and accomplish and celebrate ongoing accomplishments.** Even the most generous funding in this field is typically piecemeal and limited to particular organizations or projects in time. We need an additional layer of funding that focuses explicitly on establishing trust and interdependence. Establishing shared goals and working together to achieve them is too hard to do without steady funding.
- **Build a more ecosystem-level view of funding that helps funders coordinate on funding or shared goals.** Duplication of work is rampant in research software, and much of this is happening due to the lack of optics across the varied projects and programs that philanthropic and government funders are supporting. The most successful venture will still ultimately fail if it overlaps too significantly with other efforts in the field – they each undercut each other, making it hard for any of them to survive, let alone thrive.
- **Encourage actors to rely on each other and to build the necessary service environments.** Trust is essential to this process, and engaging in a concrete set of meaningful work that is co-owned by the group (and that is additive to rather than distracting from their individual aims) will reinforce this trust and help establish an environment for specialization and ongoing intentional collaboration.
- **Where feasible, encourage joint infrastructures for research software with those in adjacent fields like research data.** There are shared infrastructure needs across these fields that could help both to develop more efficiently; there are also shared challenges across these two fields that warrant more collaborative work. Given how closely linked research software and data are and how important both are to

⁴⁷ This work would build on the relatively strong coherence we heard around the “visions” listed in our findings above (specifically, elevating research software to a first-class research object, making research software more discoverable and reusable, and recognizing contributors and functions across the research software lifecycle). We think that a shift from platitude to measurable vision is achievable.

research integrity, we see many potential coordination and partnership opportunities between the research software and data ecosystems.⁴⁸ Funders could help incentivize and facilitate mutually beneficial collaborations between these groups. Without that additional funding incentive and support, collaboration is a volunteer effort, one that can only yield success if all partners have sufficient means to cover the expense and insecurity that comes with working together.

- **Explore resourcing (funding, services and labour) models that allow for sharing** across networks or nodes. These can support a diversity of community-based infrastructures without each needing to hire, build, license or maintain its own set of operational and technical resources. A shared resources approach mimics open source contributions and adaptations, but at the operations and financing levels.

Our recommendations are largely ongoing and mutually reinforcing. For example, coordination and community building need to be continuous, though of course the focus and intensity of the work will vary, as networks, processes, and norms are established and evolve. In a few cases, we see some preferred sequence to the recommendations. Tracking and reporting provides necessary data in our view, without which decision making on where to invest, for example, is more difficult and less likely to succeed in its goals. Having data at hand while also building and bridging communities helps facilitate collaborating on shared agendas. In parallel with all of this work, exploring and growing the “market” for research software supports recognition and capacity building for its infrastructures. These efforts are not meant to be linear, predominantly sequential or entirely distinct but having foundations for this work established (and ongoing) provides the best environment for developing shared services, which is what we recommend to best support this growing field.

Additional Observations

As this report is a preliminary investigation, a number of areas emerged in our research that would benefit from further exploration. This and similar landscape reviews necessarily focus on core functions and organizations/initiatives. Much of the proposed work in this section is broader or more granular; from the highest system levels and the most granular disciplinary and individual levels.

Reinforcement vs overcrowding

Even within the relatively small ecosystem of research software organizations and initiatives, multiple efforts perform similar functions, for example, training. This overlap, along with the challenge of organizational sustainability common to research supporting

⁴⁸ There are already some examples. The Research Data Alliance (RDA), for example, hosts software groups, including the FAIR for Research Software (FAIR4RS) Working Group and the Software Source Code Interest Group, which houses the FAIR4RS principles.

infrastructures, are key reasons for investigating the potential for shared services in this space. What is a healthy, sustainable and inclusive level of backup or overlap in the ecosystem? When does that cross over into overcrowding and unnecessarily duplicated effort? Who gets to decide?

Starting to answer these complex, system-level questions involves weighing the relative merits of mutually reinforcing work, buffering against the potential for single (or a few) points of failure, reflecting the diversity of research across disciplines and geographies and addressing the nature of organizations to be self-perpetuating. On this last point, it is worth noting that several interview participants raised the need for thoughtful, supportive assistance in helping to identify and sunset research software that is approaching the end of its lifecycle, and to recognize and support the individuals involved in this work. This is not to make the case for identifying or shuttering organizations. It is to recognize practical limitations in the ecosystem and to clearly signal that the enormous effort that goes into supporting research software across the lifecycle deserves recognition and reward, including at the end of its cycle.

