People Roadmap report

Table of Contents

List of acronyms

AGU - American Geophysical Union
ARDC - Australian Research Data Commons
CESSDA - Consortium of European Social Science Data Archives
CHAOSS - Community Health Analytics Open Source Software
CLARIN - Common Language Resources and Technology Infrastructure
CSCCE - Center for Scientific Collaboration and Community Engagement
CS&S Code for Science & Society
DANS - Data Archiving and Networked Services
DARIAH - Digital Research Infrastructure for the Arts and Humanities
DEI - Diversity, equity and inclusion
DLR - German Aerospace Center (Deutsches Zentrum für Luft- und Raumfahrt)
DOE ECP - Department of Energy Exascale Computing Project
ERIC - European Research Infrastructure Consortiums
EURISE Network - European Research Infrastructure Software Engineers' Network
EOSC - European Open Science Cloud
FAIR - Findable, accessible, interoperable, reusable
FORCE11 SCIWG - Future of Research Communications and e-Scholarship Software Citation Implementation Working Group
INTERSECT - INnovative Training Enabled by a Research Software Engineering Community of Trainers
JORS - Journal of Open Research Software
JOSS - Journal of Open Source Software
NeSI - New Zealand eScience Infrastructure
NLeSC - Netherlands eScience Center
NWO - Dutch Research Council
OECD - Organisation for Economic Co-operation and Development
ReSA - Research Software Alliance
RSE - Research Software Engineering
RSEs - Research Software Engineers
SSI - Software Sustainability Institute
UKRI - UK Research and Innovation
UNESCO - United Nations Educational, Scientific and Cultural Organization
URSSI - US Research Software Sustainability Institute
US-RSE - US Research Software Engineers Association
VSNU - Association of Universities in the Netherlands

1. Introduction

In response to the evolution of a range of research and software areas, including 1) the rise of open science (which includes open software), 2) increased understanding of the need for advanced digital skills in the research community (including research software engineering) to achieve the aims of open science, and 3) the development of the Research Software Engineering (RSE) movement to recognise and support the Research Software Engineers (RSEs) who are responsible for the development and maintenance of a significant amount of research software, the Research Software Alliance (ReSA) has created a People Roadmap activity to facilitate identification and opportunities for accelerating efforts to address major issues related to people (or personnel) challenges faced by the research software community. This People Roadmap aims to increase understanding of how to create an environment where research software personnel are recognised, have appropriate skill sets and access to inclusive communities, within policy and infrastructure environments that support their work. Significant cultural change is needed in the research sector in order to provide both the highly skilled personnel who are needed, and the cultural change needed to support their work.

In the remainder of this introduction, we provide the rationale and aims of this work, followed by methodology we used in Section 2. Section 3 provides a detailed analysis of the results obtained from profiling 28 research software initiatives. Section 4 analyses this data and concludes with some suggestions on next steps.

Focus on making research practices outputs more open has been inspired by initiatives and guidelines such as the San Francisco Declaration on Research Assessment (2013), Leiden Manifesto (Hicks et al., 2015), and Hong Kong principles for assessing researchers (Moher et al., 2020). Cultural change is needed to research practice more widely, including the inclusion of appropriate incentives for open research in the reward and evaluation system for research staff. Whilst there are moves from a range of governments, funders, research institutions, and publishers, to address these (VSNU et al., 2019; Woolston, 2021), the inclusion of research software in these efforts is still lagging behind that of research data. The Sorbonne Declaration of Research Data Rights (2020) promotes the sharing and proper use of data, but does not address other research objects such as software. Yet software is pervasive in modern research, and recognition of the efforts of its creators is equally important. The UK Research Software Survey found that more than 90% of 1,000 randomly chosen researchers acknowledged software as being important for their own research, and about 70% of these researchers said that their research would not be possible without software (Hettrick et al., 2014).

There has also been considerable emphasis on the need to improve the digital skills needed to support data-intensive science by organisations such as the European Open Science Cloud (EOSC) and the Organisation for Economic Co-operation and Development (OECD) (European Commission & EOSC Executive Board, 2021; OECD, 2020). However, research software skills comprise a very minor part of these efforts, despite research on the most-cited 100 papers in

one of the leading international science journals, *Nature*, that revealed that the vast majority describe experimental methods or software that have become essential in their fields (Noorden et al., 2014).

Alongside these broader developments across the research sector, there has been an increasing research emphasis on the people undertaking software development. This is partly due to the evolution of the RSE community (Akhmerov et al., 2019; Anzt et al., 2021; Brett et al., 2017; Cohen & Woodbridge, 2020; Hardey & Leng, 2020), but also reflects parallel efforts to increase recognition of the importance of research software (Barker et al., 2020). Recent international policy initiatives such as the G7 Research Compact (2021), the draft United Nations Educational, Scientific and Cultural Organization (UNESCO) Recommendation on Open Science (UNESCO, 2021), and the OECD Recommendation on Access to Research Data from Public Funding (OECD, 2021) to support open science have included an emphasis on research software as a fundamental and vital part of research, alongside more commonly recognised research outputs such as publications and research data. This recognition of the need to equally value research software is long overdue; findings from the 2018 OECD International Survey of Scientific Authors includes evidence that 25% of research produces new code (Bello & Galindo-Rueda, 2020).

Given these developments, and the lack of recognition of research software and it's creators, developers, and maintainers, the Research Software Alliance's (ReSA's) mission is to bring research software communities together to collaborate on the advancement of the research software ecosystem. As outlined in the ReSa Strategic Plan (2021-23), ReSA is supported by top global research infrastructure institutions to function as a backbone organisation across the research sector to increase the community's ability to collectively impact achievement of the shared vision that research software is recognised and valued as a fundamental and vital component of research worldwide. Backbone organisations enable collaboration in a synchronised effort across multiple stakeholders to achieve a common goal. ReSA supports collective impact by providing the social infrastructure needed to foster the cross-organisational communication, alignment, and collaboration required to achieve sectoral change in the research sector.

The People Roadmap is a ReSA community consultation to map the landscape of research software community initiatives related to people (including RSEs), as part of ReSA's key role in the sector to improve collaboration across national and international research software organisations and initiatives.

2. Scope and Methodology

The design of the People Roadmap consultation required definition of the scope of people-related issues, to identify initiatives to profile. The methodology for analysis of this profiling utilised both quantitative and qualitative methods to yield results.

2.1 Scope

This consultation aimed to facilitate strategic discussion across research software initiatives, with the intention of developing a collaborative approach to increasing understanding of the landscape in the following five areas:

- Policy: activities that aim to influence the policy or actions of other stakeholders, not only one's own organisation
- Incentives: ways to improve the establishment, recognition and progression of research software personnel roles
- Communities: groupings that enable people with common interests or characteristics to come together to achieve shared goals
- Skills and training: projects that support skills and training efforts beyond those for their own staff, and include designing curriculum, training trainers, identifying skills frameworks, and convening training and mentoring events
- Infrastructure: the components used to support the role of software for and in research practices, including the schemas and standards that support this

These five themes are based on an adaptation of Nosek's strategy of the five layers that need to be addressed in a comprehensive change strategy (Nosek, 2019), with Noseks' second bottom layer of user interface/experience being replaced by skills and training in Figure 1:

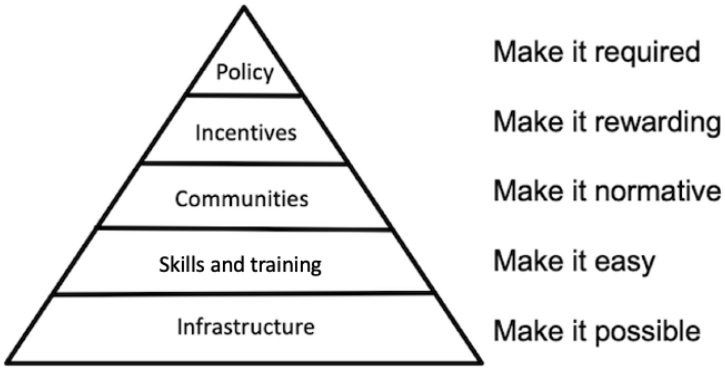

Figure 1: Adaptation of Nosek's strategy for culture change (Nosek, 2019)

This also has some overlap with three of the four pillars of RSE identified by Cohen et al. (2021) as key to providing comprehensive and sustainable support to RSE. The four pillars are software development, community, training and policy; but software development does not align with any of the five themes used here.

2.2 Profiled initiatives

The People Roadmap profiled 28 organisations, community initiatives and/or projects (hereon referred to as initiatives) in the research software community. 17 were originally identified from the ReSA's 2020 analysis of the research software community landscape, which identified 50+ organisations (Katz et al., 2020), with the People Roadmap enabling identification of additional organisations. 28 initiatives were chosen to give a cross-section of different types of organisations, including whose focus was a national/regional or disciplinary area; initiatives focused on a particular topic such as research software sustainability or citation; and initiatives providing a specific service such as a repository or publications. The initiatives also varied in structure, ranging from discrete projects to much larger government-funded initiatives. Some had existed for more than a decade and others were time-limited, or only just beginning. The final set of initiatives profiled is listed in Appendix 1.

2.3 Methodology

Each of the 28 initiatives listed in Appendix 1 was offered an interview within which to answer the questions listed in Appendix 2, and a few initiatives chose to respond to the questions in written form instead. In four instances where it was difficult to identify a lead person to correspond with, the initiative's work was summarised based on public online information.

The People Roadmap utilised Exaptive's Cognitive City software to visualise the profiled initiatives. ReSA is utilising the People Roadmap to pilot the creation of a digital twin of the research software community that maps linkages between strategic goals, projects, collaborators and funders, as a way of addressing the ReSA mission of bringing research software communities together to collaborate on the advancement of the research software ecosystem. The Cognitive City is a virtual environment that maps the initiatives according to the data contained in Appendix 3.

3. Data obtained from profiled initiatives

The 28 profiled initiatives were invited to answer the questions contained in Appendix 2, and their responses are summarised in the following six subsections, which correspond to the order and focus of the questions.

3.1 Missions and visions

All the initiatives were asked to provide their mission and/or vision. There was a lot of variation in the content of the missions and visions, which is to be expected given the breadth of types of initiatives being profiled. Some initiatives had a people-focus as a core element of their mission, whilst for others it was only part of their remit. The Cognitive City enables analysis of key words in a termscape, somewhat like a more complex version of a word cloud. This analysis was run across all the initiatives' visions and missions, with community emerging as the most common term, closely followed by software and developers, then data.

3.2 Strategic goals and projects

The initiatives were asked to provide about their strategic goals and projects, and this information has been utilised to ascertain which of the five people themes the initiative included foci on. The answers given for strategic goals and projects have been combined as it was often difficult to distinguish between the two (particularly for smaller initiatives). Figure 2 shows the results.

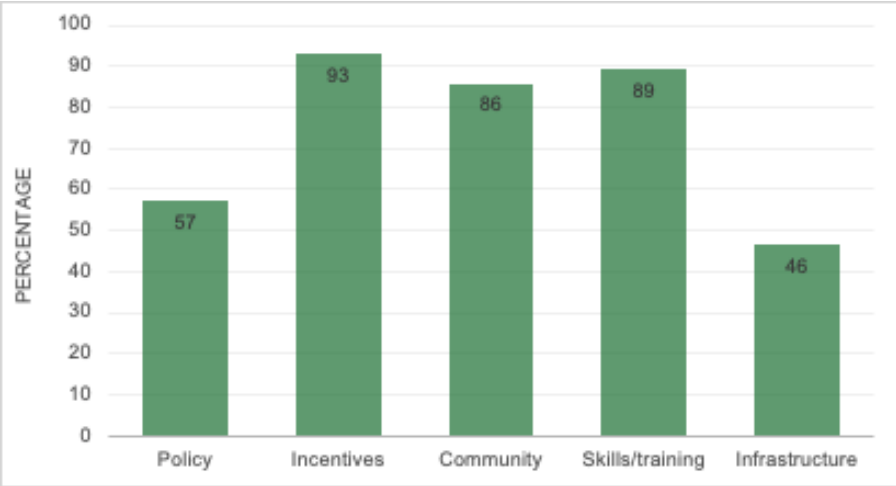

Figure 2: Percentage of initiatives addressing each theme

Figure 3 also shows which initiatives correspond to each theme. This data is also shown in table form in Appendix 3.

