

ANNUAL Report 2023

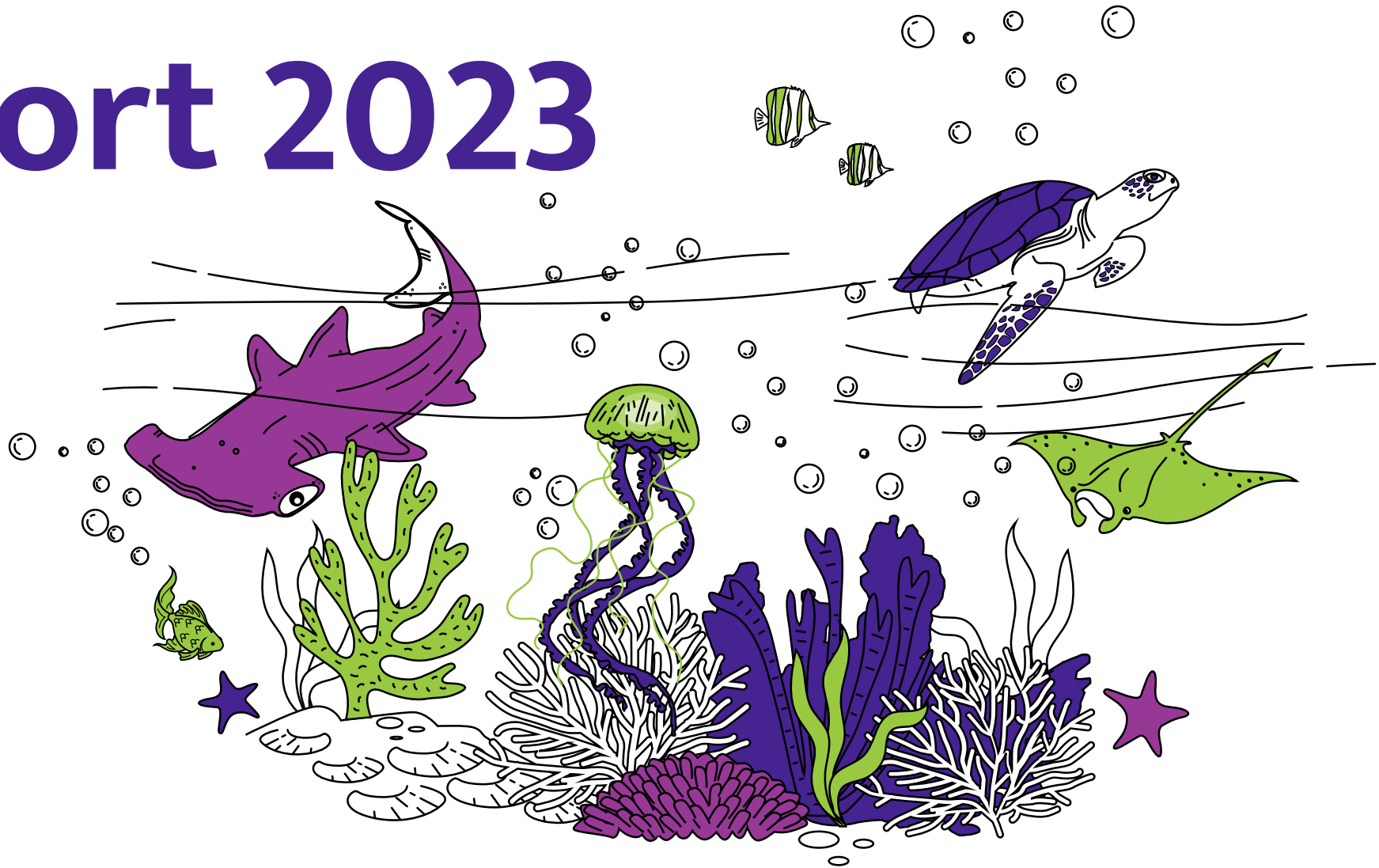


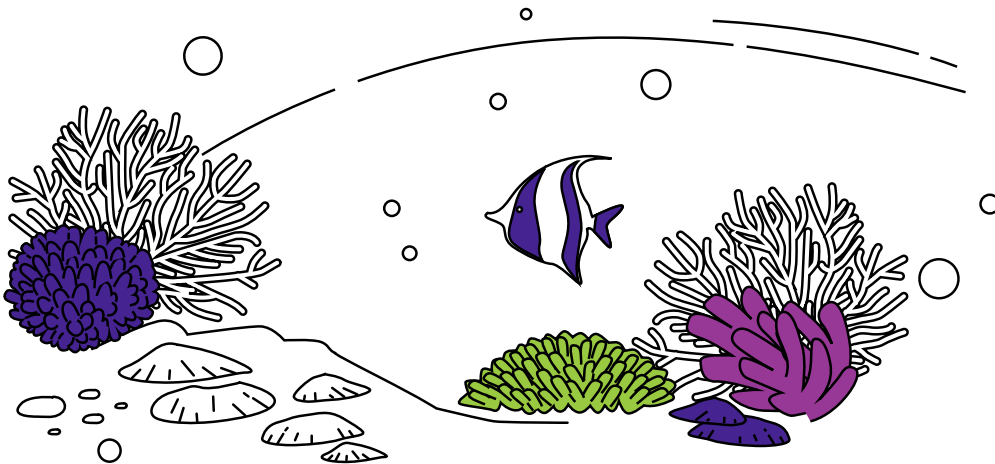
Table Of Contents

- Letter From Our Directors** 3
- About Us** 4
 - Our Mission 5
 - Our Model 6
 - Our Purpose 7
- We Are Labs** 8
- Sustainable Ecosystem** 9
 - OSS Projects We Contribute To 10
 - Project Highlights 12
 - Community Highlights 14
 - Leadership Roles in OSS Projects 15
 - Stories: Contributor Experience 16
 - Stories: JupyterLab Accessibility 17
 - Stories: Consortium for Python Data API Standards 18
 - Stories: Nebari And conda-store 20
 - Stories: Open Source Design 21
 - Community Participation And Events 22
 - Internship Program 23
 - Interns, Mentors, And Projects 24
 - Healthy Mentorship Programs 25
 - How Labs Internship Works 26
 - Collaboration With Christ Church 27
- Labs Work Culture** 28
 - Community-First Working Model 29
 - Co-Ownership Governance Model 31
 - 2023 Focus Areas 31
 - Python Packaging 32
 - Access Centered Practices 33
 - Multiple Array / Tensor Library Support 34
 - Financials 35
 - Diversity Within Labs 36
 - Building a Sense of Belonging 38
- How Far We've Come And Looking Ahead** 39
- Acknowledgements** 40
- Closing Page** 41

Letter From Our Directors

We founded Quansight Labs five years ago to challenge, reimagine, and redesign existing approaches to funding, sustaining, and developing open source infrastructure, tooling, and communities. Over those five years, Quansight Labs has grown into an efficient and impactful organization enabling and conducting essential work in the scientific computing open source ecosystem.

As with any great endeavor, we would not have been able to do this alone. Much of this growth results from a deliberate approach to building, nurturing, and strengthening partnerships and collaborations with key stakeholders — including funding bodies, industrial partners, non-profit organizations, service providers, open source communities and leaders, and many more interested groups.



The last couple of years have been particularly challenging in the tech industry. However, keeping our eyes and minds on our vision for the organization has allowed us to turn these into years of steady growth and outsized impact at Quansight Labs. We have continued to build an exceptionally talented, diverse, and thoughtful team that shares our vision of building sustainable open source communities and projects. We have also remained laser-focused on our goal of positively impacting our ecosystem.

During the past year, we continued to improve our organizational model and processes. In the spirit of our Feminist co-leadership model and its democratic principles of inclusiveness, participation, and empowerment, we have identified and supported emergent leaders within strategic areas for our organization and communities — from Data API standards and packaging to documentation, accessibility, and developer experience. We will continue to refine and build (and report, of course) on the learnings from this approach in the years to come.