The human infrastructure

A closely related topic for future work reflects a common refrain heard in our interviews: how to provide recognition, reward and support for the human infrastructure that developed and that powers this ecosystem, particularly as the landscape changes over time. Some may be well positioned to support shared services, for example. Are there cross-initiative strategies that might work for guarding against “institutional memory” loss, providing a “parachute” of sorts in cases of staff reductions or just supporting those looking to contribute their expertise in new ways?

A variation on this could be to determine what community vetting means at the system level, i.e. how much input, on something like a proposed standard, for example, is enough? What steps would improve feedback from individual community members? What are strategies to identify and combat “community washing”⁴⁹?

Intentional vs accidental infrastructure

Not all infrastructures meant to become infrastructure or originally thought of their work as “infrastructure.” The idea of “accidental infrastructure” came up in our interviews (and in an unrelated interview⁵⁰ with the founder and CEO of SciCrunch) and is worth exploring in more detail, both to help identify and support interdependencies and to better

⁴⁹ <https://copim.pubpub.org/pub/copim-perspective-on-community-led#why-does-community-led-matter> (Last accessed 02-17-2025).

⁵⁰ <https://scholarlykitchen.sspnet.org/2024/11/14/kitchen-essentials-an-interview-with-anita-bandrowski-of-sicrunch/> (Last accessed 02-17-2025).

understand sustainability approaches as organizations and initiatives transition to infrastructure, intentionally or not.

The role of institutions

Setting aside the increased understanding that systematically recognizing the work of research software can provide, as well as the potential for linking between functions like OSPOs, research administration, institutional repositories and the actual work of research software, institutions would seem to have incentives to give visibility to more of their faculty's outputs. Given the role of research funding for institutions, and the current fragility of both in the US especially, it may be instructive to explore the potential institutional role(s) in, and barriers to, providing better recognition of and support for research software.

The long tail

In research software, as in other ecosystems, there are some major players with large communities behind them. But there is also a long tail of niche software used by very small groups (or individuals) and visibility of the full, diverse range of research software is necessary to understand needs, resourcing and impact on infrastructures. Understanding strategies for situating the long tail in the overall funding and recognition of research software may be useful generally and in pursuit of more diverse, inclusive and equitable support for research.

Analyses of all of these issues must be tailored to particular landscapes and contexts but, at a high level, these questions are valuable explorations for all areas of open infrastructure.

Conclusion

Research software is a key function of contemporary research but it is not yet supported or recognized as such, at least not at a level that reflects its ubiquity and contributions across disciplines. The field has made great strides to rectify this gap over the last 20 years, but it is still maturing. ***Particularly given that we are entering into a critical moment of defunding in open science, we highly recommend shifting funding and investment attention toward immediate coordination of actors in this field, helping them to come together, make choices and commitments, and build an ecosystem of interdependent infrastructure to support research software.*** With greater coordination and targeted funding, this ecosystem can avoid some of the pitfalls of more established, adjacent environments. Being intentional and coordinated in advocating for policy and building and sharing infrastructures can, for example, reduce unnecessary duplication of effort and increase buy-in for shared solutions. Establishing flexibility and resilience in the ecosystem can help put research software into the position its champions have long

worked for: that of a first class research object, well supported, visible, used, and maintained to the benefit of science, social science, and the humanities.

Acknowledgments

We would like to thank all interview participants for sharing their experiences and insights with us, including: Alice Allen, Astrophysics Source Code Library (ASCL) / University of Maryland; Elton Barker, The Pelagios Network; Michelle Barker; Neil Chue Hong, Software Sustainability Institute / University of Edinburgh; Roberto Di Cosmo, Juan Pablo Flores, GitHub; Ralf Gommers, Quansight Labs; Morane Gruenpeter; Hilary Hanahoe, Research Data Alliance; Daniel S. Katz, University of Illinois Urbana-Champaign; Allen Lee; Claudia Bauzer Medeiros, University of Campinas, Brazil; James Munroe, 2i2c; Fotis Psomopoulos, Institute of Applied Biosciences (INAB) / Centre for Research and Technology Hellas (CERTH); Noam Ross, rOpenSci; and Anelda Van der Walt, Talarify.