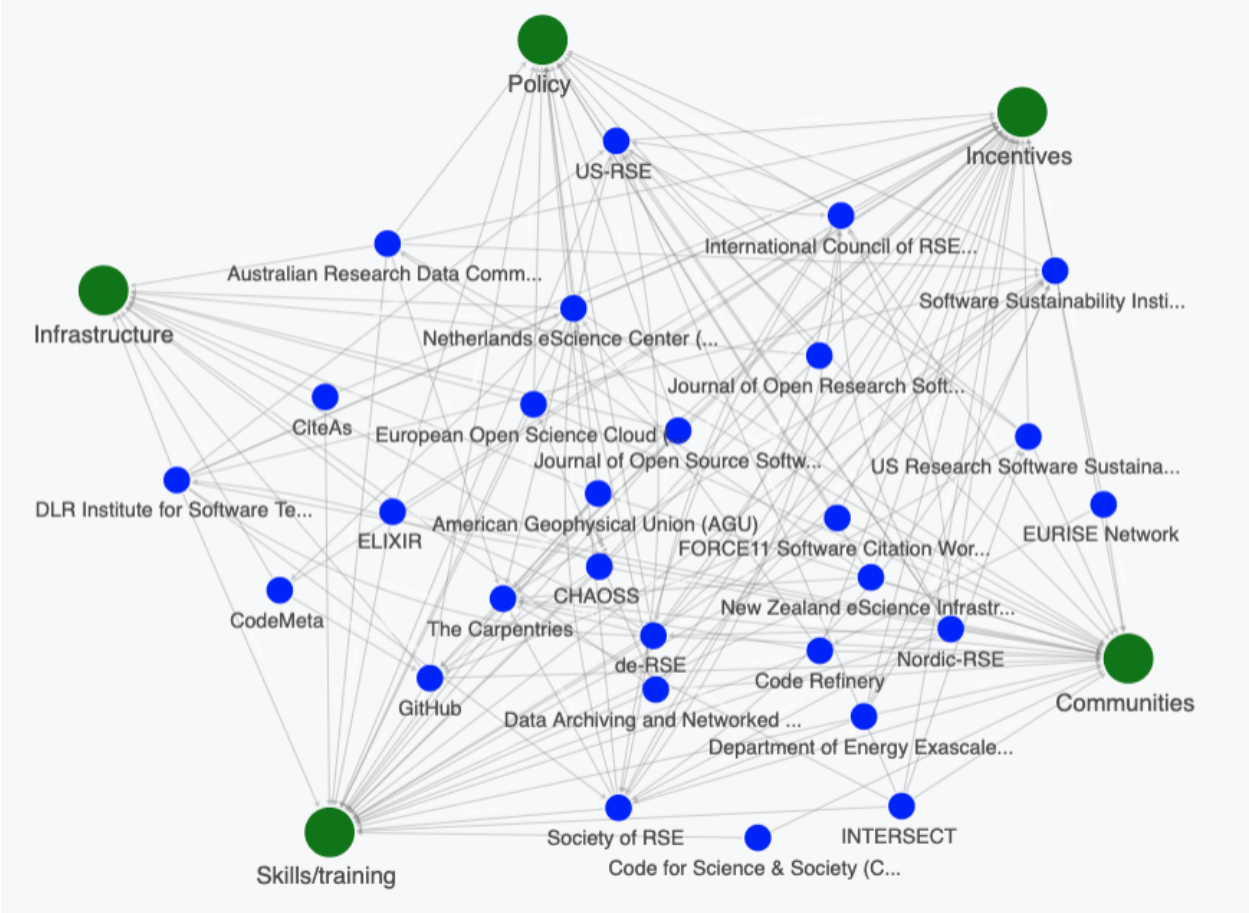

Figure 3: Initiatives linked to each theme

3.2.1 Policy

Sixteen of the profiled initiatives (57%) address policy issues related to people themes. Policy is defined here as activities that aim to influence the policy or actions of other stakeholders, not only one's own organisation. Some initiatives include this as a key strategic goal:

- Australian Research Data Commons (ARDC): The Australian national agenda for research software created by ARDC includes actions to create the policy and incentives environment to encourage code availability; that recognises research software as a first class output of research; and that supports the development and maintenance of critical research software infrastructure (Honeyman & Treloar, 2021).
- International Council of RSE Associations: Strategic goals include communication and coordination, which focuses on coordination of advocacy, developing a common argument for advocating for the implementation of RSE for institutions, policy makers, funders, etc.
- Society of RSE: Strategic goals include creating a research environment which recognises software and its contributors, the remit of which encompasses "engaging and

lobbying UK Research and Innovation (UKRI) and other UK funding councils to ensure funding guidance around the inclusion of research software and the people who develop it to promote a recognition and reward structure within research institutions", and "working with software focused partners such as the Software Sustainability Institute (SSI) to lobby for the adoption of policy that encourages recognition of software as (and as a key part of) research practice and outputs (including publication and impact)." Other strategic goals also mention policy; the goal on increasing awareness and opportunities for the role of RSE includes an aim "to work with stakeholders to lobby for the adoption of policy which encourages visibility of the people behind software to raise awareness of the important work that they do" (Society of RSE, 2020).

Examples of policy development on people-themed issues related to research software include:

- American Geophysical Union (AGU): Recent publication policy updates require data citation and highly encourage software citation (AGU, 2021).
- de-RSE: Activities include development of position papers that discuss relevant issues around research software in Germany and involve the community in their creation and review, including "An environment for sustainable research software in Germany and beyond: current state, open challenges, and call for action" (Anzt et al., 2021).
- Future of Research Communications and e-Scholarship (FORCE11) Software Citation Implementation Working Group (SCIWG): Aims to achieve adoption of the software citation principles (Smith et al., 2016) and consider how to implement them in different contexts. Their vision is that people who develop software are given the scholarly credit they deserve by having their software directly cited. Policy development occurs through task force collaborations with registries, repositories and journals. These task forces have different roles in achieving the goals to endorse the software citation principles; develop sets of guidelines for implementing the principles; help implement the principles - provide meta-guidance (on how to use guidance) and guidance on adoption to help them feel comfortable in adopting the principles; and test specific implementations of the principles. Publications include best practices for software repositories and registries and guidance for journals (Katz et al., 2021; Monteil et al., 2020).
- Netherlands eScience Center (NLeSC): Collaborations with national research bodies such as the Association of Universities in the Netherlands (VSNU) and the Dutch Research Council (NWO) include focus on increasing recognition of RSE recognition.
- Journal of Open Research Software (JORS): Incorporates policy for relevant areas such as software citation features, and submissions are required to adhere to these. It is part of the goals of JORS to ensure that the software and the papers will be citable, and that reuse will be tracked.
- Journal of Open Source Software (JOSS): Similar to JORS in incorporation of policy into submission and review requirements, to increase the citability of software.
- SSI: The policy team works with the research community to understand the issues surrounding the use of research software, and then campaigns to raise awareness and solve those issues. They have produced a range of outputs (Software Sustainability Institute, 2019).

It should be noted that the definition adopted here is of policy work that focuses on organisations beyond one's own is a fine line to draw. For example, large organisations such as the German Aerospace Center (DLR) and Department of Energy (DOE) Exascale Computing Project (ECP) can influence a large number of stakeholders internally, and showcase best practice to others. Some examples of their internal policies to support research software personnel are included in the next section.

It can also be difficult to differentiate between policy influence or advocacy, and outreach activities. For example, the US-RSE Association includes advocacy as one of its four main goals, with the aim for this including to promote RSEs' impact on research, highlighting the increasingly critical and valuable role RSEs serve. Similarly, some initiatives have activities or outputs that can be used to influence policy, but the initiative itself does not use them in this way.

3.2.2 Incentives

All but two of the profiled initiatives (93%) highlighted incentives as part of their goals, even if the initiative's main focus was on the interim steps needed to support this broader agenda. For example, CiteAs provides infrastructure that can improve the visibility of research software work; however, at a high level CiteAs aims to change the incentive environment so that the recognition for people doing this work changes. The only two initiatives that didn't have this focus, Code for Science and Society, and GitHub, are arguably more focused on code and developers in a general context rather than a research software domain.

A range of new roles are emerging for digitally skilled professionals in research, including that of RSEs, research software project managers and community managers, who may come from varied educational backgrounds. These roles need metrics to enable recognition of their achievements, and defined career paths. Programs related to incentives are defined here as those that focus on ways to improve the establishment, recognition and progression of research software personnel roles, as this requires both understanding of the importance of these roles, and need for ways to implement them within personnel structures in ways that ensure recognition and career paths are provided.

Some initiatives include incentives as a major focus in their vision or mission:

1. Society of RSE: Our mission is to establish a research environment that recognises the vital role of software in research. We work to increase software skills across everyone in research, to promote collaboration between researchers and software experts, and to support the creation of an academic career path for Research Software Engineers.
2. JOSS: Aims to provide a pathway to academic credit for researchers who develop software, and to improve the quality of open source software through peer review.

Some of the incentives projects of specific initiatives include:

3. DLR Institute for Software Technology: Have advanced internal personnel policies by developing a key performance indicator which recognises staff effort in developing open source software.
4. Department of Energy Exascale Computing Project (DOE ECP): Identify roles, careers and recognition as important to their work, as their hundreds of million dollar budget does not include development of exascale computers, but is for the personnel who are building the software that enable these computers to function. DOE ECP has RSE positions with associated merit levels, and has introduced new roles in its software teams to improve their commitment to software quality. DEE ECP is also considering how to develop a research software scientist role, reflecting the benefits of having cognitive or social scientists incorporated into teams, whose role is to integrate social science methodology to strengthen understanding of how people can improve software development.
5. EOSC: Includes in the objectives of the forthcoming EOSC Task Force on Infrastructures for Quality Research Software: “Increase recognition to software developers and maintainers of research software as a valuable research result, on a par with publications and data, in the Open Science landscape.” The Task Force will include an education, training and career working group whose remit includes exploration of best practices to recognise software development in careers and evaluations (these practices differ from discipline to discipline, and from country to country) (EOSC, 2021a). This builds on a 2020 EOSC working group which made recommendations on work needed to improve systems for providing credit to research software personnel (European Commission, 2020).
6. Society of RSE: Held an RSE Careers Pathway event in 2021 which attracted over 200 registrations, aimed to explore and expose some of the boundaries to career progression for RSEs within the UK. The event provided an opportunity to explore possible alignment with the Research Concordat and the Technician Commitment, which set out principles and obligations to be followed by research institutions in supporting talent development (Richmond et al., 2021).
7. US-RSE: Strategic goals including provision of useful resources to multiple demographics: “For current and future RSEs we strive to provide technical and career development resources to support their professional development” (US-RSE, 2021).

3.2.3 Communities

24 of the profiled initiatives (86%) include community in their remit, which is consistent with the termscape analysis in Section 4.1, identifying community as the most commonly mentioned term in missions and/or visions. The exceptions include CiteAS and CodeMeta, which are time-limited projects to develop specific services or infrastructure. Communities are defined as groupings that enable people with common interests or characteristics to come together. The focus of these communities varies, and includes sharing of best practice in particular areas, such as software citation or development of RSE career paths; and provision of mutual support

for people identifying as part of emerging communities, such as RSEs, or particular types of trainers, learners, and leaders.

Some initiatives include development of community as a key strategic goal, including:

- The Carpentries: Our vision is to be the leading inclusive community teaching data and coding skills. The Carpentries Strategic Plan for 2020-2025 includes as its first goal to build regional and local capacity to empower sustainable communities (The Carpentries, 2020).
- de-RSE: It is our vision to establish research software as a first-class citizen in academia, establish and support careers for RSEs across institutions, support the education relevant for Research Software Engineering to increase the quality of research software, and grow and facilitate the Research Software Engineering community in Germany.

Some of the community projects of specific initiatives include:

- ARDC: Actions contained in the national research software strategy to develop community include actions to build and sustain a national community of practitioners interested in informatics infrastructure and in the behaviours that inform it; and build and sustain a national community of professional RSEs.
- Council of International RSE Associations: Provides a formal open forum for established national and multinational RSE associations to talk and coordinate regularly, and thus sustain international collaboration.
- European Research Infrastructure Software Engineers' Network (EURISE Network): Formed by the three Social Sciences and Humanities European Research Infrastructure Consortiums (ERICs): Consortium of European Social Science Data Archives (CESSDA), Common Language Resources and Technology Infrastructure (CLARIN) and Digital Research Infrastructure for the Arts and Humanities (DARIAH), to create an umbrella where research infrastructures meet research software engineers, the EURISE Network promotes knowledge exchange between experts from different areas of infrastructure development and operation.
- INnovative Training Enabled by a Research Software Engineering Community of Trainers (INTERSECT): Goals include creation of a community of RSEs through the development of a training curriculum to educate the next generation of RSEs. The project will include activities to grow the community in meaningful ways, such as workshops designed to bring trainers together to build the curriculum.
- NeSI: Supports community development including events such as the annual New Zealand Research Software Engineering Conference, and eResearch New Zealand conference.
- SSI: A range of community building activities are undertaken, including the SSI Fellowship program. Evaluation has shown that this plays a wide-ranging role in supporting communities of best practice and skills transfer (Sufi & Jay, 2018).

- US Research Software Sustainability Institute (URSSI): Their four main areas include community and outreach, with URSSI aiming to connect researchers or developers in the role of an RSE to peers and provide access to beneficial material, resources, and contacts. This includes aims on promoting new career paths for those who develop and maintain research software; and developing and advocating for research software usage and impact metrics to be a factor in the hiring and promotion of software developers and maintainers
- US- RSE: Convenes events to increase community engagement, such as a two-day community building event scheduled for April 2022 (US-RSE, 2021).