Once again, we are proud to share highlights from the Quansight Labs team's work over the past year, the new learnings we encountered along the way, and a preview of our goals for this year.

Tania Allard & Ralf Gommers
Directors, Quansight Labs

About Us



Our Mission

The mission of Quansight Labs is to sustain and grow community-driven open source projects and ecosystems, focusing on the core of the PyData stack and tools and digital infrastructure for data science, Machine Learning (ML) and Artificial Intelligence (AI), and scientific computing.

Open source has become an integral part of our daily lives. Millions of people worldwide use open source software and infrastructure day in and day out.

We fulfill our mission through these activities:

1. Quansight Labs facilitates advances and innovations in science, research, education, government, and industry.

- Our team comprises individuals with diverse backgrounds and skills, including developers, community leaders, designers, accessibility experts, and technical writers. Together, we support the scientific computing ecosystem and its ever-growing community.
- We develop and maintain open source software for ML/AI, data science, scientific research and technical communication. Through this, we foster the creation and dissemination of knowledge.
- We contribute to many community-driven open source projects and support organizations like NumFOCUS to support the long-term sustainability of the ecosystem.

2. Quansight Labs empowers the people who make up the open source community and fosters growth and inclusion.

- We run an annual three-month-long open source internship program for individuals from historically underrepresented groups interested in contributing to and participating in open source.
- Quansight Labs members are actively involved in community building through organizing community calls, contribution sprints, mentoring individuals, organizing community events, and helping coordinate resources and initiatives within the open source ecosystem.
- Quansight Labs has built an access-centered practice to help make scientific computing tools accessible to a broader range of individuals, particularly those with disabilities.

Our Model

A core focus of Quansight Labs is to develop innovative yet sustainable funding mechanisms to live up to our mission — **sustaining community-driven open source projects and their communities**. As such, we have spent the last five years developing and improving our funding and operational models.

Through these, we ensure open source maintainers and community leaders have a significant amount of time dedicated to working on community-driven open source projects while also spending part of their time on consulting projects with Quansight's commercial clients. Such a model includes benefits such as learning and upskilling opportunities for our team, bringing back concrete user feedback and needs, and ideas for open source projects. It also gives us the flexibility to quickly scale up our contributions to the open source projects we support when we receive direct financial support for them.



Our Purpose

The majority of our efforts are aimed at maintaining and evolving existing open source projects that are critical to the PyData stack, including NumPy, Jupyter, SciPy, pandas, and conda-forge, among many others. We aim to give talented maintainers time for essential day-to-day maintenance and tackling complex problems that require weeks or months of dedicated time. Additionally, we complement this maintenance work with innovation activities to identify and fill critical gaps in the PyData and broader scientific computing ecosystems.

In doing so, we acknowledge that for open source projects and communities to be successful and sustainable, many skills and perspectives are required beyond software development. Expertise in community building, technical writing, graphic design, UI/UX design, marketing, accessibility, fundraising, and project management, among others, are essential to open source. Yet, many of these are often underrepresented or nonexistent in community-driven open source projects. At Quansight Labs, we hire designers, community managers, developer experience engineers, and many others interested in open source to grow the diversity of skills in open source communities and support the projects' growth, sustainability, and maturity. And at the same time, we spearhead and support grassroots and advocacy initiatives aimed at formalizing and recognizing the fundamental role these activities play in our ecosystem.

“My favorite part of the role is being able to collaborate with amazing people on projects I intrinsically care about.”

Marco Gorelli

pandas and Polars sustainer, Senior Software Engineer
at Quansight Labs

“It’s important for us to be able to plan for the future and for having healthy communities. That means we need to focus on making sure that projects are sustainable, that people can actually come and go from the projects, and be able to enjoy their time there. I love that about Quansight.”

Melissa Weber Mendonça

NumPy, SciPy, Matplotlib, pandas sustainer, Senior Developer
Experience Engineer at Quansight Labs

“Since I started being involved in the Python community, I’ve really cared about its longevity and overall success. Joining Quansight Labs has allowed me to work on impactful projects that are vital to the community, all while working alongside and learning from brilliant people!”

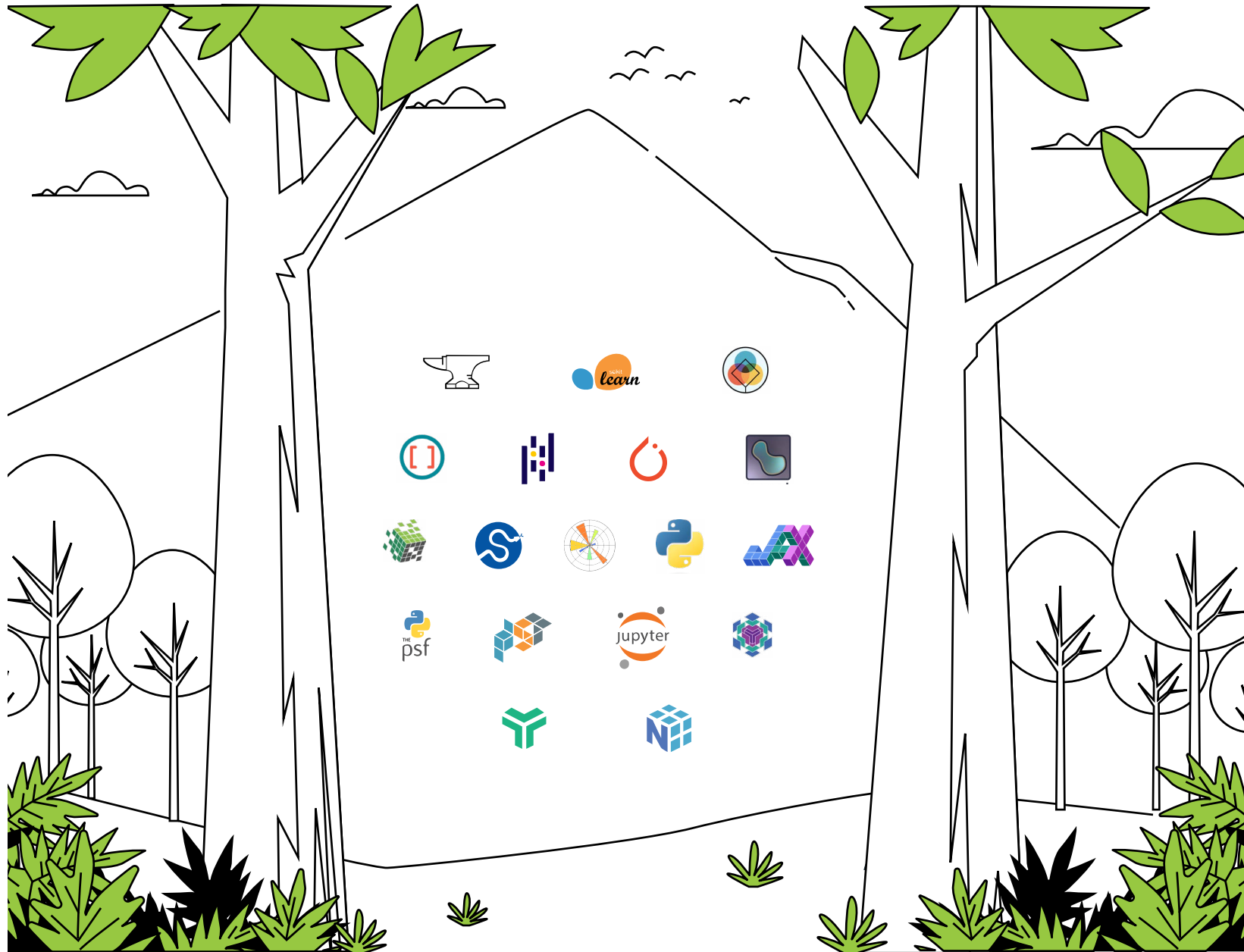
Lysandros Nikolaou

CPython and NumPy sustainer, Senior Software Engineer
at Quansight Labs

We Are Labs



Sustainable Ecosystem



OSS Projects We Contribute To

Our Quansight Labs team actively sustains and contributes to many foundational projects in the PyData and scientific computing ecosystems and beyond, including the Python programming language and [stdlib.js](#).