This work is funded by The Alfred P. Sloan Foundation (Grant ID: G-2024-22595).

Additional information about how the work of IOI is funded can be found here:

<https://investinopen.org/about/how-were-funded/>.

The authors also thank: Elton Barker, Michelle Barker, Neil Chue Hong, Juan Pablo Flores, Daniel S. Katz, Allen Lee, James Munroe, Fotis Psomopoulos, Noam Ross, and Anelda Van der Walt for reviewing a draft of this paper; their insights helped to strengthen this final version and any/all errors and shortcomings are solely the responsibility of the authors. We also thank IOI senior researcher Gail Steinhart for preparing an extensive collection of desk research sources and notes (and reviewing the paper).

In addition, IOI formally recognizes the importance of research software through inclusion of research software infrastructures in the Infra Finder tool.⁵¹ As this report was being finalized, eight organizations had been added with a ninth pending.

Appendix A: Bibliography

de los Ángeles Orfila, M. (2024). 'Scienticide': Argentina's science workforce shrinks as government pursues austerity [Dataset]. <https://doi.org/10.1126/science.zjcq1nt>

Arp, L., & Forbes. (2018, February). Guidebook—It Takes A Village.
<https://itav.lyrasis.org/guidebook/>

Barcelona Declaration on Open Research Information, Kramer, B., Neylon, C., & Waltman, L. (2024). Barcelona Declaration on Open Research Information.
<https://doi.org/10.5281/ZENODO.10958522>

⁵¹ <https://infrafinder.investinopen.org/solutions> (Last accessed 02-17-2025).

- Barker, M., Chue Hong, N. P., Katz, D. S., Lamprecht, A.-L., Martinez-Ortiz, C., Psomopoulos, F., Harrow, J., Castro, L. J., Gruenpeter, M., Martinez, P. A., & Honeyman, T. (2022). Introducing the FAIR Principles for research software. *Scientific Data*, 9(1), 622. <https://doi.org/10.1038/s41597-022-01710-x>
- Bourdieu, P. (1993). *The field of cultural production: Essays on art and literature* (R. Johnson, Ed.). Columbia University Press.
- Chue Hong, N. P., Katz, D. S., Barker, M., Lamprecht, A.-L., Martinez, C., Psomopoulos, F. E., Harrow, J., Castro, L. J., Gruenpeter, M., Martinez, P. A., Honeyman, T., Struck, A., Lee, A., Loewe, A., van Werkhoven, B., Jones, C., Garijo, D., Plomp, E., Genova, F., ... RDA FAIR4RS WG. (2022). FAIR Principles for Research Software (FAIR4RS Principles) (1.0). Zenodo. <https://doi.org/10.15497/RDA00068>
- Eghbal, N. (2016, July 14). *Roads and Bridges: The Unseen Labor Behind Our Digital Infrastructure*. Ford Foundation. <https://www.fordfoundation.org/work/learning/research-reports/roads-and-bridges-the-unseen-labor-behind-our-digital-infrastructure/>
- Garijo, D., Ménager, H., Hwang, L., Trisovic, A., Hucka, M., Morrell, T., Allen, A., Task Force on Best Practices for Software Registries, & SciCodes Consortium. (2022). Nine best practices for research software registries and repositories. *PeerJ Computer Science*, 8, e1023. <https://doi.org/10.7717/peerj-cs.1023>
- Gruenpeter, M., Katz, D. S., Lamprecht, A.-L., Honeyman, T., Garijo, D., Struck, A., Niehues, A., Martinez, P. A., Castro, L. J., Rabemanantsoa, T., Chue Hong, N. P., Martinez-Ortiz, C., Sesink, L., Liffers, M., Fouilloux, A. C., Erdmann, C., Peroni, S., Martinez Lavanchy, P., Todorov, I., & Sinha, M. (2021). Defining Research Software: A controversial discussion. Zenodo. <https://doi.org/10.5281/ZENODO.5504015>
- Hettrick, S. (2016, August 17). A not-so-brief history of Research Software Engineers | Software Sustainability Institute. <https://www.software.ac.uk/blog/not-so-brief-history-research-software-engineers-0>
- Hopkins, K., Adema, J., Barnes, L., Deville, J., Fathallah, J., Gatti, R., Hughes, A., Sanders, K., & Steiner, T. (2024). The Copim perspective on Community-Led. Copim. <https://doi.org/10.21428/785a6451.cc2195f3>
- Kemp, J., & Tsang, E. (2024). Interview Protocol—Research Software. <https://doi.org/10.5281/ZENODO.14141211>
- Kemp, J., & Tsang, E. (2024). [IOI] Research Brief: Community Infrastructure to Further Open Research Software. <https://doi.org/10.5281/ZENODO.14178127>
- Knowles, R., Mateen, B.A. & Yehudi, Y. We need to talk about the lack of investment in digital research infrastructure. *Nat Comput Sci* 1, 169–171 (2021). <https://doi.org/10.1038/s43588-021-00048-5>