Many of the profiled initiatives specifically incorporate a diversity, equity and inclusion (DEI) focus in their community (and/or including their own staff), some of which are included below:

- The Carpentries: The Carpentries Strategic Plan's second goal is to intentionally incorporate equity, inclusion, and accessibility to support a diverse community (The Carpentries, 2020). Their projects include convening Carpentries workshops to increase the number of certified Instructors who are Latinx, Black, and Indigenous, and development of culturally-relevant training for data literacy and computational skills. The Carpentries have also partnered with Understood, an organisation committed to shaping the world for difference, which includes creating workplaces that are inclusive and supportive for people with disabilities. As part of that partnership, Understood will be providing disability inclusion training to staff at The Carpentries.
- Code for Science & Society (CS&S): Projects include a conference fund to support events that promote inclusion and broaden participation in open data science that drives scholarship.
- Community Health Analytics Open Source Software (CHAOSS): Encourages projects and events to obtain CHAOSS diversity and inclusion badges for reasons of leadership, self-reflection, and self-improvement on issues critical to building the Internet as a social good. The project aims to increase understanding of the open source project and event practices that encourage greater diversity and wider inclusion of people from different backgrounds (CHAOSS, 2020).
- GitHub: Provide a range of initiatives across platform, people, philanthropy and policy to fulfill their high level goal: As home to the largest developer community in the world, we believe deeply in the potential and the power of a diverse open source community. We feel it is our responsibility to promote diversity and inclusion, and integrate these values into every aspect of what we do." (GitHub, 2021b)
- US-RSE: Recently updated its mission to include a fourth goal on diversity, equity and inclusion: "We will actively promote, encourage, and improve diversity throughout the broader US Research Software Engineer Community consistent with our full diversity, equity, and inclusion mission statement. We will ensure we provide an inclusive environment with equitable treatment for all and we will prioritize a program of diversity, equity, and inclusion activities for our organization, led by a dedicated team of active community members." (US-RSE, 2021)

- URSSI: Utilised a survey to identify some areas where URSSI could play a key role in advancing the sustainability of research software in the United States. Findings included that most projects lack a formal diversity plan, and that URSSI could help by providing template diversity plans and support for developing appropriate plans for individual projects.

3.2.4 Skills and training

25 of the profiled initiatives (89%) incorporate a focus on skills and training. This is defined here as initiatives that support skills and training efforts beyond those for their own staff, and include designing curriculum, training trainers, identifying skills frameworks, and convening training and mentoring events. The skills focused on can include “soft” skills such as teamwork, communication and leadership. Some initiatives focus on this in their mission:

- The Carpentries: “builds global capacity in essential data and computational skills for conducting efficient, open, and reproducible research. We train and foster an active, inclusive, diverse community of learners and instructors that promotes and models the importance of software and data in research. We collaboratively develop openly-available lessons and deliver these lessons using evidence-based teaching practices. We focus on people conducting and supporting research.” (The Carpentries, 2021a)

Some examples of initiatives with skills and training strategic goals and/or projects include:

- CS&S: The Digital Incubator Infrastrastructure program supports open source project leaders implementing best practices in sustainability, governance, and community health.
- Code Refinery: Are working with students, researchers, RSEs and national e-infrastructure partners to advance FAIRness of software management and development practices so that research groups can collaboratively develop, review, discuss, test, share and reuse their codes. Aims to become a Lesson Program of The Carpentries, which are collections of lessons comprising one or more Carpentries workshops, serving distinct goals and audiences.
- Data Archiving and Networked Services (DANS): Is integrating software elements into their training portfolio and some training provided by partners.
- DLR Institute for Software Technology: Provides a range of training opportunities to the Helmholtz Association of German Research Centres, sometimes based on curriculum from The Carpentries (HIFIS, 2021). Training focuses on the software development skills needed by staff who undertake research (rather than professional research software developers), and include two-day training workshops that have been convened for more than eight years. There have been a number of publications discussing how their software engineering guidelines and community have been developed (DLR, 2019)
- ELIXIR: ELIXIR activities are divided into five platforms, which include both a Tools Platform to help communities utilise software best practices, and a Training Platform that

is focused on the coordination, management and delivery of training across all ELIXIR nodes. The ELIXIR Training Platform aims to strengthen national training programmes, grow bioinformatics training capacity and competence across Europe, and empower researchers to use ELIXIR's services and tools. There is also an extensive train the trainer programme that aims to build a network of trainers to allow them to benefit from reciprocal support and discussion (ELIXIR, 2021).

- EOSC: The Strategic Research and Innovation Agenda of EOSC identifies as the first of its three general objectives: “Ensure that open science practices and skills are rewarded and taught, becoming the ‘new normal’” (EOSC, 2021b). There are a number of projects that address different elements of this goal, including EOSC Synergy, a Horizon 2020 project that is developing a free digital learning environment to support skills for EOSC.
- INTERSECT: Development of an open source modular training framework for RSEs is one of the goals of this three-year project. The curriculum will target postgraduate researchers who have no formal training in RSE, to fill gaps in their knowledge. It will provide a level of training beyond that of The Carpentries. Activities have included collection of RSE training material, workshops, and resources (INTERSECT, 2021).
- JORS has two major types of submissions: Software Metapapers describing research software with high reuse potential; and longer research papers that cover different aspects of creating, maintaining and evaluating open source research software. The aim of the latter is to promote the dissemination of best practice and experience related to the development and maintenance of reusable, sustainable research software.
- New Zealand eScience Infrastructure (NeSI): Objectives include growing capability in the sector by lifting research capability and research artefact management. NeSI participates in community leadership of a range of communities including RSE-Australian/New Zealand (RSE-AUNZ), The Carpentries, and Women in High Performance Computing (WHPC). NeSI supports a regional Carpentries coordinator, whose focus includes national leadership on train the trainer programs.
- Nordic RSE: Aims to provide a range of activities including advanced training for RSEs, ResearchSoftwareHour, a mentoring program to support RSE career development, and assistance with code checks.
- SSI: Skills and training programs include free online Research Software Camps held twice a year over the course of two weeks, and Carpentry programmes. SSI coordinates activities of The Carpentries in the UK, including organising and running workshops, and nurturing the UK community by bringing in and training new instructors and members.
- US-RSE Education and Training Working Group: Has begun describing the types of people who work as RSEs and what their training needs might be, to enable definition of the list of skills often needed by RSEs.

DOE ECP and DLR both also have a focus on applying the scientific method to understanding how people develop and use software to do research. This approach, sometimes called the science of science, or metascience, utilises the approaches of social and cognitive science to improve individual and team behaviours to improve software outcomes.

3.2.5 Infrastructure

There are thirteen initiatives (46%) that include an infrastructure focus related to people-themed issues. Infrastructure is defined here as the components used to support the role of software for and in research practices, including the schemas and standards that support this. Some of the profiled initiatives have infrastructure related to people-themes as a key strategic goal, including:

- CiteAs: Is a tool for obtaining the correct citation for diverse research products including, software, datasets, preprints, and traditional articles. By making it easier to cite software and other "alternative" scholarly products, CiteAs aims to help the creators of such products get full credit for their work.
- CodeMeta: Strives to promote the citation and reuse of software authored for scientific research by developing a mechanism to assist the transfer of software and software metadata between the entities that author, archive, index and distribute and use the software. Their approach is not to create a new metadata standard or schema, but instead to define a crosswalk between existing software metadata schemas, and to provide a uniform method to package and transfer this metadata between entities.

Example of projects on people-themed issues related to infrastructure include:

- AGU: Incorporate best practices around data and software sharing in AGU journals, including use of appropriate, trusted repositories and how to incorporate citations in papers (AGU, n.d.). Data and software citations allow for automated linking to the paper, author, and funder, and support credit for the creator.
- CHAOSS: Creates tools such as Augur, a software suite for collecting and measuring structured data about free and open-source software communities. CHAOSS is a Linux Foundation project focused on creating analytics and metrics to help define community health.
- DANS: Aims to ensure their infrastructure enables software citation and the measurement of usage and impact metrics for software.
- DLR Institute for Software Technology: Software engineering guidelines are supported by analytics that enable examination of how coding behaviour changes. For example, the Back-bone Catalogue of Relational Debris Information (BACARDI) project applied methods to structure, communicate, and utilise the diverse skills, knowledge, and experience in the team concisely and precisely. After one year of practical utilisation, these were analysed based on repository data to assess and prove the effects of the introduced process on the development of a software (von Kurnatowski et al., 2020).
- ELIXIR: Coordinates and develops life science resources across Europe so that researchers can more easily find, analyse and share data, exchange expertise, and implement best practices. TeSS is ELIXIR's training platform, providing a one-stop shop for trainers and trainees to discover online information and content, including training materials, events and interactive tutorials. Another example is bio.tools, which strives to provide a registry of software and databases, facilitating researchers from across the

spectrum of biological and biomedical science to find, understand, utilise and cite the resources they need in their day-to-day work.

- GitHub: Includes built-in support for software citation (GitHub, 2021a). This feature enables academics and researchers to let people know how to correctly cite their work, especially in academic publications/materials, and is integrated with archives including Zenodo and Figshare. It is built on a standardised citation format created by the research software engineering community (Druskat et al., 2017).

The definition of infrastructure utilised here makes classification of some initiatives challenging. For example, Code Refinery provides a code repository for Nordic research software, for collaborating on code and scripts. Similarly, DOE ECP created The Extreme-scale Scientific Software Stack, a community effort to provide open source software packages for developing, deploying and running scientific applications on high-performance computing platforms. It supports people-related themes as it improves the ability of teams to work together, and includes community policies in areas like team behaviour. Should these be classified as people-related infrastructure?

3.3 Collaborators

The questionnaire asked each initiative to identify its collaborators. The 28 initiatives identified another 98 initiatives that they collaborated with, as shown in Figure 4. However, it should be noted that this is only an indication of some of the collaborators, as many initiatives did not list all of their members (EOSC has 100+ members and ELIXIR has 39 national nodes that are not listed here), or identified a broad grouping such as “industry”. There was also variance in consideration of what constituted a collaboration, with some only listing partners with whom formal agreements are in place, whilst others included informal collaborations.

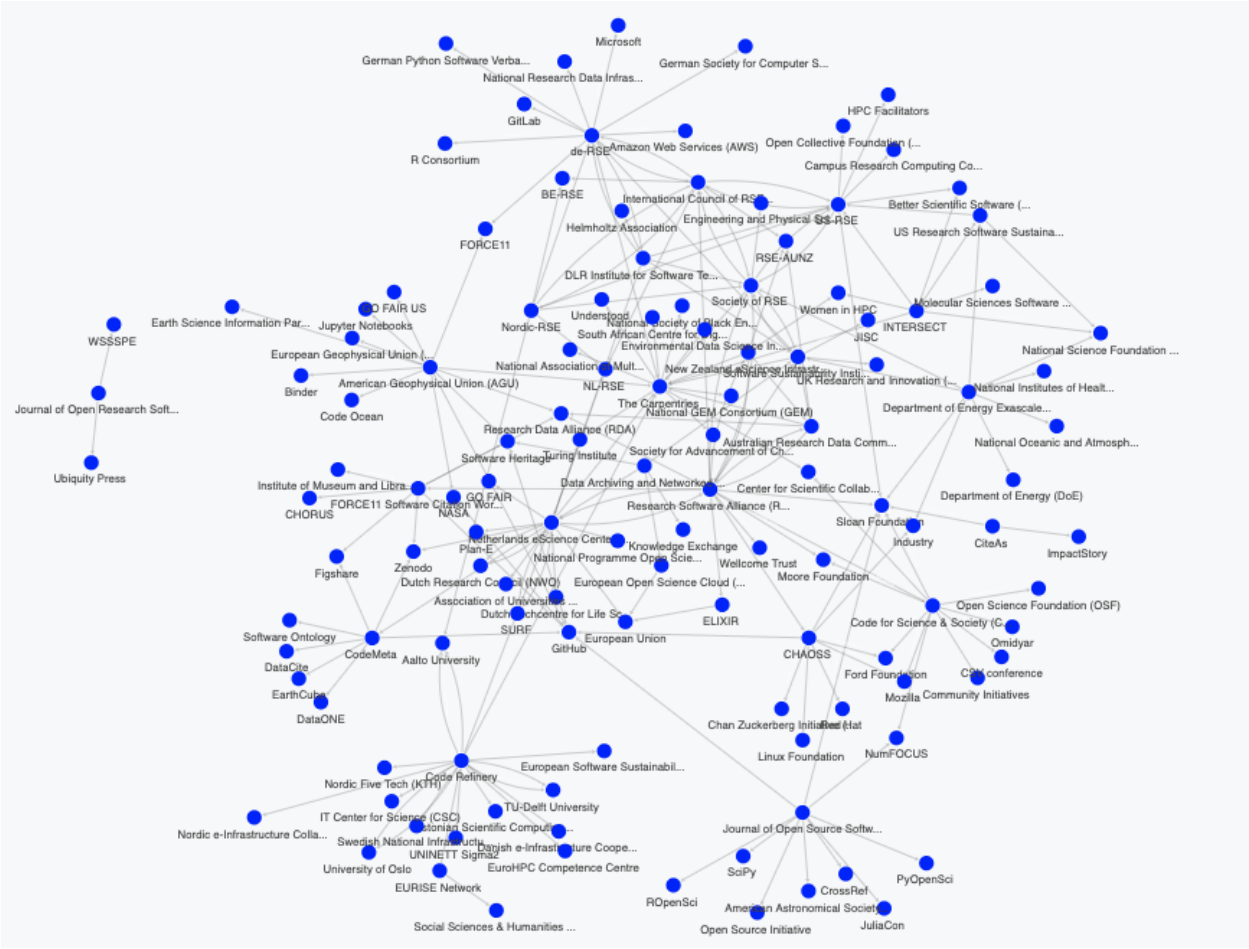

Figure 4: Collaborations between initiatives

This illustrates the breadth of initiatives of potential relevance. Initiatives supporting the research software community include those that focus on open science, reproducibility, roles and careers for people who are less visible in research, publishing and review, and other types of scholarly products and digital objects.