Last year, we sustained our contributions to the libraries at the core of the PyData ecosystem, including NumPy, pandas, SciPy, Matplotlib, scikit-learn, PyTorch, and JupyterLab to name but a few. We also extended beyond the core of PyData by contributing to high-performance computing, visualization projects, and infrastructure projects like Numba, Ray, Napari, and [Nebari](#).

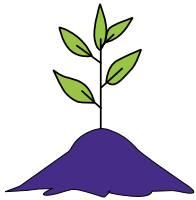
We also continued our already substantial support of the Python packaging ecosystem, from continuing our contributions to projects like conda, conda-forge, and meson-python to adopting new projects like [conda-store](#).

The main goal of these continued efforts remained improving the end-user and maintainer's experience and workflows by focusing on the tooling for packaging and distributing scientific software.

As part of our mission of building inclusive and sustainable projects and communities, Labs members continued participating in projects focused on Diversity, Equity, Inclusion, and Accessibility within our communities. For example, the [NumFOCUS DISC](#) committee aims to make the PyData projects, events, and overall initiatives more welcoming, safe, and accessible to historically underrepresented and underserved communities in our ecosystem. The [Jupyter Accessibility](#) project aims to improve the accessibility of tools and spaces within the Jupyter ecosystem for developers, contributors, and users.



OSS Projects We Contribute To

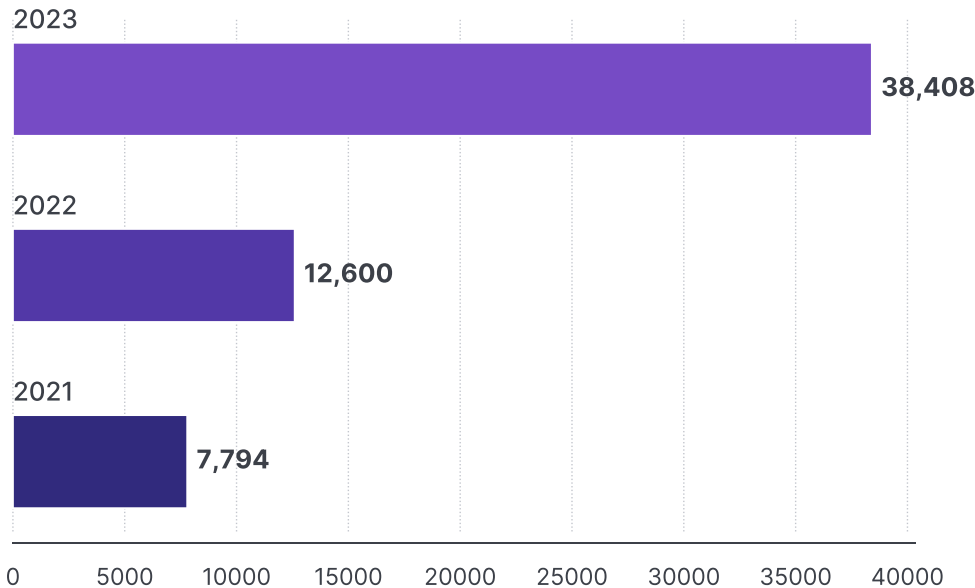


In 2023, we contributed to **35+** open source projects.



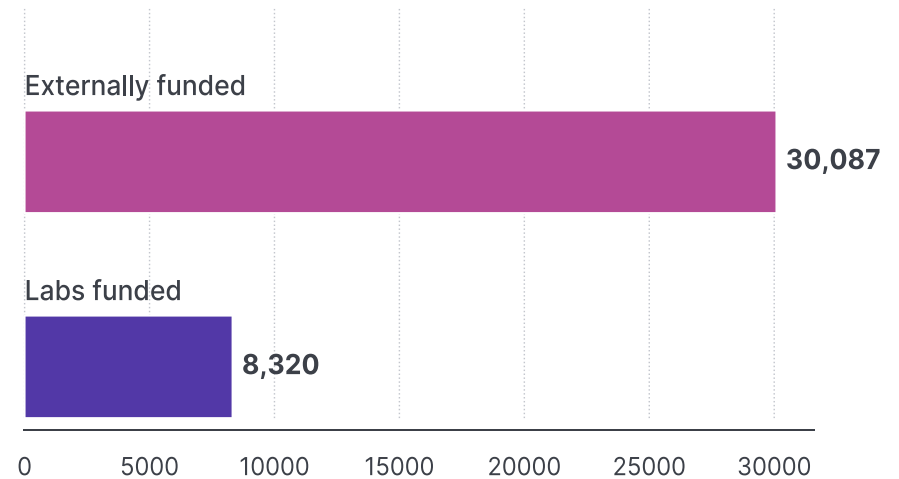
We spent **38,000+** hours working on these projects.

Hours spent in OSS project



We recently revisited how we account for hours spent on OSS projects - hours for 2022 and 2021 had been significantly underestimated.

Hours spent per funding source



Out of those **38,000+** hours worked in community open source projects, **~22%** (or 8320 hours) were directly funded by Quansight Labs.

Project Highlights

A lot of the impact of our work comes from the breadth and depth of our contributions to community-driven open source projects, as highlighted on the preceding pages. It is also worth celebrating individual project highlights that Labs team members were responsible for or made significant contributions to this year:

- The [OpenBLAS](#) team worked towards alignment on standardizing a 64-bit reference BLAS and LAPACK API.
- PyTorch 2.0 and TorchVision 0.15 included a complete revamp of the augmentation pipelines that now natively support object detection, semantic and instance segmentation, and more through the same interface.
- The CPython team released [Python 3.12](#) with better and more [flexible f-strings \(PEP 701\)](#).
- The Scientific Python project created a new Sphinx directive and UI elements for interactive documentation.
- [Accessible-pygments](#), a collection of accessible themes for syntax highlighting, was adopted by several documentation themes and became a critical project (i.e., top 1% installations) on PyPI.
- The [PyData Sphinx Theme](#) adopted a new color palette and design system developed with an accessibility-first approach and that conforms with WCAG 2.0 color contrast criteria.
- Napari gained support for asynchronous image-slicing functionality and layer control design.
- [Papyri](#) is working on the adoption of the MyST Abstract Syntax Tree (AST) to build documentation across several key PyData libraries.
- PyTorch-Ignite's Code-Generator web application [enhanced the code templates and added a Nebari integration](#).
- The Data APIs consortium launched and created an initial RFD for the DataFrame API standard, for cross-dataframe collaboration in the PyData ecosystem.
- The array-api-compat project, which provides a compatibility layer for the Array API standard, added support for PyTorch.
- The [meson-python](#) team did four feature releases, including the long-awaited support for editable installs in 0.13.0.
- [conda 23.10 adopted conda-libmamba-solver as the new default solver](#), dramatically reducing solver, times for all conda users.
- conda-forge created a [Streamlit app](#) to stream and serve conda package metadata online.
- NumPy's Python API setup was completely refactored with a single location for all public API symbols, and an adequately hidden private API.

- ☑ NumPy completed the transition away from `numpy.distutils` to `meson`, and is upstreaming improvements to Meson's BLAS/LAPACK detection.
- ☑ SciPy and scikit-learn added support for the Array API standard to part of their code base, with several SciPy submodules and scikit-learn estimators now working with PyTorch tensors and CuPy arrays — yielding significant performance gains.
- ☑ After three years of development, pandas versions 2.0 and 2.1 were released with [improvements to Copy-on-Write](#), [datetimes parsed in a consistent format](#), and [support for nullable dtypes](#), among [many other enhancements, bug fixes, and API stability improvements](#).
- ☑ The CuPy team implemented support for large parts of SciPy's signal processing and interpolation capabilities in CuPy and helped upstream `cuSignal` into CuPy — as described in [this blog post](#).
- ☑ The stdlib team released version 0.1, which was a major development milestone, including over 400 new packages, with added support for fancy indexing and slicing in `ndarrays`.
- ☑ Numba added support for [dynamic exceptions](#) (i.e., exceptions with runtime arguments), the [NumPy Polynomial class](#), and [several ufunc attributes along with `ufunc.reduce`](#).