- Lippincott, S., & Skinner, K. (2022). FOREST Framework for Values-Driven Scholarly Communication. <https://doi.org/10.5281/ZENODO.6557301>
- Meadows, A. (2024, November 14). Kitchen Essentials: An Interview with Anita Bandrowski of SciCrunch. The Scholarly Kitchen. <https://scholarlykitchen.sspnet.org/2024/11/14/kiitchen-essentials-an-interview-with-anita-bandrowski-of-sci crunch/>
- National Institutes of Health (NIH). (n.d.). Best Practices for Sharing Research Software | Data Science at NIH. Retrieved January 27, 2025, from <https://datascience.nih.gov/tools-and-analytics/best-practices-for-sharing-research-software-faq>
- Office of Science and Technology Policy. (2022, August 25). *Ensuring Free, Immediate, and Equitable Access to Federally Funded Research*. <https://bidenwhitehouse.archives.gov/wp-content/uploads/2022/08/08-2022-OSTP-Public-Access-Memo.pdf>
- Open Science NL budget cut by half: "The transition to open science is as important as ever" | Open Science NL. (2024, September 17). <https://www.openscience.nl/en/news/open-science-nl-budget-cut-by-half-the-transition-to-open-science-is-as-important-as-ever>
- Pain, E. (2024). Belt-tightening budget derails France's multiyear research funding plan [Dataset]. <https://doi.org/10.1126/science.z6dcmg8>
- Research Software Alliance. (2024). Amsterdam Declaration on Funding Research Software Sustainability (1.1). Zenodo. <https://doi.org/10.5281/zenodo.13735888>
- Saenen, B. (2025). Developing and Aligning Policies on Research Software: Recommendations for Research Funding and Research Performing Organisations. Zenodo. <https://doi.org/10.5281/ZENODO.13740998>
- Sochat, V., May, N., Cosden, I., Martinez-Ortiz, C., & Bartholomew, S. (2022). The Research Software Encyclopedia: A Community Framework to Define Research Software. *Journal of Open Research Software*, 10(1), 2. <https://doi.org/10.5334/jors.359>
- Strasser, C., Hertweck, K., Greenberg, J., Taraborelli, D., & Vu, E. (2022). Ten simple rules for funding scientific open source software. *PLOS Computational Biology*, 18(11), 1010627. <https://doi.org/10.1371/journal.pcbi.1010627>
- Thaney, K., & Skinner, K. (2024). Grant proposal: Community infrastructure to further open research software: A landscape analysis of needs, dependencies, and opportunities. Zenodo. <https://doi.org/10.5281/ZENODO.13961596>
- UNESCO. (2021). UNESCO Recommendation on Open Science. UNESCO.

<https://doi.org/10.54677/MNMFH8546>

The full list of resources used in this project is available in IOI's public Zotero library.⁵²

Appendix B: Interviews brief

The project team interviewed 21 people across 19 interviews. This interviews brief is a companion to the research project brief.⁵³

1. Summary

At Invest in Open Infrastructure (IOI), we are working to understand the landscape of research software and to identify the gaps, challenges, and opportunities for advancing its sustainability and resilience. This brief is one output of a broader project (funded by the Sloan Foundation, October 2024-January 2025) in which we conducted desk research and interviews to inform a set of findings and recommendations regarding the health and future development opportunities for this rapidly growing landscape.