Figure 5 narrows down the collaborators to only those 28 initiatives that were profiled, to show the interlinkages between them.

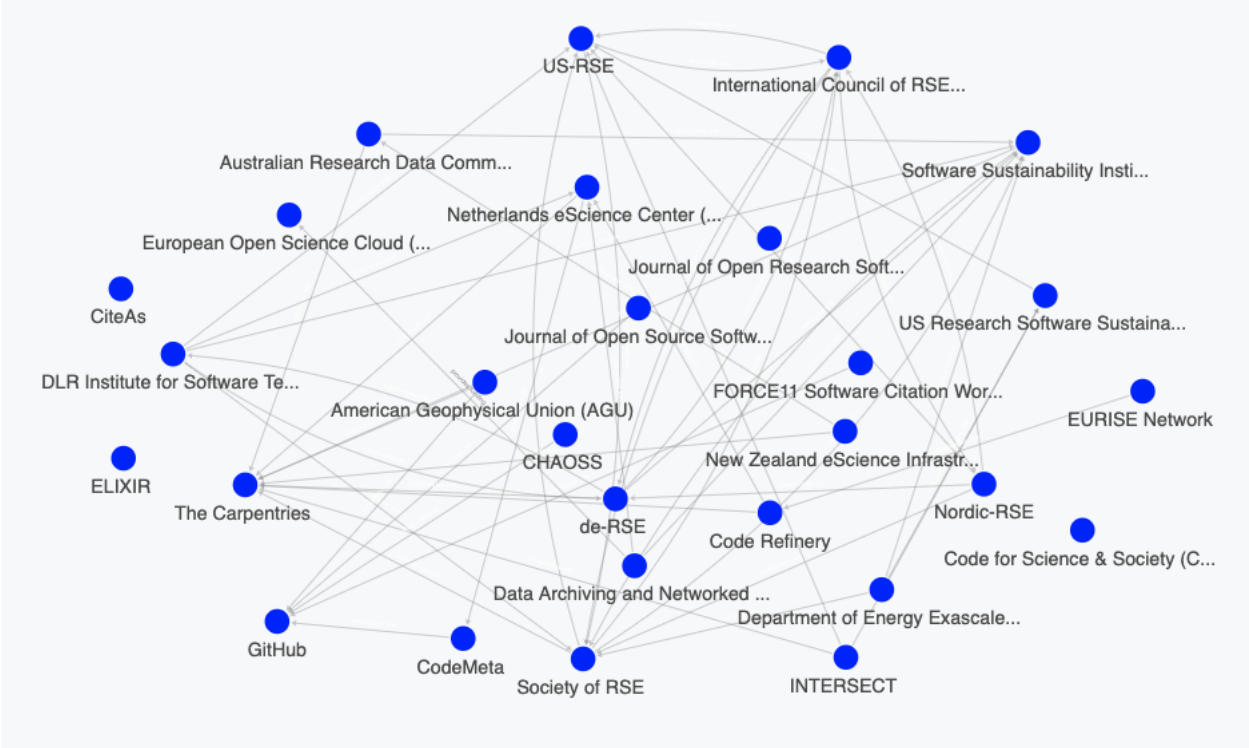

Figure 5: Collaboration between profiled initiatives

This shows that some initiatives are more interlinked, particularly SSI, the Society of RSE, and The Carpentries, as shown below in Figures 6-8. Figures 6-8 also detail if an initiative both identified itself as collaborating with one of these three initiatives, and was identified by that initiative as well (by showing two lines linking the two initiatives), although if only one of these is the case then the diagram does not clarify which.

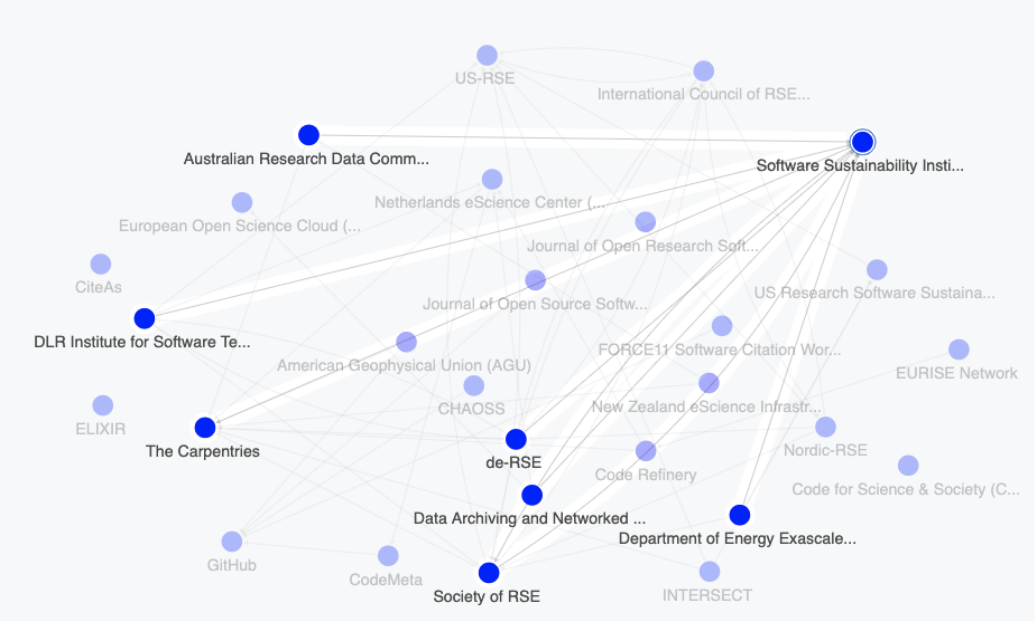

Figure 6: SSI's collaborations with other initiatives

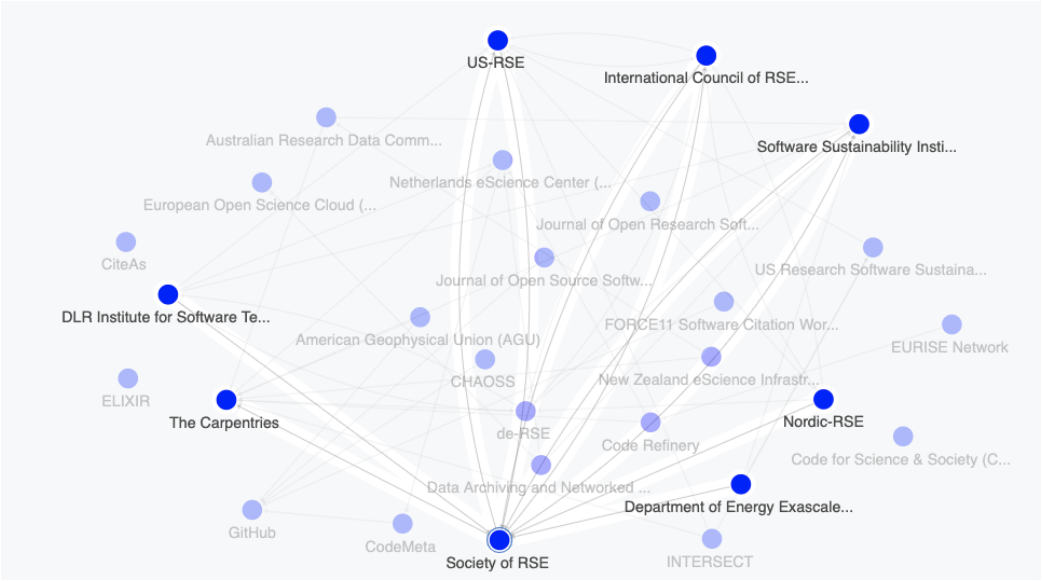

Figure 7: Society of RSE's collaborations with other initiatives

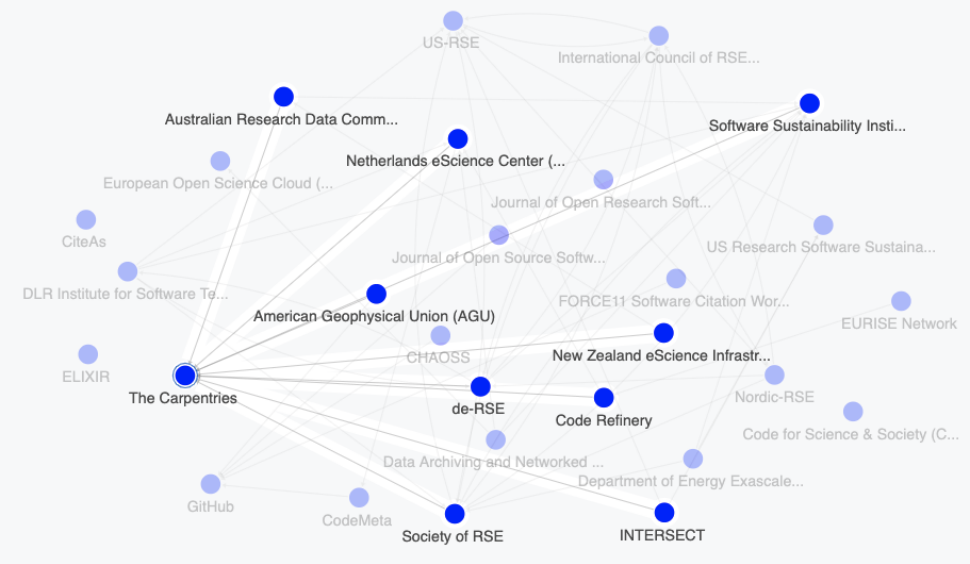

Figure 8: The Carpentries' collaborations with other initiatives

It should be noted that the four RSE organisations that were profiled (de-RSE, US-RSE, Nordic-RSE and Society for RSE) are all members of the International Council of RSE, as this explains the strong links between these.

3.4 Resourcing

Each initiative where asked where the resources came for their people-themed projects, although most answered in terms of where the initiative's overall funding came from. 24 different funders were identified, with almost all funding only one initiative. The exception was the Alfred P. Sloan Foundation which supports at least six initiatives. However, this analysis is incomplete as it does not show where membership fees were included, thus omitting a number of organisations that do support the 28 initiatives profiled. Many initiatives also receive significant in-kind support from the employers of key staff and/or community members, which was also not included in this analysis. However, this does show that there are a variety of funders for this type of work.

3.5 Demonstrating benefits

The initiatives were asked what information/evidence (data, policies) they use to support investment in this area, and how they track impact or outcomes. There were some commonalities in their answers:

- Membership-based or community organisations often use growth in member numbers as a key metric.

- A common response was around skills and training metrics was a focus on demographics of participants, post-workshop impact, and qualitative case studies, such as that provided by the Carpentries (The Carpentries, 2021b).
- Initiatives with an infrastructure focus, such as CHAOSS, utilise metrics such as downloads, contributors, clones, forks, people contributing to projects, etc.

Other responses from individual initiatives included:

- DLR Institute for Software Technology: Created software analytics to examine how use of their software engineering guidelines was impacting on behaviour. The analytics extract metadata info from GitLab to analyse how coding behaviour changes, and also surveys developers. (von Kurnatowski et al., 2020).
- FORCE11 SCIWG: Tracks implementation of their recommendations by tallying which stakeholders (such as journals) have software citations policies, and how the number of software citations is increasing in repositories such as DataCite, Freya and Zenodo.
- GitHub: Have analytics on organisational users (GitHub, 2021c).
- NeSI: Utilise the international RSE survey (which is presumably also a source of information for national RSE groups).
- NLeSC: Track reuse of software through their research software directory. This includes a manually-curated metrics section with information about software citation in articles, blogs, conference presentations, etc. (NLeSC, 2021).