Community Highlights

- ✓ OpenBLAS was moved into a new GitHub organization: [OpenMathLib](#), taking a step towards sustainability.
- ✓ CuSignal's license was changed from Apache-2 to MIT, to better facilitate upstream collaboration with CuPy.
- ✓ NumPy, SciPy, Matplotlib, and pandas focused on improving the new contributor experience, and [created a handbook for community reference](#).
- ✓ The Jupyter Accessibility team presented workshops and conference talks, and conducted semi-weekly calls to build awareness around accessibility best practices in the PyData community.
- ✓ The Napari team actively [contributed to the enhancement and maintenance of conda/constructor](#), an upstream conda library, which led to Constructor becoming the first community-maintained project of the conda organization.
- ✓ Conda-forge implemented a new security policy, and is in the process of [redesigning their website and documentation](#).
- ✓ NumPy is preparing its first major release since NumPy 1.0, which involved a lot of community-wide coordination and collaboration to make sure downstream is able to gracefully handle breaking changes. This effort was also aided by the scientific python community's effort to establish a community-wide repository for nightly uploads.
- ✓ [NumPy's documentation](#) was translated into Japanese and Portuguese, to reach more community members.



Leadership Roles in OSS Projects



Tania Allard was named a Jupyter Distinguished Contributor for the 2023 cohort. Distinguished contributors are recognized for their substantial and impactful contributions to the Jupyter ecosystem.



Nathan Goldbaum was added as maintainer of NumPy, after over six months of contributing to the NumPy C API.



Marco Gorelli was given commit rights to Polars, in recognition of several valuable volunteer contributions.



Gabriel Fouasnon became a member of the JupyterLab Council (maintainers with voting rights) and the elected Accessibility representative in the Jupyter Software Steering Council.



Dharhas Pothina joined the inaugural HoloViz Steering Council, which is now a NumFOCUS-sponsored project.

Members of the Consortium for Python Data API Standards published [“Python Array API Standard: Toward Array Interoperability in the Scientific Python Ecosystem”](#) in the Proceedings of the 22nd Python in Science Conference (SciPy 2023).

Stories: Contributor Experience

Getting started with open source contributions can be intimidating. Often, information about contributing is incomplete, scattered, or invisible, creating barriers for those looking to participate in open source communities.

With this in mind, we experimented with creating the role of “Contributor Experience Leads” to focus on onboarding contributors, organizing, documenting, and streamlining contribution processes for four open source communities: NumPy, SciPy, Matplotlib, and pandas.

This project, [funded by the CZI’s EOSS Program](#), involved several of our team members and community members in an effort to experiment and document ideas for improving contributor experience and, hopefully, lowering barriers for contributors from historically marginalized groups.

Our team of Contributor Experience Leads prioritized the following actions:

- ☑ Improving **contributor onboarding guides**, including, for example, PR review guidelines, internal roles and processes, policies and general guidance for new contributors and new maintainers.
- ☑ Improving **developer tooling** (CI, Cloud Development Environments).
- ☑ Setting up and hosting **community meetings and additional communication channels** for each project, as needed, with documented procedure.
- ☑ Establishing (and documenting) regular **mentoring sessions for new contributors** for each project.
- ☑ Collect and analyze **contributor activity and contributor demographic data** for each project. In collaboration with the CHAOSS project, we have developed [a draft of community health metrics](#) to be used in open source projects.

- ☑ In collaboration with our external DEI consultant, May Ireland, through interviews with project maintainers and key personnel, we have identified priorities to work on and [created a DEI action plan](#). All of this work has been organized around DEI Culture Labs, co-organized with May Ireland, with the goal of connecting with other communities facing the same issues and working on shared solutions.
- ☑ Established a **Code of Conduct working group** in collaboration with NumFOCUS.
- ☑ Launched an interview series to **highlight opportunities for professional development** and growth through contributing to the projects we focused on.
- ☑ Organizing sprints and participating in other **outreach activities** such as talks, live coding events, or events targeting historically underrepresented groups in open source.

Documenting these procedures across all projects has sometimes proved to be challenging in the absence of each project’s desire for governance changes and establishing written procedures. Flexibility is often seen as an advantage, and in many cases, the pathways to project leadership or maintainership are still unclear or decided on a case-by-case basis. However, the CEL team has mentored, advocated for, and successfully onboarded several new maintainers in each project.

Our [Contributor Experience handbook](#) is a comprehensive collection of the work we have done in these two years, and it is our hope that we can continue to further these discussions in an effort to both involve more projects and exchange experiences and ideas for the future.

Stories: Jupyter Accessibility

Despite [representing over one billion individuals globally](#), people with disabilities are frequently overlooked in the design and development of technology. Aabled developers often create projects without considering fundamental accessibility features, inadvertently excluding disabled users from the outset. This results in a self-perpetuating cycle: projects gain traction among abled users yet fail to attract disabled users due to inherent accessibility barriers. Over time, this creates an isolated “bubble” where the project’s community remains detached from the realities faced by individuals with disabilities. Unfortunately, this concerning trend is prevalent across the entire technology ecosystem, including open source, as evidenced by projects like Jupyter.

The Jupyter Accessibility project, which includes several Labs members, was funded as an attempt to bridge the disability inclusion gap within Project Jupyter. For the last two years, thanks to funding from the [Chan Zuckerberg Initiative \(CZI\)](#) and the [Space Telescope Science Institute \(STScI\)](#), the team focused on the technology — to make Jupyter tools more accessible and the people building that technology — to raise awareness around accessibility. Some noteworthy achievements over the past two years include:

1. To make the technology more accessible:

- ✓ Conducted [accessibility audits](#), as a preliminary step to identify and address accessibility issues.
- ✓ Created [custom, automated accessibility tests that run in CI](#) (to maintain accessibility standards).
- ✓ Created an [extension that adds accessible themes to JupyterLab](#).
- ✓ Created a package with [accessible themes for Pygments](#) to fix low contrast in code blocks.
- ✓ Improved JupyterLab’s compliance with standards through [increased color contrast](#), [visible focus states](#), [improved layouts](#), and many more improvements listed in the [project board](#).
- ✓ Conducted accessibility-focused usability [testing](#) and [surveys](#) to gather information and feedback for users of assistive technologies.

2. To create a more accessibility-minded community:

- ✓ Successfully argued for accessibility-related concerns to have representation and a vote in Jupyter’s governance.
- ✓ Delivered several [workshops with invited guests](#) and talks to share better accessibility practices with Jupyter and the broader scientific Python community.
- ✓ Published the [JupyterLab Accessibility Statement](#), [Accessibility: A JupyterLab Developer’s Guide](#), and in collaboration with STScI, [Notebook authoring accessibility checklist](#).

There is still much work to be done to meet international accessibility standards, but JupyterLab and other Jupyter projects are now developed with disabled people in mind.

Stories: Consortium for Python Data API Standards

The Scientific Python Ecosystem (SPE) landscape has shifted dramatically in recent years. The rise of deep learning and the emergence of new hardware has led to a proliferation of new libraries within the PyData array and dataframe ecosystem. This proliferation has been a boon to Python users, who now have a wealth of choices for libraries and frameworks for numerical computing, data science, machine learning, and deep learning, with new frameworks pushing forward the state of the art appearing every year.

One consequence of all this activity has been fragmentation in the fundamental building blocks—multidimensional arrays (also known as tensors) and dataframes—that underpin the SPE. This fragmentation comes with high costs, from reinvention and re-implementation of arrays and dataframes and associated application programming interfaces (APIs) to siloed technical stacks targeting only one array or dataframe library to the proliferation of user guides focused on how to convert between and interoperate among libraries.