Research software as a primary output in its own right has gained attention in recent years, particularly as the focus on data, spurred in part by mandates like the US OSTP Nelson memo,⁵⁴ has also increased. We are witnessing a process of formalization and institutionalization of this open science-related field, with the development of best practices (e.g. FAIR Principles for Research Software), newly recognized roles (e.g., Research Software Engineer), thriving forums (e.g. US-RSE, Software Sustainability Institute, Journal of Open Source Software), and functions both in and beyond the academic community. With this process, there is growth in initiatives, organizations, and services — or “infrastructures” — working to support the development, maintenance, discovery, sharing, reviewing, and citation of research software, with the aim of furthering open, reproducible, and equitable research. Many of these infrastructures are community-governed and/or open to varying extents; this signals that the research community's constituents are involved in steering the growth and evolution of the research software field. Our research seeks to understand the degree to which existing research software infrastructure is open and what might help this ecosystem to thrive and sustainably develop. Though this infrastructure is the focus of this work, desk research for the project involved a look across the entire landscape of related issues and players.

⁵² https://www.zotero.org/groups/4377072/invest_in_open/collections/JX8XY6G4 (Last accessed 02-17-2025).

⁵³ Kemp, J., & Tsang, E. (2024). [IOI] Research Brief: Community Infrastructure to Further Open Research Software. Zenodo. <https://doi.org/10.5281/zenodo.14178127> (Last accessed 02-17-2025).

⁵⁴ <https://bidenwhitehouse.archives.gov/wp-content/uploads/2022/08/08-2022-OSTP-Public-Access-Memo.pdf> (Last accessed 02-17-2025).

This brief summarizes our interviews in the context of the broader project, which includes a full report and [webinar](#) on our preliminary findings and recommendations (23 January 2025).

2. Research objectives

We want to understand what is happening to date in community infrastructure for research software across disciplines. Our guiding objectives were to identify:

- Existing connections between different research software infrastructure initiatives, and where new connections can be built to advance sustainability and resilience of the ecosystem
- Areas in the ecosystem where there is overcrowding/excessive functional overlap, and where there are gaps, blockers, vulnerabilities, and/or critical dependencies that need to be addressed
- Indications of what helps infrastructure services and initiatives succeed in adding value to the research software ecosystem

2.1 Qualitative approach

We followed a qualitative approach combining desk research and analysis, and community engagements. We conducted 19 interviews with key stakeholders, identified in part from the desk research we conducted. Interviews serve, in part, to validate, correct and/or round out information gathered from desk research.

3. Desk research

To inform the project direction and approach, IOI senior researcher Gail Steinhart prepared an extensive collection of desk research sources and notes, including identifying:

- Relevant organizations, services, and initiatives, categorized by type, for example, documentation and workflow tools, software publishers and repositories and communities of practice
- A core, annotated bibliography of more than 50 sources, to which additional sources and notes were added by project team members during the project.

The desk research aided in identifying potential interview participants and developing the interview instrument and questions.

3.1 Desk research findings summary

Several themes, areas of agreement, and underserved needs surfaced through the desk research (these are more fully described in the research brief⁵⁵ and in this report).

- There is broad agreement on and efforts toward the need to recognize and reward research software as a primary research output
- The research software community is converging on principles and best practices to ensure that the development of research software advances open, collaborative science
- The lack of common, formal recognition for this work has been a disincentive for busy researchers, who generally have had no/little formal software development training, to take the (often considerable) time necessary to prepare software for sharing and (re)use
- Funding and sustainability concerns are rampant and unresolved
- Popular, convenient options for sharing that may be free to use are not necessarily "open"
- Software citation is a challenge for researchers and publishers
- Research software includes a variety of concerns and needs that are distinct from but situated within the broader context of:
 - Software discovery and use (including the extent to which software may or may not be designed for reuse, and software metadata standards and management)
 - Software preservation
 - Data generated or analyzed by software
 - Differentiating between code (e.g. for statistical analyses), vs. full-fledged software and other stages of research software work
 - Maintenance and open source software management
 - Research reproducibility
 - Concerns about equity, diversity, inclusion, and accessibility
 - Community building and engagement

4. Interviews

- Complement desk research
- Inform follow up research, including additional desk research resources and potential additional interview participants
- Provide current information and personal perspectives
- Are confidential
- Are coded and summarized to inform the research report

⁵⁵ <https://doi.org/10.5281/zenodo.14178127> (Last accessed 02-17-2025).