3.6 Priority areas

The initiatives were asked about priority areas for the sector, either where they'd like to see more focus occurring across the research software sector, or where they'd like their initiative to be able to do more. 40% of the initiatives stated that career paths and recognition for research software personnel was important, and 25% supported valuing research software as a first class research object. Other priorities that were identified are incorporated in the next section.

4. Analysis

This section analyses the data from Section 4 and concludes with suggestions on future work. While the discussion continues to be framed in the five areas of people, incentives, communities, skills and training and infrastructure, these areas overlap.

4.1 Policy

Policy development is a key part of cultural change. The OECD report on digital capacity building and skills for data-intensive science identifies enablers for digital workforce capacity development as one of its five action areas. The report provides recommendations to a range of stakeholders on how to integrate digital workforce capability development into broader science policy frameworks and action (OECD, 2020). Figure 9 shows the profiled initiatives that have a

policy focus related to the people theme who collaborate with which other (although not necessarily on policy goals or projects)

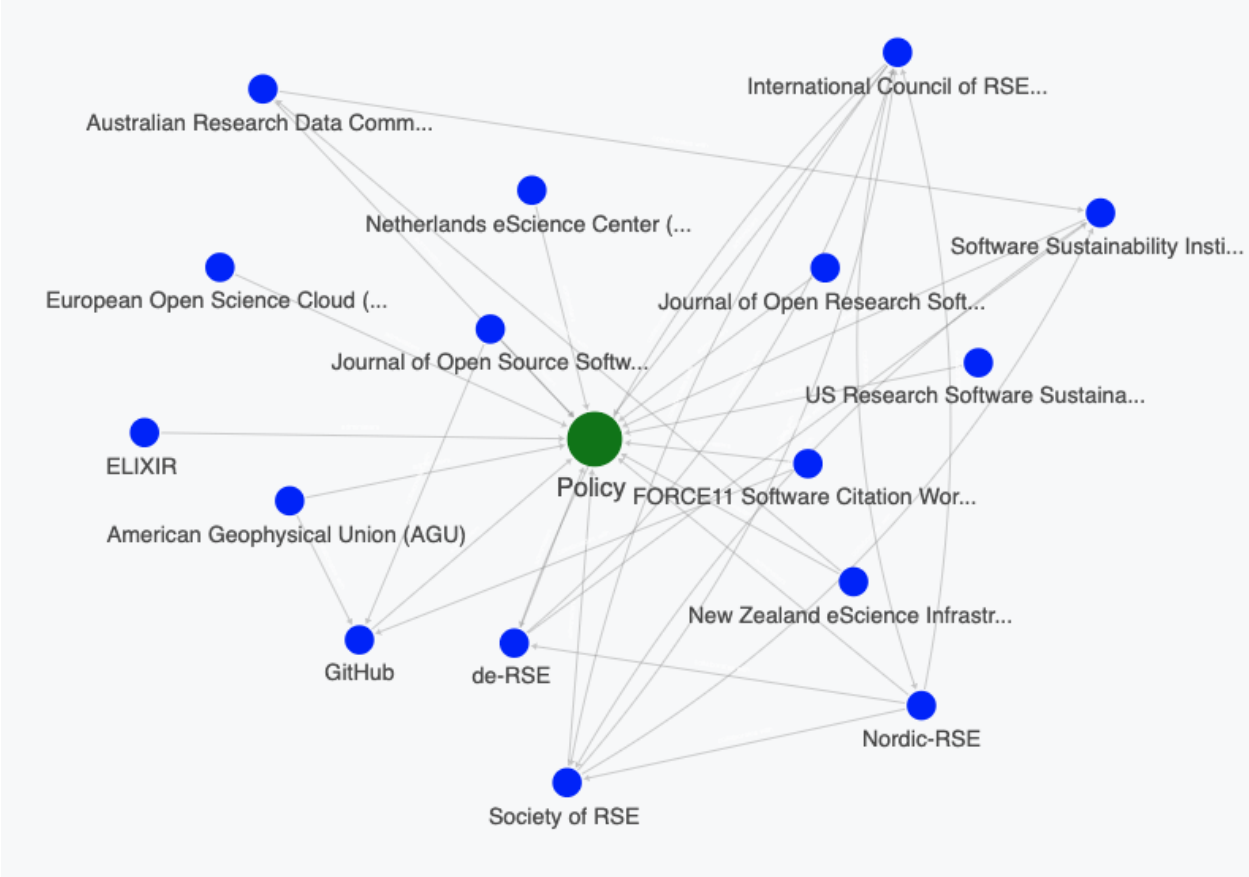

Figure 9: Initiatives with a policy focus

It can be seen that there is some collaboration between the initiatives, although this analysis does not differentiate whether the collaboration is particularly on policy development, or other goals and projects.

It is likely that more integration would assist in the achievement of common goals. URSSI's identification of common challenges and dilemmas in the USA found a need for coordination and targeted leadership around policy for sustainable research software:

> Participants in URSSI workshops described a need for coordinating communities around emerging national, institutional, and even disciplinary specific policies that have a downstream impact on sustainability, such as software citation principles, tenure and promotion guidelines that recognise research software contributions, sustained funding or financial support for research software maintenance, and software management plans that explicitly document expectations around software development and archiving.

Currently, community members take on many of these policy activities as additional professional service, or as volunteer work. This secondary focus on any one policy issue, in turn, leads to slow progress, high turnover of volunteers, and does not allow any one person or institution to develop the deep expertise needed for effective sustained analysis or advocacy (Ram et al., 2021).

URSSI aims to fill this gap for the USA, to facilitate data-driven policy research that could lead to stronger advocacy positions, and benefit funding agencies, universities, and research institutions that seek to adopt new policy aimed at improving research software sustainability.

ReSA aims to enable more collaboration in this area at the international level, and has undertaken activities including:

- Influencing international OECD and UNESCO community consultations and expert groups to include software in key policy documents (OECD, 2021; UNESCO, 2021).
- Coordination of community development of research software sharing guidelines for policy makers, funders, publishers and researchers for inclusion in the Research Data Alliance COVID-19 Guidelines and Recommendations (Research Data Alliance, 2020).

The profiled initiatives also identified a range of priority areas and possible solutions, particularly related to funder policy:

- Support equal emphasis on each of data, software and publications as critical components of research.
- Develop funding for team science to improve practices. There has been an enormous investment in software infrastructure, but very little on people elements such as coding practices, team behaviours or decision-making processes.
- Encourage emerging digital roles such as RSEs, data stewards and data scientists, to work together to gain recognition.
- Require inclusion of research software professionals on grants with a software development component, to assist recognition of key personnel and provide stable funding.
- Increase investment in community management roles for research software communities. A lot of the capacity building work in areas such as training and promoting DEI, must be financially supported by funding bodies and individual institutions. It is not sustainable to ask RSEs or individuals in other roles to do this work “on the side.”
- Increase understanding by funders of how to assess research better and provide better opportunities to recognise and reward the diversity of roles within research.
- Analyse funder policies on the referencing of outputs, including software.
- Enable RSE initiatives to move beyond survival or subsistence mindsets around funding. This can lead to conservatism in risk taking in some areas, and moving beyond this could increase the ability to take risks and regard failure in a less terminal fashion. Increase sustainability of research software beyond the product funding cycle, as this has consequences for the developers and maintainers.

- Enable easier access to resourcing that has less restrictions and bureaucracy, to make it more equitable.

4.2 Incentives

Effort is needed to change academic evaluation and reward systems to attract and retain diverse personnel. Figure 10 shows the profiled initiatives with a focus on incentives who collaborate with which other (although not necessarily on incentives-related goals or projects).

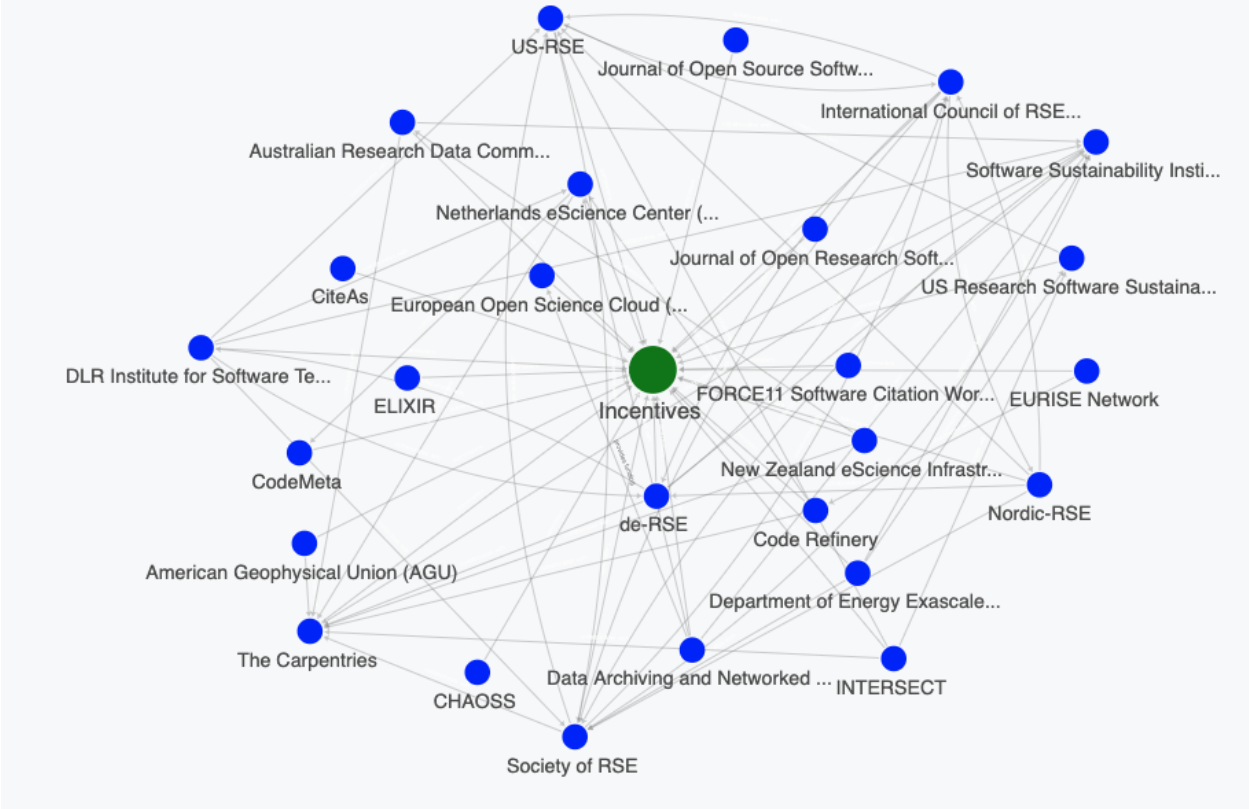

Figure 10: Initiatives with an incentives focus

This is often identified as a key issue. The Society of RSE held an RSE careers pathway event which attracted over 200 registrations (Richmond et al., 2021). Research undertaken to identify key issues in RSE across people, policy and infrastructure areas resulted in the largest number of questions being related to the people theme, and that all of the people-themed questions are centered around RSE career paths, with training, recruitment and retention of talent being identified as an issue. RSE careers were also central to a majority of the policy questions (and also play a role in infrastructure) (Lamprecht et al., 2021).

The profiled initiatives also identified a range of priority areas and possible solutions related to incentives that are particularly relevant to employers:

- Increase understanding of career recognition and metrics issues across a wider range of countries to consider cultural differences.
- Consider career recognition and metrics issues from alternative viewpoints (not just technical). For example, open source software functions as a community.
- Encourage emerging digital roles such as RSEs, data stewards, and data scientists to work together to gain recognition.
- Increase understanding across the entirety of the research community of the RSE role and its collaborative nature. Change perceptions of what research is, to embrace team science and the role of RSEs.
- Lobby institutions to create RSE roles and career paths.
- Provide better pay scales for research software developers, to compete with industry.
- Run competitions that highlight contributions to science through software. If professional organisations highlighted these contributions, alongside existing elements such as the best research paper, then this could change incentives by highlighting important behaviour.

To identify next steps, it could be useful to frame these issues in terms of the next steps that research software personnel could undertake to continue to identify new professions (based on work such as Abbott, 1988), or using the 12 thematic challenges identified as relevant in building the research innovation workforce: diversity and inclusivity; fostering the development and support of the workforce ecosystem and talent pipeline; establishing viable career paths and normative role descriptions in the workforce; enhancing internal and external communications and education for stakeholders; compensation; workforce sustainability; the establishment of an identity of the field as a discipline; the position of research computing within institutional organizations; and the need for continuing training and education for professionals (Arafune et al., 2020).

4.3 Communities

Communities are defined here as groupings that enable people with common interests or characteristics to come together to achieve shared goals. Professional communities enable diverse groups of people to come together to achieve change at various levels through knowledge transfer, networking, etc. Figure 11 shows the profiled initiatives with a focus on communities who collaborate with each other (although not necessarily on communities-related goals or projects).