In recognition of the growing need for array and dataframe interoperability, in May 2020, Quansight Labs, together with industry partners and maintainers of the key array and dataframe libraries, created the Consortium for Python Data API Standards (“the Consortium”).

Since its founding, the Consortium has worked in close collaboration with stakeholders across the SPE¹ to standardize a critical set of

¹ Direct stakeholders include the maintainers of Python array and dataframe libraries and organizations which sponsor library development. Indirect stakeholders include maintainers of libraries which consume array and dataframe objects (“consuming libraries”), developers of compilers and runtimes with array- and dataframe-specific functionality, and end users, such as data scientists and application developers.

functionalities: interchange protocols, vectorized APIs, linear algebra, fast Fourier transforms (FFTs), statistical reductions, array creation, and manipulation, broadcasting, and indexing. These efforts have led to several achievements:

- ✓ Initial publication of the [2021 revision](#) of the Array API standard.
- ✓ Standardization of an [array interchange protocol](#) based on DLPack.
- ✓ Standardization of a [dataframe interchange protocol](#).
- ✓ Initial [reference implementation](#) of the Array API Standard in NumPy.
- ✓ The release of the 2022 revision of the Array API Standard that adds support for complex number data types and FFTs.
- ✓ Creating of a [test suite](#) for measuring compliance with the Array API Standard.
- ✓ Creating of a [test suite](#) for measuring compliance with the dataframe interchange protocol.
- ✓ Creating of an Array API [compatibility layer](#) to facilitate Array API Standard adoption among array-consuming libraries.

Critically, NumPy, CuPy, PyTorch, JAX, Dask, and oneAPI have all committed to adopting the Array API Standard, allowing for consistent APIs and behavior across their public interfaces. And maintainers and key stakeholders of pandas, Arrow, Modin, Vaex, polars, cuDF, and others have all been hard at work defining the contours of a Dataframe API Standard.

Consequently, for 2023, the Consortium sought to continue building on the momentum of the previous years and further two key objectives:

1. Accelerate Array API Standard adoption among downstream array-consuming libraries and
2. Release an initial RFC for a dataframe API Standard.

For the former objective, Quansight Labs took a leading role by working with the scikit-learn and SciPy communities to adopt the Array API Standard and thus decouple their implementations from NumPy and expand support for non-CPU-based execution. We are proud to say that not only were we able to add support for CuPy and PyTorch arrays, but our efforts

enabled **20-50x performance gains over the original NumPy implementations**. These efforts provide critical validation for the importance of the Array API Standard and will serve as a foundation for our efforts to further facilitate Array API Standard adoption throughout 2024.

In comparison to arrays, dataframes have presented significant challenges for standardization, given their implementation heterogeneity and complexity. In fact, among dataframe libraries, the very definition of what constitutes a dataframe varies.

Despite all these challenges, Quansight Labs continued to play a leading role by collaborating closely with industry partners and key dataframe library stakeholders to reduce dataframe ecosystem fragmentation. These efforts culminated in a [draft dataframe standard](#) released in May for public review. Since the draft's release, we have received considerable positive feedback and continued to iterate on the standard's design to finalize an initial release in 2024.



Stories: Nebari And conda-store

Quansight Labs supports community-driven open source projects that align with our community and sustainability values. In 2023, two Quansight-incubated open source projects, [Nebari](#) and [conda-store](#), adopted a community-first governance and development model and transitioned to Quansight Labs-supported projects.

Nebari is a Jupyter-based project for the deployment and maintenance of collaborative data science platforms.

It migrated to a new home at github.com/nebari-dev, and the development team spent the last year establishing community-first governance principles, supporting external contributions, creating a public roadmap, and organizing regular community meetings. It also adopted a new plugin system for quickly integrating other libraries in the data science ecosystem, and there was a major focus on maintenance to upgrade the latest versions of JupyterHub, JupyterLab, Kubernetes, and more.

Nebari remains a strategic project for some of Quansight's consulting clients. This creates opportunities for funding and integration of user feedback. We understand it can sometimes be challenging to balance goals important for the community with organizational motivations toward funding. The team has been steadfast in its commitment to Labs' community-governed principle, where funding always follows the community roadmap.

In 2023, most of the sustenance and development was sponsored by the Chief Digital and Artificial Intelligence Office (CDAO) for the Joint AI Test Infrastructure Capability (JATIC) project. A lot of the work was also funded by OpenTeams' Open Source Assurance initiative. In total, the team spent over 2700 hours contributing to Nebari in a community-first manner.

conda-store is a toolkit for managing data science environments in collaborative settings. In July 2023, [conda-store joined the conda community as an incubated \(federated\) project](#), adopting conda governance and robust contribution and development processes.

Since then, the project has seen many significant improvements:

- ☑ Windows and macOS support for individual users, in addition to the pre-existing support for Linux, to better serve the open source community.
- ☑ Improving the graphical user interface (GUI), i.e., conda-store-ui, with stronger user-experience principles and a standardized design system that also meets [WCAG 2.0 \(Web Content Accessibility Guidelines\)](#).
- ☑ Documentation revamp for improved, organized, and up-to-date information for all conda-store projects in a single place following the [Diátaxis framework](#).
- ☑ Research and development of a minimum viable cross-platform reproducible environment mechanism, in collaboration with the broader packaging community.

Conda-store was able to rapidly transition and grow as a community-driven project thanks to two rounds of support from the [Sovereign Tech Fund \(STF\)](#) through their [Contribute Back Challenges program](#).

Stories: Open Source Design

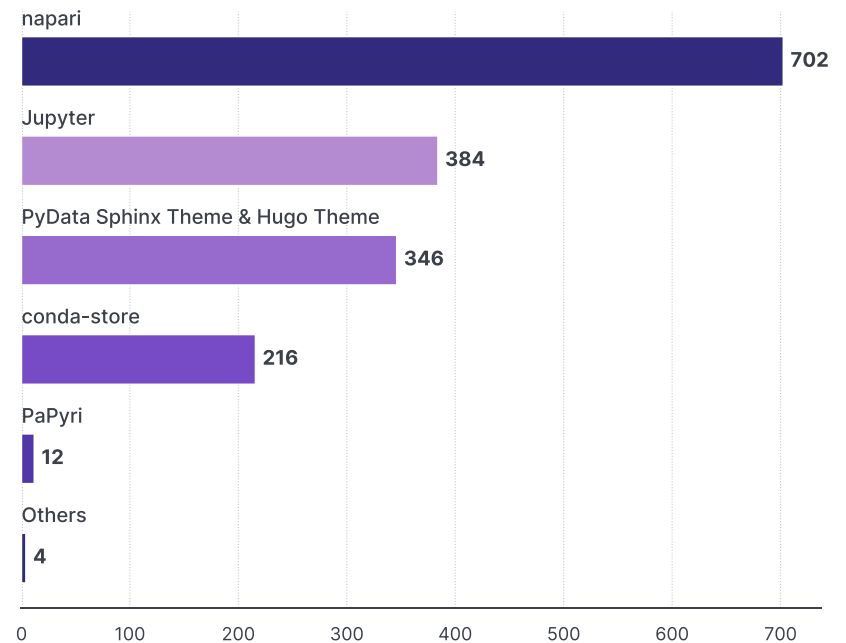
Regardless of the field or application, including design expertise — whether user experience (UX), user interface (UI), or graphic - helps create projects with consistent, accessible, and user-friendly interactions and presentations. Adopting a design-first approach to software development ensures it is also user-first, thus centering all aspects of software development on our users' needs, challenges, and motivations. Therefore, integrating design practices is particularly significant for the open source ecosystem due to its enormous scale, impact, and reach.

Open source design is an open, transparent, equitable, and collaborative design process aligned with open source software development and its ethos.

At Labs, open source design has become a non-negotiable and integral part of all our projects and initiatives. In 2023, we spent **over 1500 hours** making design-focused contributions to a range of projects like Napari - an interactive image analyzer desktop application; JupyterLab - a browser-based interactive development environment; accessible pygments, and PyData Sphinx Theme - a theme for project documentation.