4.1 Instrument and topics

We developed a semi-structured interview protocol⁵⁶ informed by the desk research to facilitate the interviews. It includes five main sections, with final question sets selected for each. To fit the available time and to tailor the discussion to the role(s) of the individuals and initiatives involved, the actual questions asked in each interview were a subset of the full protocol, summarized here:

- Grounding questions
 - Participants' motivations for their involvement in research software
 - Participants' definition(s) of research software
- Big picture questions
 - State of the ecosystem
 - Major players and foundational infrastructure
 - Pros and cons of overlapping efforts
 - Gaps in the landscape
- Questions on the individual researcher/discipline perspective
 - How various functions (e.g. software discovery, review, citation) do and should work
 - Key challenges
 - Training and support needs
 - Support role(s) of institutions
- Questions on the potential for shared needs and services
 - How niche/granular efforts need to be to improve recognition and sustainability
 - What efficiencies might exist if organisations and services work together more formally
 - Key landscape vulnerabilities
 - Barriers to collaboration
- Wrap-up and reflecting questions
 - Suggestions for other interview participants
 - What questions we should have asked but did not
 - Common misperceptions about research software
 - Whether or how research software overlaps with open source, open access and/or open science
 - Identifying and crediting various research software roles

4.2 Interview findings

Interview participants, most from academia and industry, represent a range of organisations and initiatives that focus on a variety of research software-related efforts,

⁵⁶ <https://doi.org/10.5281/zenodo.14141211> (Last accessed 02-17-2025).

including training, technical solutions, advocacy, convenings and more, across discipline types.

There is general agreement among participants that, while still in its early days, the research software community has made gains in elevating the profile of research software and supporting researchers working with it.

Strengths identified and observed include:

- Deep expertise, commitment and a practical mindset within organizations and initiatives working to improve research software. For example:
 - Recognition of the need for greater maturity in the system and both near- and longer-term solutions and strategies
 - Interest in learning from the work of related communities
- Efforts are largely grassroots and hands-on
- Interest in and progress toward developing best practices
- Agreement on areas that have the potential for shared services (explored in more detail in our forthcoming report), such as marketing
- The role of research software engineers (RSE) has emerged as its own profession, supporting researchers in some areas and generally indicating a recognition of expertise involved in doing the work

Participants reflected broad agreement on the challenges they see:

- Funding has improved but remains a challenge:
 - funders do not always have a good grasp of the landscape or understand the work in enough detail and do not coordinate on their approaches and strategies
 - legal barriers to cross-border funding inhibit broader representation and support as well as a more international approach that would better reflect how research is conducted
- The various stages that comprise the work of research software are not addressed evenly or often treated as part of an overall lifecycle. For example:
 - maintenance is notoriously underfunded and undervalued and other ongoing work that can be difficult, such as metadata, does not get sufficient attention or resourcing from various stakeholders
 - similarly, preservation and technical sustainability are exceptions to what is otherwise a generally healthy current technical state
 - product management-type functions, for example, user experience work, is lacking
- Resources to improve engagement and participation, for example community managers, are a frequently cited need

- There is a need for greater awareness of where expertise and resources exist, to improve work within the landscape as well as for greater recognition outside of it

In addition, participants frequently raised issues that may be a mix of challenges and opportunities and/or that warrant further exploration:

- There is significant interest in increasing and improving collaboration among advocates, practitioners and organizations working in this space that may need more or different resources and strategies
- There is a long tail of software that is often discipline-specific. How to give visibility and support to this work and whether or how to factor such niche software into broader approaches came up frequently
- Attention is needed on how best to support the human infrastructure working with research software. Examples include:
 - how to approach sunsetting software
 - potential changes brought by Artificial Intelligence (AI)
 - overall resourcing and the need to avoid burnout

There are some areas where views or priorities diverge or that are less clear overall:

- Open source: The extent to which research software should overlap or align with open source (and indeed open science) is somewhat murky or perhaps seen as secondary to the primary work of research software
- Broader participation: Though there is considerable interest in increasing collaboration and frequent observations about the lack of diversity in the community, participants indicated differences in priorities and practicalities of increasing representation
- Financial support: While funding is an area of near-universal concern, there are some differing perspectives on the relative merits of relying on volunteers versus paid staff

The landscape is rapidly developing and these findings are broad strokes that primarily reflect the landscape in the US, Australia and Europe. We caution readers from extrapolating too much to other regions or too far into the future.