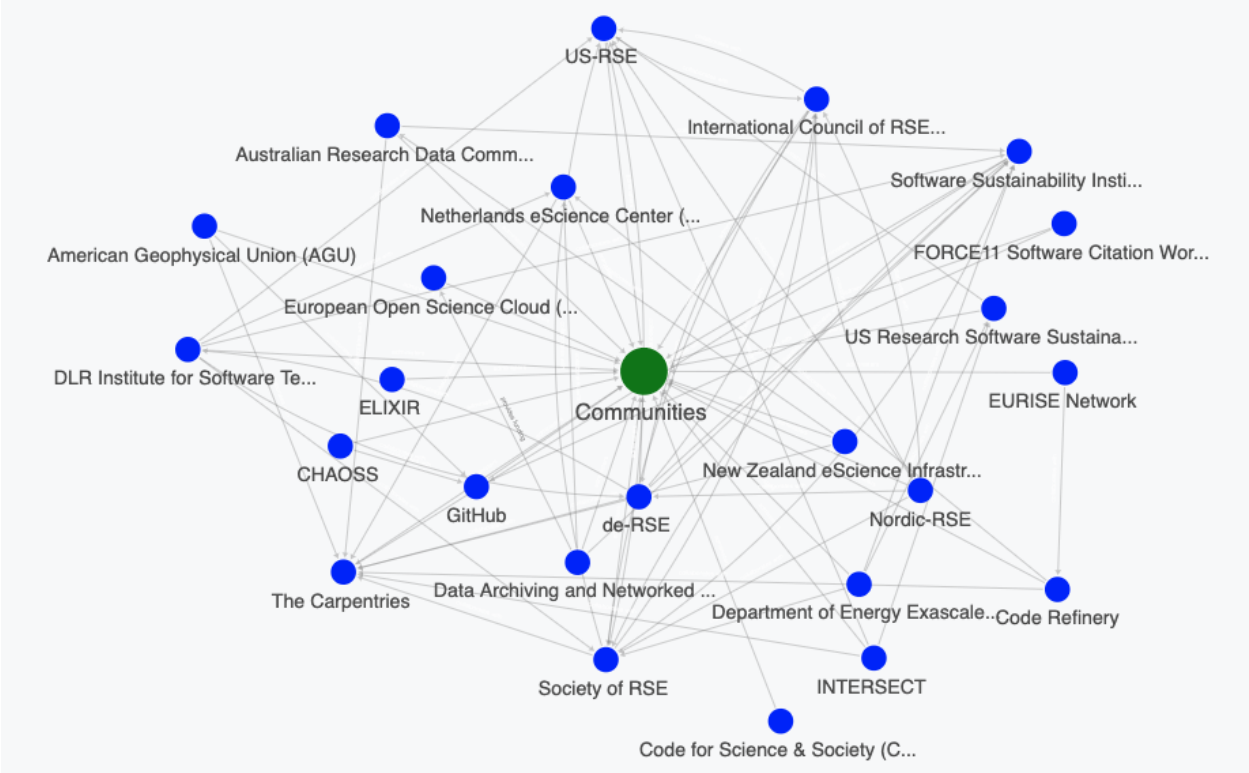

Figure 11: Initiatives with a communities focus

The finding that almost all of the initiatives include a community focus is in line with research documenting the growing range of grassroots organisations and projects that aim to improve software quality, productivity, and sustainability (Katz et al., 2019). Katz et al. propose that these endeavors ensure the integrity of research results and enable more effective collaboration. The growing emphasis on communities is also demonstrated by the rise of new roles such as the community manager, and organisations such as the Center for Scientific Collaboration and Community Engagement (CSCCE), a research and training center to support and study the emerging field of scientific community engagement.

In recognition of the continuing emphasis on DEI, and proliferation of initiatives in this area, ReSA has been conditionally approved to co-convene a Lorentz workshop in April 2022 (both online and in-person at the Lorentz Centre in the Netherlands). This workshop would consider how research software engineering could be reframed to place DEI as a central organising principle in research software. This workshop will bring together a wide range of stakeholders who have not previously collaborated, to analyse best practice and available data to identify both research opportunities and solutions to their own local challenges, to create an informal network of champions in this area.

The research software community is well situated to evolve to frame DEI at its centre due to its dependency on community involvement for innovation and sustainability. Expanding the pool of

research software contributors is a concrete, desired outcome that improved DEI could contribute to, with benefits including:

1. Increasing innovation: research has found that diverse teams can improve scientific outputs (Campbell et al., 2013; Liang et al., 2007). Research software work usually occurs in teams, and this work will continue to require the ability to maintain critical relationships with diverse stakeholders in the research community. However, Chue Hong et al. also identify research that demonstrates there can also be drawbacks, and maximisation of benefits depends on effective implementation (2021).
2. Increasing sustainability: Research software is often open source software, which is characterised by development by a network of people working together. However, the culture of open source software faces similar challenges in improving DEI (Benjamin, 2019; Dunbar-Hester, 2020; Vasilescu et al., 2015).

The profiled initiatives also identified a range of priority areas and possible solutions related to incentives, that are particularly relevant to employers:

- Enable metrics on community health to be easily accessible and meaningful.
- Increase involvement with open science communities.
- Increase engagement with the broader software engineering community (not just RSE).
- Establish cross-disciplinary RSE groups (within institutions)
- Enlarge RSE community to include non-typical collaborators such as humanities and social sciences.
- Increase diversity in research software engineering, which could also identify innovative ways to solve challengers around career paths and metrics.
- Recognise that DEI and the perennial labour shortage in open source software are a solution and problem that could be brought together.
- Enhance knowledge and use of political understandings, such as, awareness of where communities are ideologically. A healthy diversity of ideological positions is needed.
- Support initiatives like URSSI, SSI and ReSA. Expanding their services to enable broader engagement with more people would be a net benefit.

4.4 Skills and training

Ongoing work is needed to identify the key competencies and skills for research software personnel, alongside mechanisms to provide appropriate training. Figure 12 shows the profiled initiatives with a focus on skills and training who collaborate with which other (although not necessarily on skills and training goals or projects).

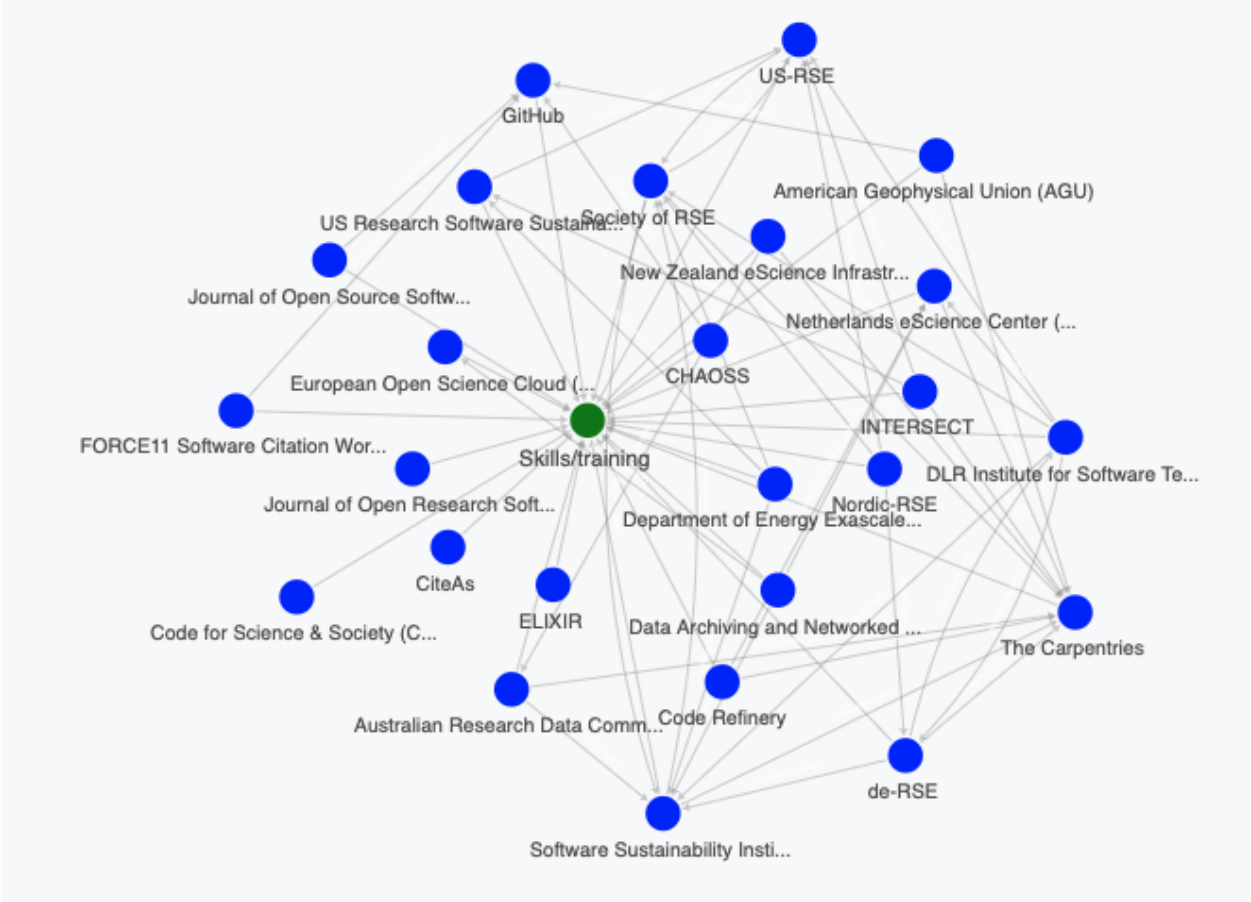

Figure 12: Initiatives with a skills and training focus

With the majority of initiatives providing some kind of effort in this area, it would be useful to nuance analysis of this landscape further. For example, whilst INTERSECT and Code Refinery both aim to create training material that is more advanced than that of The Carpentries, what other sections of a training pipeline might still need to be addressed, to ensure that all levels of training are available?

The profiled initiatives also identified a range of priority areas and possible solutions related to skills and training:

- Increase understanding of the ethical aspects of software creation, such as the potential for biases in algorithms incl ethics.
- Encourage good software creation practices by developers of virtual research environments, science gateways and eResearch platforms.
- Enable better integration of R markdown files, Jupyter notebooks, etc., into research to improve sharing and reproducibility.
- Increase understanding of FAIR (Findable, Accessible, Interoperable, Reusable) software.

- Teach people to contribute to each other's projects, and to also push fixes upstream to software with dependencies. This kind of changes would also assist in developing values on ecosystem stewardship, where developers are conscious of other software that is dependent on their work.
- Increase findability of research software with appropriate metadata, and education on how to enable software reuse.
- Create degree programs in RSE, as is already happening for data science.
- Provide more online teaching to reach participants who do not have a local community offering in-person events.

4.5 Infrastructure

Infrastructure is needed to ensure that people-themed goals can be achieved, that there are methods to implement policies and measure adherence; record the data needed to advance career paths and recognition; support communities; and to deliver and measure impact of skills and training. Figure 13 shows the collaborations between the initiatives with an infrastructure focus (although not necessarily on infrastructure goals or projects).

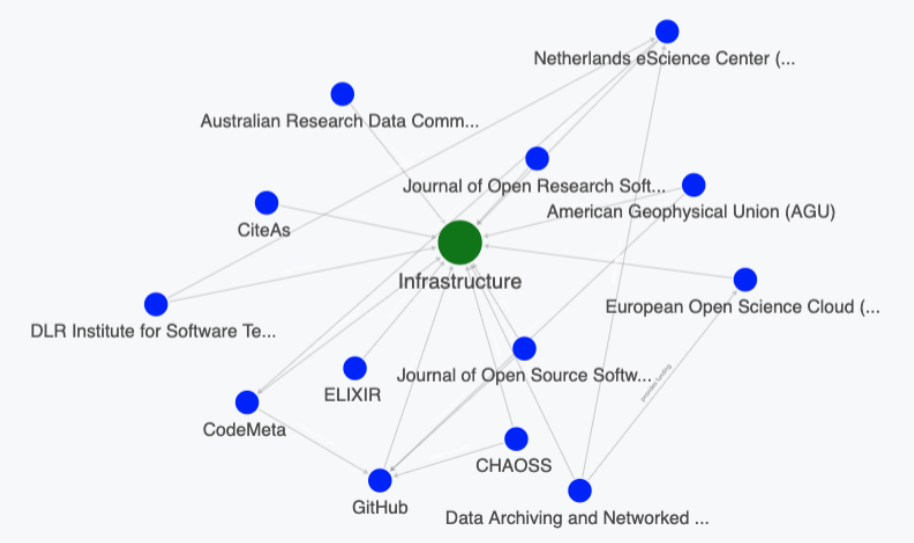

Figure 13: Initiatives with an infrastructure focus

The profiled initiatives also identified priority areas related to infrastructure:

- Enable deeper integration with the scholarly ecosystem, including better identity management.
- Provide software dependency graphs by relevant infrastructure.