Additionally, most of our design projects and approaches intersect with accessibility initiatives - some of which are described in other sections in this report.

Design hours spent per project



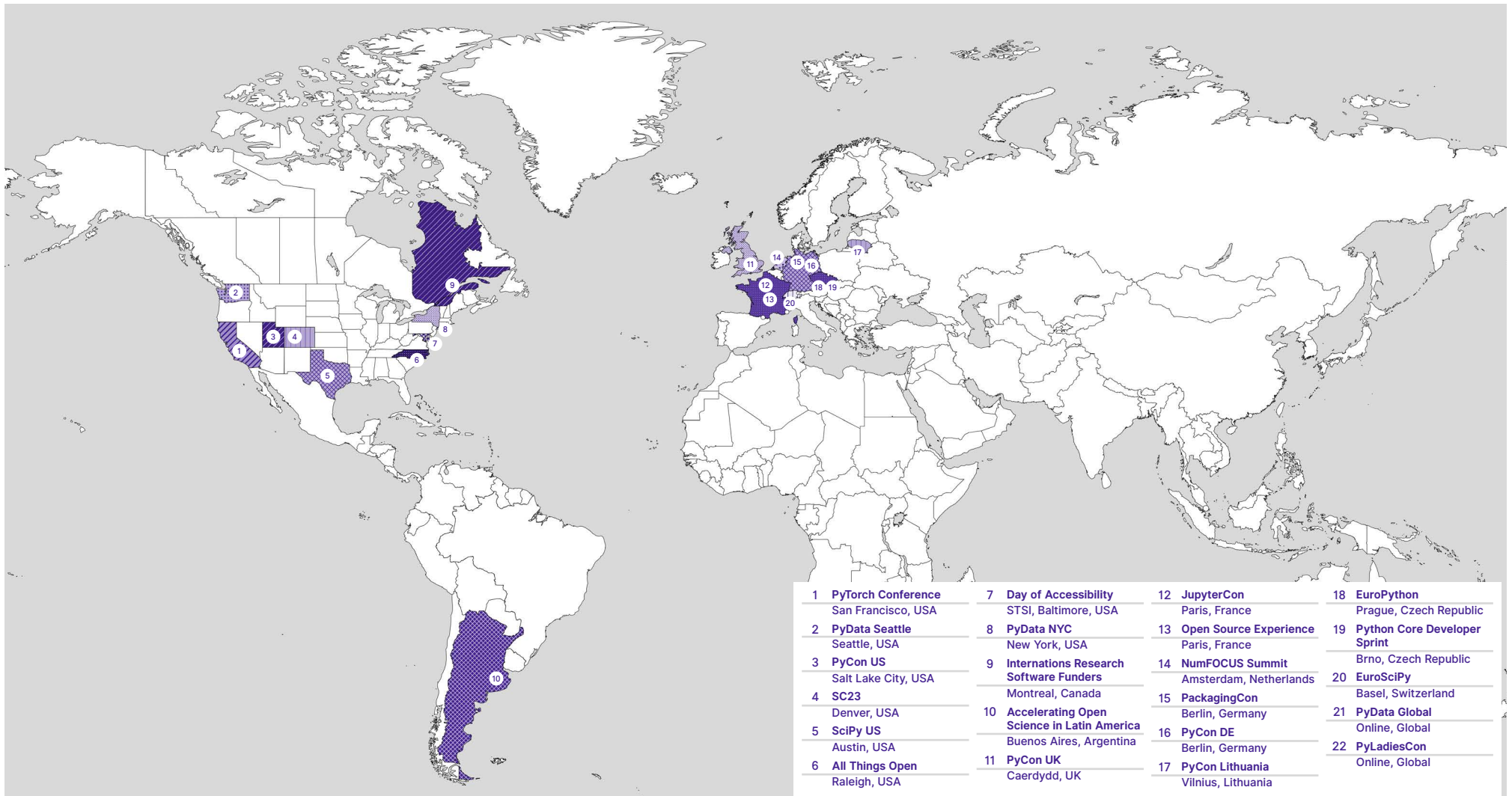
“As a designer at Quansight Labs, I’ve been thrilled to merge creativity with technology, especially in an open source environment that deeply values design. This unique focus not only enhances our projects but also vividly illustrates how design can transform the accessibility and impact of technology.”

Smera Goel

UI/UX Designer at Quansight Labs

Community Participation And Events

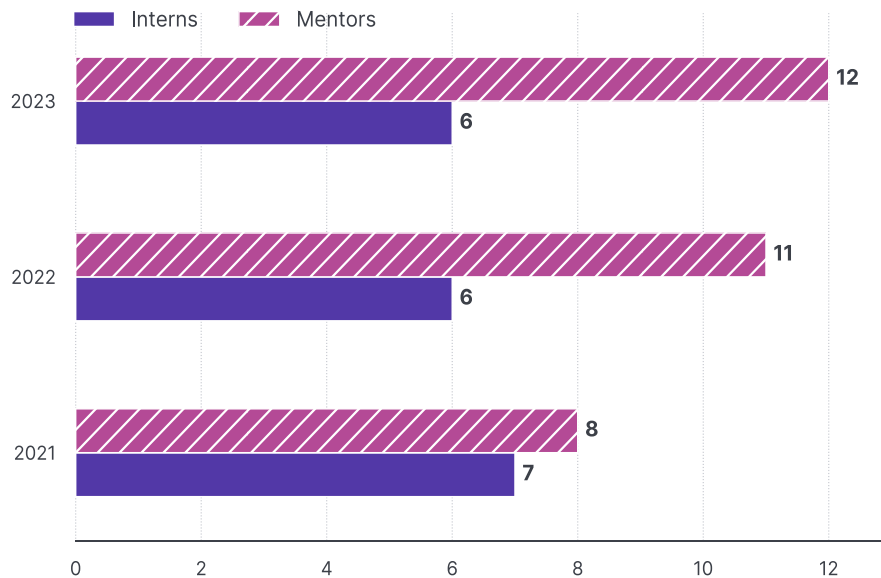
Labs members participated in 22 community conferences and events worldwide, including PyCon US, SciPy US, JupyterCon, PackagingCon, PyLadiesCon, and many more!



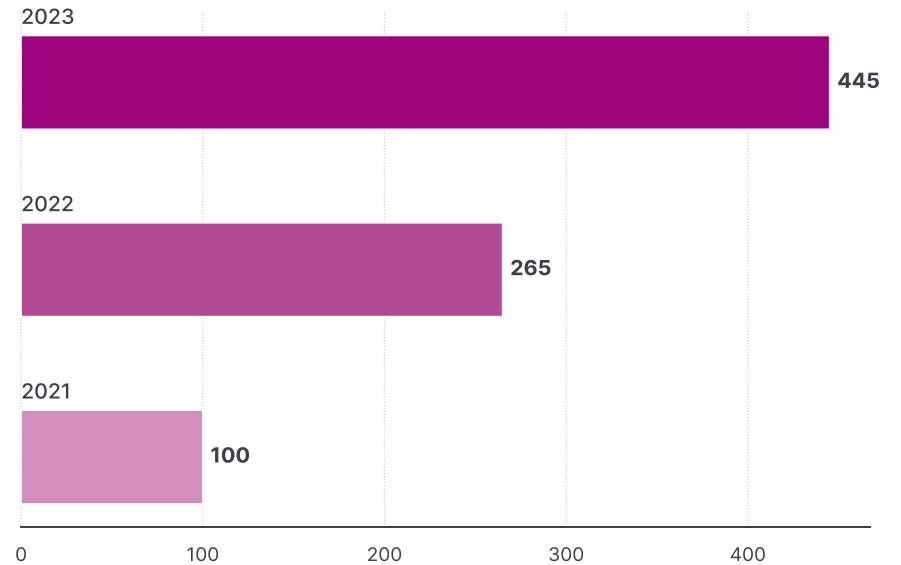
Internship Program

For the past three years, Quansight Labs has run an annual open source internship program to mentor and support new open source sustainers. 2023 was our most popular iteration so far, with ~450 total applicants, a 40% increase over the previous year. In July, we welcomed six interns and twelve mentors.