5. Overall analysis

The study aimed to assist in answering a range of questions, as follow:

- What is the range of strategic goals and main projects, and what are these common and/or unique themes?
- How much collaboration and/or niche focus is occurring on strategic goals?
- Do projects for each of the five people-themes interlink?
- Where do resources come from for this work?
- What are the priority areas across initiatives?
- Consider if information collected enables answers to more specific questions, e.g., how are the needs of personas with different career paths being addressed; what is needed to advance recognition of RSEs; what improvements are needed to link infrastructure to recognition systems?

More detailed analysis is needed to fully answer these questions, particularly the last point. The maturity matrices developed by Science Europe in relation to sustainable research data could be applied here to nuance understanding of the status of people-related challenges in the research software community (Boccali et al., 2021). The Science Europe matrices present three progression steps for each of the key areas they identified:

- Plans to develop: The organisation has acknowledged the need to take action in a given area and is developing/has developed plans on how to proceed.
- Development ongoing: The organisation has done the groundwork in a given area to achieve the sustainability of research data, though more refinement is needed.
- Developed on organisational level: The respective area is addressed on a mature level within the organisation.

The Future of Open Scholarship project being led by Invest in Open Infrastructure could also provide a valuable model for next steps. Their work involved similar steps in engaging stakeholders to identify and analyse the current landscape, then utilised co-design frameworks to make recommendations on a range of interventions that would address identified short-term outcomes and long-term impacts (Goudarzi & Thaney, 2020).

ReSA will present the outcomes of the People Roadmap to the research software community at the Workshop on Sustainable Software Sustainability (WoSSS) on 8 October 2021, and in two public webinars. This will enable an opportunity for more inputs and validation of the analysis and encourage the community to consider how to to increase knowledge transfer or collaboration on initiatives of relevance. The People Roadmap has enabled piloting of an approach to mapping the research software ecosystem, and it would be valuable to extend this to include the many other relevant initiatives to better understand linkages between strategic goals, projects, collaborators and funders. The next phase of development should also encourage representation from the Global South.

Acknowledgements

ReSA is grateful to the Wellcome Trust for support to complete this consultation and analysis, all of the project participants for their involvement, and the ReSA Steering Committee for their input.

References

Abbott, A. (1988). *The System of Professions*. University of Chicago Press. https://press.uchicago.edu/ucp/books/book/chicago/S/bo5965590.html

AGU. (2021). *Data and Software for Authors*. https://www.agu.org/Publish-with-AGU/Publish/Author-Resources/Data-and-Software-for-Authors

Anzt, H., Bach, F., Druskat, S., Löffler, F., Loewe, A., Renard, B. Y., Seemann, G., Struck, A., Achhammer, E., Aggarwal, P., Appel, F., Bader, M., Brusch, L., Busse, C., Chourdakis, G., Dabrowski, P. W., Ebert, P., Flemisch, B., Friedl, S., … Weeber, R. (2021). An environment for sustainable research software in Germany and beyond: Current state, open challenges, and call for action. *F1000Research*, *9*, 295. https://doi.org/10.12688/f1000research.23224.2

Arafune, L., Brunson, D., Hacker, T., & Smith, P. (2020). *Building the Research Innovation Workforce: A workshop to identify new insights and directions to advance the research computing community.* https://www.rcac.purdue.edu/ciworkforce2020/report/report.pdf

Boccali, T., Sølsnes, A. E., Thorley, M., Winkler-Nees, S., & Timmermann, M. (2021). *Practical Guide to Sustainable Research Data*. https://doi.org/10.5281/ZENODO.4769703

CHAOSS. (2020). *Diversity and Inclusion Badging*. https://github.com/badging/diversity-and-inclusion

Cohen, J., Katz, D. S., Barker, M., Chue Hong, N., Haines, R., & Jay, C. (2021). The Four Pillars of Research Software Engineering. *IEEE Software*, *38*(1), 97–105. https://doi.org/10.1109/MS.2020.2973362

DLR. (2019). *References*. https://rse.dlr.de/02_references.html

EOSC. (2021a). *Charter of the EOSC Task Force on Infrastructures for Quality Research Software*. https://www.eosc.eu/sites/default/files/tfcharters/eosca_tfinfrastructureforqualityresearchsoftware_draftcharter_20210614.pdf

EOSC. (2021b). *Strategic Research and Innovation Agenda of the European Open Science Cloud* (Version 1.0 15). https://www.eosc.eu/sites/default/files/EOSC-SRIA-V1.0_15Feb2021.pdf

European Commission. (2020). *Scholarly infrastructures for research software: Report from the EOSC Executive Board Working Group (WG) Architecture Task Force (TF) SIRS.* Publications Office. https://data.europa.eu/doi/10.2777/28598

GitHub. (2021a). *About CITATION files*. https://docs.github.com/en/repositories/managing-your-repositorys-settings-and-features/customizing-your-repository/about-citation-files

GitHub. (2021b). *Global diversity, inclusion, and belonging at GitHub*. https://github.com/about/diversity

GitHub. (2021c). *Who's using GitHub?* https://government.github.com/community/

Goudarzi, S., & Thaney, K. (2020). *Future of Open Scholarship: Preliminary Findings*. Zenodo. https://doi.org/10.5281/ZENODO.5153627

Hettrick, S., Antonioletti, M., Carr, L., Chue Hong, N., Crouch, S., De Roure, D., Emsley, I., Goble, C., Hay, A., Inupakutika, D., Jackson, M., Nenadic, A., Parkinson, T., Parsons, M. I., Pawlik, A., Peru, G., Proeme, A., Robinson, J., & Sufi, S. (2014). *UK Research Software Survey 2014* [Data set]. Zenodo. https://doi.org/10.5281/zenodo.14809

Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). Bibliometrics: The Leiden Manifesto for research metrics. *Nature*, *520*(7548), 429–431. https://doi.org/10.1038/520429a

Honeyman, T., & Treloar, A. (2021). *A National Agenda for Research Software*. https://doi.org/10.5281/ZENODO.4940273

INTERSECT. (2021). *Existing Research Software Engineering Training Material*. https://intersect-training.github.io/training-links/

Katz, D. S., Barker, M., Martinez, P. A., Anzt, H., Gonzalez-Beltran, A., & Bakker, T. (2020). *The Research Software Alliance (ReSA) and the community landscape*. https://doi.org/10.5281/ZENODO.3699950

Katz, D. S., Chue Hong, N. P., Clark, T., Muench, A., Stall, S., Bouquin, D., Cannon, M., Edmunds, S., Faez, T., Feeney, P., Fenner, M., Friedman, M., Grenier, G., Harrison, M., Heber, J., Leary, A., MacCallum, C., Murray, H., Pastrana, E., … Yeston, J. (2021). Recognizing the value of software: A software citation guide. *F1000Research*, *9*, 1257. https://doi.org/10.12688/f1000research.26932.2

Moher, D., Bouter, L., Kleinert, S., Glasziou, P., Sham, M. H., Barbour, V., Coriat, A.-M., Foeger, N., & Dirnagl, U. (2020). The Hong Kong Principles for assessing researchers: Fostering research integrity. *PLOS Biology*, *18*(7), e3000737. https://doi.org/10.1371/journal.pbio.3000737

Monteil, A., Gonzalez-Beltran, A., Ioannidis, A., Allen, A., Lee, A., Bandrowski, A., Wilson, B. E., Mecum, B., Du, C. F., Robinson, C., Garijo, D., Katz, D. S., Long, D., Milliken, G., Ménager, H., Hausman, J., Spaaks, J. H., Fenlon, K., Vanderbilt, K., … Morrell, T. (2020). Nine Best Practices for Research Software Registries and Repositories: A Concise Guide. *ArXiv:2012.13117 [Cs]*. http://arxiv.org/abs/2012.13117

NLeSC. (2021). *Research Software Directory*. https://research-software.nl/

Nosek, B. (2019). *Strategy for Culture Change*. https://www.cos.io/blog/strategy-for-culture-change

OECD. (2020). *Building digital workforce capacity and skills for data-intensive science* (OECD Science, Technology and Industry Policy Papers No. 90; OECD Science, Technology and Industry Policy Papers, Vol. 90). https://doi.org/10.1787/e08aa3bb-en

OECD. (2021). *Recommendation of the Council concerning Access to Research Data from Public Funding*. https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0347

Ram, K., Carver, J., Gesing, S., Katz, D. S., & Weber, N. (2021). *URSSI Implementation Plan*. https://plan.urssi.us

Research Data Alliance. (2020, April 23). *RDA COVID-19 Guidelines and Recommendations*.

RDA. https://www.rd-alliance.org/group/rda-covid19-rda-covid19-omics-rda-covid19-epidemiology-rda-covid19-clinical-rda-covid19-0

San Francisco Declaration on Research Assessment. (2013). https://sfdora.org/read/

Smith, A. M., Katz, D. S., Niemeyer, K. E., & FORCE11 Software Citation Working Group. (2016). Software citation principles. *PeerJ Computer Science*, *2*, e86. https://doi.org/10.7717/peerj-cs.86

Society of RSE. (2020). *The Society of Research Software Engineering: Strategy 2020/2021*. https://society-rse.org/the-society-of-research-software-engineering-strategy-2020-2021/

Software Sustainability Institute. (2019). *Policy research*. https://www.software.ac.uk/tags/policy-research

Sorbonne Declaration on Research Data Rights. (2020). https://sorbonnedatadeclaration.eu/

The Carpentries. (2020). *Our Strategic Plan for 2020-2025*. https://carpentries.org/strategic-plan/

The Carpentries. (2021a). *About Us*. https://carpentries.org/about/

The Carpentries. (2021b). *Assessment and Impact*. https://carpentries.org/assessment/

UNESCO. (2021). *Draft text of the UNESCO Recommendation on Open Science*. https://unesdoc.unesco.org/ark:/48223/pf0000378381.locale=en

US-RSE. (2021). *Mission*. https://us-rse.org/mission/

von Kurnatowski, L., Stoffers, M., Weigel, M., Meinel, M., Wasser, Y., Rack, K., & Fiedler, H. (2020). Scientific Software Engineering: Mining Repositories to gain insights into BACARDI. *2020 IEEE Aerospace Conference*, 1–10. https://doi.org/10.1109/AERO47225.2020.9172261

VSNU, NFU, KNAW, NWO, & ZonMw. (2019). *Room for everyone's talent*. http://vsnu.nl/recognitionandrewards/wp-content/uploads/2019/11/Position-paper-Room-for-everyone%E2%80%99s-talent.pdf

Woolston, C. (2021). Impact factor abandoned by Dutch university in hiring and promotion decisions. *Nature*, *595*(7867), 462–462. https://doi.org/10.1038/d41586-021-01759-5

Appendix 1: Profiled initiatives

This is a list of the initiatives profiled for this project, and the interviewees for each (where applicable).

1. American Geophysical Union (AGU) - Chris Erdmann and Shelley Stall
2. Australian Research Data Commons (ARDC) - Andrew Treloar
3. CiteAs - James Howison
4. Community Health Analytics Open Source Software (CHAOSS) - Matt Germonprez and Sean Goggins
5. Code for Science & Society (CS&S) - Rayya El Zein, Jessica Hardwick and Danielle Robinson
6. CodeMeta - online sources
7. Code Refinery - Radovan Bast, Richard Darst and Anne Claire Fouilloux
8. Council of International RSE Associations - online sources
9. Data Archiving and Networked Services (DANS) - Gerard Coen
10. de-RSE - Stephan Druskat and Frank Löffler
11. German Aerospace Center (DLR) Institute for Software Technology - Carina Haupt
12. Department of Energy (DOE) Exascale Computing Project (ECP) - Mike Heroux
13. ELIXIR - Peter Maccallum and Fotis Psomopoulos
14. European Research Infrastructure Software Engineers' Network (EURISE Network) - Carsten Thiel
15. European Open Science Cloud (EOSC) - Mathew Dovey
16. FORCE11 Software Citation Implementation Working Group (SCIWG) - Neil Chue Hong and Martin Fenner
17. GitHub - Arfon Smith
18. INnovative Training Enabled by a Research Software Engineering Community of Trainers (INTERSECT) - Ian Cosden and Jeffrey Carver
19. Journal of Open Research Software (JORS) - Matthew Turk
20. Journal of Open Source Software (JOSS) - Kyle Niemeyer
21. Netherlands eScience Center (NLeSC) - Carlos Martinez-Ortiz
22. New Zealand eScience Infrastructure (NeSI) - Nick Jones, Nooriyah Lohani and Georgina Rae
23. Nordic-RSE - Radovan Bast, Richard Darst and Anne Claire Fouilloux
24. Society of RSE - Paul Richmond and Claire Wyatt
25. Software Sustainability Institute (SSI)
26. The Carpentries - Alycia Crall, Toby Hodges and Kari Jordan
27. US Research Software Engineer Association (US-RSE) - Ian Cosden and Sandra Gesing
28. US Research Software Sustainability Institute (URSSI) - online sources

Appendix 2: Initiative profile questions

Each initiative was profiled using the following questions:

1. Organisational mission and vision.

2. Strategic goals relevant to the five people-themes above.

3. Current and future projects relevant to the five people-themes

4. Collaborators on these projects. (This doesn't have to be a list of all organisations you collaborate with, it's more about major partners or types of collaborators e.g., industry, universities, government.)

5. What is the source of your resourcing for these projects? (Where does funding for people-themed projects come from e;g., membership fees, government, philanthropic grants? How much of the organisational budget is spent on people-themed projects?)

6. Other initiatives involved with (may be less formal collaborations than those discussed in the previous question on main projects).

7. What information/evidence (data, policies) is used to support investment in this area? (How do you track impact or outcomes?)

8. What does your organisation consider the priority areas in the people-theme for the research software community as a whole? (This could include priorities that you think other organisations in the sector should address, or things you'd do with unlimited resources.)

Appendix 3: Cognitive City data

Data can be downloaded from
https://docs.google.com/spreadsheets/d/1XqTxFWDW-YowFiqckS9t-DvATgbIe7i2JI8PCGz3YSE/edit#gid=0

Name	Strategic Goal	Collaborators	Funded by	Administers
American Geophysical Union (AGU)	To support and inspire a global community of individuals and organizations interested in advancing discovery in Earth and space sciences and its benefit for humanity and the environment.	Earth Science Information Partners (ESIP), Research Data Alliance (RDA), European Geophysical Union (EGU), Software Heritage, NASA, Jupyter Notebooks, GitHub, FORCE11, The Carpentries, GO FAIR US, Binder, Code Ocean		Incentives, Skills/training, Communities, Infrastructure, Policy
Australian Research Data Commons (ARDC)	To accelerate research and innovation by driving excellence in the creation, analysis and retention of high-quality data assets.	Research Software Alliance (ReSA), RSE-AUNZ, SSI, The Carpentries, Research Data Alliance (RDA)		Incentives, Skills/training, Communities, Infrastructure, Policy
CHAOSS	CHAOSS is a Linux Foundation project focused on creating analytics and metrics to help define community health	Linux Foundation, GitHub, Industry	Sloan Foundation, Ford Foundation, Mozilla, Chan Zuckerberg Initiative (CZI), Red Hat	Incentives, Skills/training, Communities, Infrastructure

CiteAs	CiteAs is an effort to improve the visibility of research software work, in particular the ability to cite and make requests for citation. At a high level, it aims to change the incentive environment so the recognition for people doing this work changes.	ImpactStory	Sloan Foundation	Incentives, Skills/training, Infrastructure
Code for Science & Society (CS&S)	We are a US-based 501(c)(3) nonprofit supporting open collaboration in public interest technology through fiscal sponsorship and other programs supporting sustainable open source. We envision a future where communities build technology infrastructure that reflect their values.	Community Initiatives, NumFOCUS, CSV conference, Center for Scientific Collaboration and Community Engagement (CSCCE)	Sloan Foundation, Ford Foundation, Moore Foundation, Mozilla, Omidyar, Open Science Foundation (OSF)	Skills/training, Communities

CodeMeta	The CodeMeta project strives to promote the citation and reuse of software authored for scientific research by developing a mechanism to assist the transfer of software and software metadata between the entities that author, archive, index and distribute and use the software.	Figshare, GitHub, DataONE, Zenodo, DataCite, EarthCube, Software Ontology		Incentives, Infrastructure
Code Refinery	We are working with students, researchers, Research Software Engineers from all disciplines and national e-infrastructure partners to advance FAIRness of Software management and development practices so that research groups can collaboratively develop, review, discuss, test, share and reuse their codes.	The Carpentries, Netherlands eScience Center (NLeSC), European Software Sustainability Institute (EUSSI), Estonian Scientific Computing Infrastructure (ETAIS), UNINETT Sigma2, Aalto University, TU Delft University, Swedish National Infrastructure for Computing (SNIC) , IT Center for Science (CSC), Danish e-Infrastructure Cooperation (DeiC), Nordic Five Tech (KTH), Nordic	University of Oslo, Aalto University, TU-Delft University, Nordic e-Infrastructure Collaboration (NeIC)	Roles/careers/recognition, Skills/training, Communities

		e-Infrastructure Collaboration (NeIC), EuroHPC Competence Centre, University of Oslo		
Data Archiving and Networked Services (DANS)	DANS is the Dutch national centre of expertise and repository for research data. We help researchers make their data available for reuse. This allows researchers to use the data for new research and makes published research verifiable and reproducible	Software Heritage, Knowledge Exchange, Software Sustainability Institute (SSI), Netherlands eScience Center (NLeSC)	European Open Science Cloud (EOSC)	Incentives, Skills/training, Communities, Infrastructure
de-RSE	It is our mission to represent the German Research Software Engineers as a community, in all relevant contexts. It is our vision to establish research software as a first-class citizen in academia, establish and support careers for RSEs across institutions, support the education	German Society for Computer Science, National Research Data Infrastructure (NFDI), Software Sustainability Institute (SSI), International Council of RSE Associations, FORCE11, The Carpentries, DLR Institute for Software Technology	Amazon Web Services (AWS), Microsoft, GitLab, R Consortium, German Python Software Verband (PUG)	Policy, Incentives, Skills/training, Communities

	relevant for Research Software Engineering to increase the quality of research software, and grow and facilitate the Research Software Engineering community in Germany.			
Department of Energy Exascale Computing Project (DoE ECP)	Developing special software needed for exascale platforms	Software Sustainability Institute (SSI), Department of Energy (DoE), US Research Software Sustainability Institute (URSSI), National Science Foundation (NSF), National Oceanic and Atmospheric Administration (NOAA), National Institutes of Health (NIH), Industry, Society of RSE	Sloan Foundation	Incentives, Skills/training, Communities

DLR Institute for Software Technology	DLR is the Federal Republic of Germany's research centre for aeronautics and space. The mission of the Institute for Software Technology is research and development in software engineering technologies, and the incorporation of these technologies into DLR software projects. Current activities focus on software for distributed systems and intelligent systems, artificial intelligence, software technologies for embedded systems, visualization, and high-performanc e computing, and digitalisation.	de-RSE, Helmholtz Association, Software Sustainability Institute (SSI), Netherlands eScience Center (NLeSC), Society of RSE, US-RSE		Incentives, Skills/training, Communities, Infrastructure
ELIXIR	ELIXIR coordinates and develops life science resources across Europe so that researchers can more easily find, analyse and share data, exchange	European Union		Incentives, Skills/training, Communities, Policy, Infrastructure

	expertise, and implement best practices. This makes it possible for them to gain greater insights into how living organisms work.			
European Open Science Cloud (EOSC)	The ambition of the European Open Science Cloud (EOSC) is to provide European researchers, innovators, companies and citizens with a federated and open multi-disciplinary environment where they can publish, find and re-use data, tools and services for research, innovation and educational purposes		European Union	Incentives, Skills/training, Communities, Infrastructure, Policy
EURISE Network	The EURISE Network has been formed by the Social Sciences and Humanities ERICs CESSDA, CLARIN & DARIAH, with OPERAS, to create an umbrella where research infrastructures meet research software engineers.	Social Sciences & Humanities Open Cloud (SSHOC), Code Refinery		Incentives, Communities

FORCE11 Software Citation Implementation Working Group (SCIWG)	A FORCE11 Working Group aiming to achieve adoption of software citation principles and consider how to implement them in different contexts. Our vision is that people who develop software are given the scholarly credit they deserve by having their software directly cited.	Research Data Alliance (RDA), Software Heritage, Zenodo, GitHub, Figshare, CHORUS	Institute of Museum and Library Services (IMLS)	Policy, Incentives, Skills/training, Communities
GitHub	Be the home for all developers. Accelerate human progress through developer collaboration.	NASA, Software Heritage		Policy, Skills/training, Community, Infrastructure
International Council of RSE Associations	The International Council of RSE Associations provides a formal open forum for established national and multinational RSE associations to talk and coordinate regularly, and thus sustain international collaboration.	Society of RSE, de-RSE, US-RSE, RSE-AUNZ, NL-RSE, BE-RSE, Nordic-RSE		Incentives, Policy, Communities

INTERSECT	To address the lack of software development and engineering training for intermediate and advanced developers of research software, we are developing Innovative Training Enabled by a Research Software Engineering Community of Trainers (INTERSECT)	The Carpentries, Molecular Sciences Software Institute (MolSSI), US Research Software Sustainability Institute (URSSI), Women in HPC, Better Scientific Software (BSSw), US-RSE	National Science Foundation (NSF)	Incentives, Skills/training, Communities
JORS	The Journal of Open Research Software (JORS) features peer reviewed Software Metapapers describing research software with high reuse potential. We are working with a number of specialist and institutional repositories to ensure that the associated software is professionally archived, preserved, and is openly available. Equally importantly, the software and the papers will be	Ubiquity Press, WSSSPE		Policy, Incentives, Skills/training, Infrastructure

	citable, and reuse will be tracked. JORS also publishes full-length research papers that cover different aspects of creating, maintaining and evaluating open source research software. The aim of the section is to promote the dissemination of best practice and experience related to the development and maintenance of reusable, sustainable research software.			
JOSS	The Journal of Open Source Software (JOSS) is an academic journal with a formal peer review process that is designed to improve the quality of the software submitted.	CrossRef, Open Source initiative, NumFOCUS, ,ROpenSci, PyOpenSci, American Astronomical Society (AAS), GitHub, JuliaCon, SciPy	Sloan Foundation	Policy, Incentives, Skills/training, Infrastructure

New Zealand eScience Infrastructure (NeSI)	NeSI helps researchers, institutions and universities conduct successful research endeavours by providing expertise and capability in computational and data intensive research	RSE-AUNZ, The Carpentries, Women in HPC, Australian Research Data Commons (ARDC), Research Software Alliance (ReSA)		Policy, Incentives, Skills/training, Communities
Netherlands eScience Center (NLeSC)	A robust research community, in which all investigators in all domains are able to exploit advanced digital technologies to answer curiosity-driven questions, keeping the Netherlands at the forefront of cutting- edge international research.	NL-RSE, Turing Institute, Association of Universities in the Netherlands (VSNU), Dutch Research Council (NWO), SURF, Dutch Techcentre for Life Sciences (DTL), Zenodo, CodeMeta, The Carpentries, Research Software Alliance (ReSA), Plan-E, GO FAIR, National Programme Open Science (NPOS)	SURF, Dutch Research Council (NWO), European Union	Incentives, Skills/training, Communities, Infrastructure, Policy
Nordic-RSE	Non-profit organisation to bring together everyone interested in Research Software Engineer activities	de-RSE, Society of RSE, International Council of RSE Associations, NL-RSE, US-RSE, Aalto University, BE-RSE		Policy, Incentives, Skills/training, Communities

Society of RSE	The Society of Research Software Engineering was founded on the belief that a world which relies on software must recognise the people who develop it. Our mission is to establish a research environment that recognises the vital role of software in research. We work to increase software skills across everyone in research, to promote collaboration between researchers and software experts, and to support the creation of an academic career path for Research Software Engineers.	Research Software Alliance (ReSA), International Council of RSE Associations, Software Sustainability Institute (SSI), The Carpentries, US-RSE, Engineering and Physical Sciences Research Council (EPSRC)		Policy, Incentives, Skills/training, Communities
Software Sustainability Institute (SSI)	The Software Sustainability Institute cultivates better, more sustainable, research software to enable world-class research. We	Society of RSE, The Carpentries	JISC, UK Research and Innovation (UKRI)	Incentives, Skills/training, Communities, Policy

	help people build better software, and we work with researchers, developers, funders and infrastructure providers to identify key issues and best practice in scientific software.			
The Carpentries	Our vision is to be the leading inclusive community teaching data and coding skills.	Understood, National Association of Multicultural Engineering Program Advocates (NAMEPA), National Society of Black Engineers (NSBE), National GEM Consortium (GEM), Society for Advancement of Chicanos/Hispanics and Native Americans in Science (SACNAS), Center for Scientific Collaboration and Community Engagement (CSCCE), Turing Institute, South African Centre for Digital Language Resources (SADiLaR), Environmental Data Science Inclusion Network (EDSIN)	Sloan Foundation, California Digital Library (CDL), Chan Zuckerberg Initiative (CZI), Institute of Museum and Library Services (IMLS), Moore Foundation, Mozilla, R Consortium, Code for Science & Society (CS&S)	Incentives, Skills/training, Communities

US-RSE	The US-RSE Association is centered around four main goals: community, advocacy, resources, and diversity, equity and inclusion	Society of RSE, Better Scientific Software (BSSw), International Council of RSE Associations, Campus Research Computing Consortium (CaRCC), HPC Facilitators, Open Collective Foundation (OCF)	Sloan Foundation	Roles/careers/recognition, Skills/training, Community
US Research Software Institute (URSSI)	Developing a pathway to research software sustainability	US-RSE	National Science Foundation (NSF)	Policy, Incentives, Skills/training, Communities