No. of interns and mentors



No. of applicants

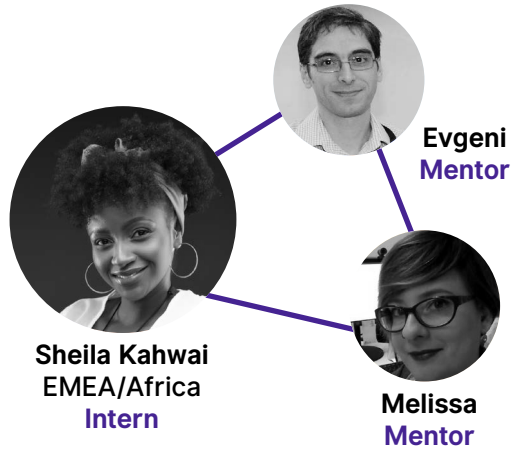


“My internship at Quansight gave me unique insight into the scientific Python community and allowed me to work on meaningful contributions to its projects. During these three months, during which I focused mainly on NumPy Python API, I received full mentoring and assistance that helped me scope my task to meet planned deadlines. The Quansight team creates an extraordinary environment where you can easily connect with and learn from the authors and core contributors of fundamental software, such as NumPy, SciPy, or IPython. I truly recommend applying to the next edition to anyone who wants to take the first steps in their Open Source journey!”

Mateusz Sokół

Quansight Labs 2023 intern turned Software Engineer at Quansight

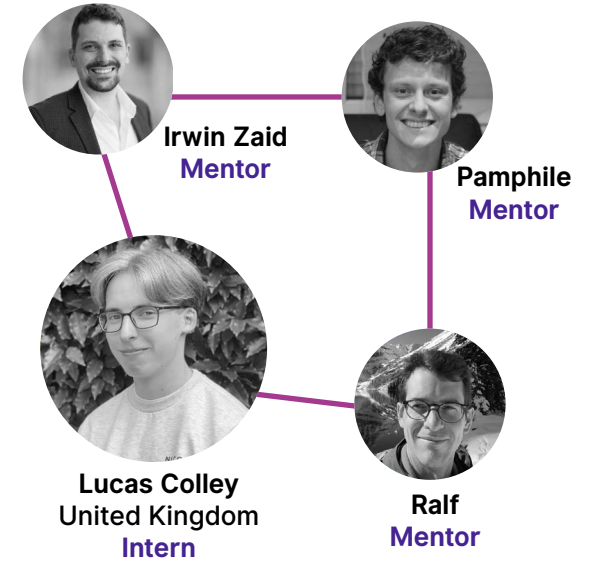
Interns, Mentors, And Projects



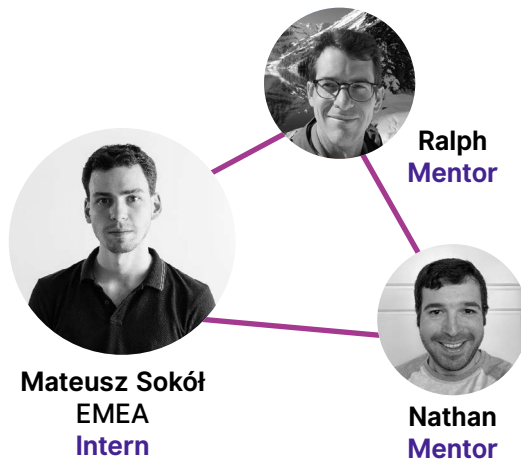
Project: [Enhanced doctesting for PyData libraries](#)



Project: [Integrate hypothesis testing into Sympy](#)



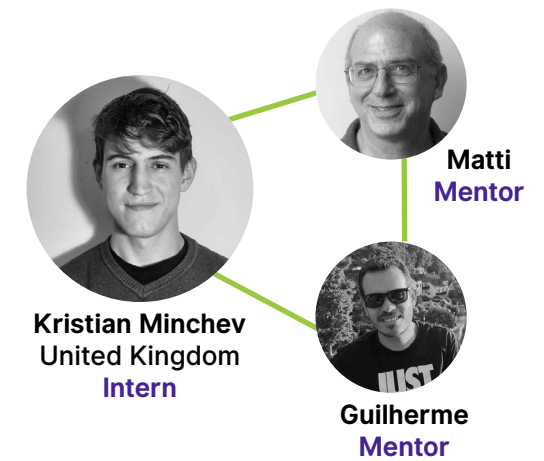
Project: [Array API standard, CuPy and PyTorch support for SciPy](#)



Project: [Python API Improvements for NumPy 2.0](#)



Project: [PyTorch-Ignite Code-generator DL/ML templates and integrations](#)



Project: [Improve Numpy support in Numba](#)

Healthy Mentorship Programs

Several large-scale open source internship and contribution programs like [Google Summer of Code](#), [Google Season of Docs](#), and [Outreachy](#) provide opportunities and exposure to hundreds of new open source contributors every year and serve as a springboard for many open source sustainers. During 2023, Labs members mentored three interns through these programs across conda-forge, Bokeh, and PyTorch Lightning.

At the same time, these programs' current scale and reach can be strenuous for open source communities (or projects) with volunteer sustainers, as the maintainers or mentors are often responsible for several logistics and organizational tasks on top of mentorship. As part of our mission of building sustainable open source communities, Labs started its open source internship program to onboard new contributors from historically underrepresented communities in the tech industry to open source in a friendly and safe environment.

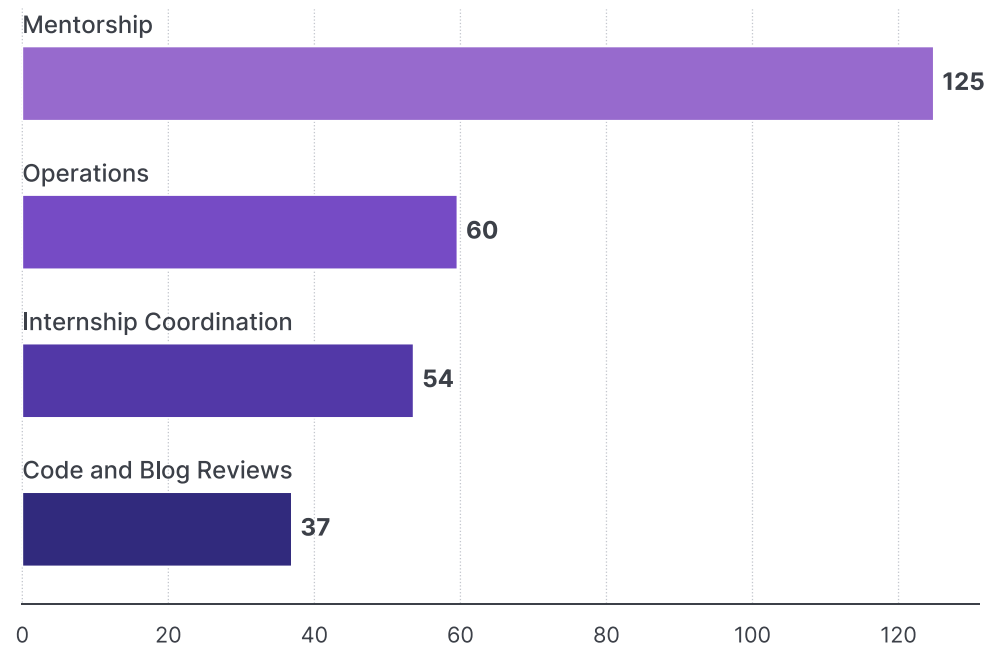
For instance, Labs members spent ~280 hours on the success of the Labs internship program in 2023. Labs' members spent less than half the time mentoring interns; Labs' members and Quansight HR and operational staff spent the rest of the time on essential organizational and administrative tasks such as reviewing internship applications, onboarding interns, supporting mentors, coordinating the program and regular check-ins, and preparing and delivering training sessions.

Our internship program focuses on establishing healthy collaborations with open source communities and building new opportunities for emerging open source leaders. We believe that a healthy and sustainable approach to mentorship in open source includes support for mentors, coordinators,

and administrators in addition to interns (through mentorship, feedback, compensation, etc.).

As such, all the aspects of the Labs' internship program are fully funded by Quansight, and members participating in external internship programs are also encouraged to spend their Labs-funded contributing hours as needed.

Hours per task - OSS Internship



Operations tasks include intern recruitment (application reviews and interviews), mentor-mentee pairing, HR tasks, and intern onboarding.

How Labs Internship Works

- ☑️ OSS sustainers at Labs propose well-scoped internship projects that align with their community roadmap. The projects can involve all aspects of open source, including but not limited to documentation, design, and community management.
- ☑️ Interns are selected through a resume review and interview and are evaluated for project-skill-fit and interest fit.
- ☑️ During the internship, interns are encouraged to join the OSS project community as supportive and collaborative members.
- ☑️ As part of the Quansight team, interns participate in company and team meetings, and have access to the Quansight spaces and resources, and some paid time off (PTO) days.
- ☑️ There are weekly cohort meetings as well as synchronous and asynchronous check-ins between interns and their mentors to ensure interns have sufficient support on their projects.
- ☑️ There are a couple of training sessions conducted for interns. Last year, we had sessions on OSS Contribution and a CV Clinic.
- ☑️ Towards the end of the internship, interns present their work at a Quansight All Hands meeting and write blog posts to share their projects and learning with the broader community.
- ☑️ Interns often also stay in touch after their internship, as Quansight alumni.

Collaboration with Christ Church College

Last year marked the initiation of our collaborative partnership with Christ Church College (University of Oxford). Partnering with academic institutions enables us to broaden the program's influence and share valuable knowledge with students about the significance and influence of OSS. Moreover, such collaborations provide avenues for exploring diverse funding opportunities, ensuring the sustained growth and success of the program.

This year, in partnership with Dr. Irwin Zaid, Lecturer in Computer Science at Christ Church College (University of Oxford), we piloted our first collaborative internship with a higher education institution. Thanks to that partnership, we welcomed two students from Christ Church College: Lucas Colley and Kristian Minchev. You can read about [Lucas' experience in this article](#). Their hard work, excellent contributions, and Dr. Irwin's continued support made this partnership a resounding success.

“Open source software is essential to our digital infrastructure, playing a part in everything from advancing discovery in modern science to providing the operating system that runs almost all mobile devices. And with the rise of data science and machine learning, this is only becoming truer, especially so for the fundamental tools of the PyData ecosystem. Software libraries like NumPy and SciPy, which were already widely popular, now find themselves powering innovation throughout science and technology. This means they have a great need for further development and support, which requires more people to contribute to make it happen.

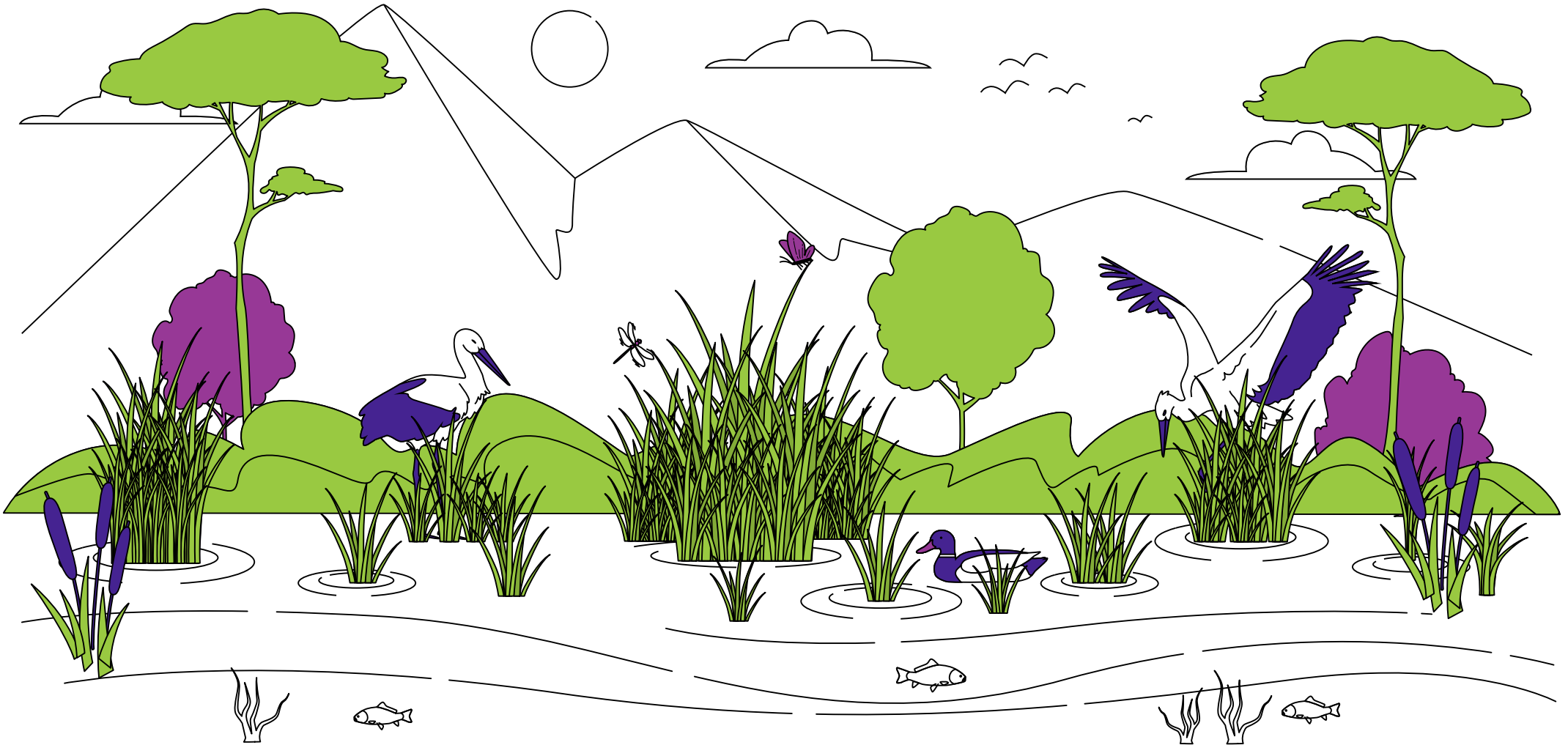
Day to day, I teach computer science at Christ Church, one of the constituent colleges of the University of Oxford. There, we are fortunate to have a truly diverse community of amazing undergraduate students, many of whom would like nothing more than to spend their summer working on a challenging project for a worthy cause. Last year, we helped two of our computer science students do just that, by supporting them with summer bursaries so they could contribute to open source projects.

They were guided by a mentor from Quansight Labs and myself. It was truly a great success, both in terms of the code written and the learning offered to the students themselves. I'm excited to do so again starting this summer, and very keen to make our focus be impactful contributions to the fundamentals of the PyData ecosystem.”

Dr Irwin Zaid

Lecturer in Computer Science, Christ Church, University of Oxford

Labs Work Culture



Community-First Working Model

Our goal at Labs is to facilitate the long-term sustainability of our ecosystem. As a collective, Labs members have vast experience as users, contributors, sustainers, and leaders within critical open source projects and their communities — sometimes spanning a couple of decades.

The following are the core principles guiding how we think about and approach open source sustainability and collaborations and engagement with the open source community.

The OSS community is the primary stakeholder

We collaborate with other community members and leaders when working on community-driven projects. That means we favor decision-making processes such as consent or consensus-seeking instead of making unilateral decisions. Additionally, as individuals employed to work on community-driven open source projects, we must be aware of work/load distribution and ensure we are not overloading volunteers in the project. Instead, our work should be making their life easier.

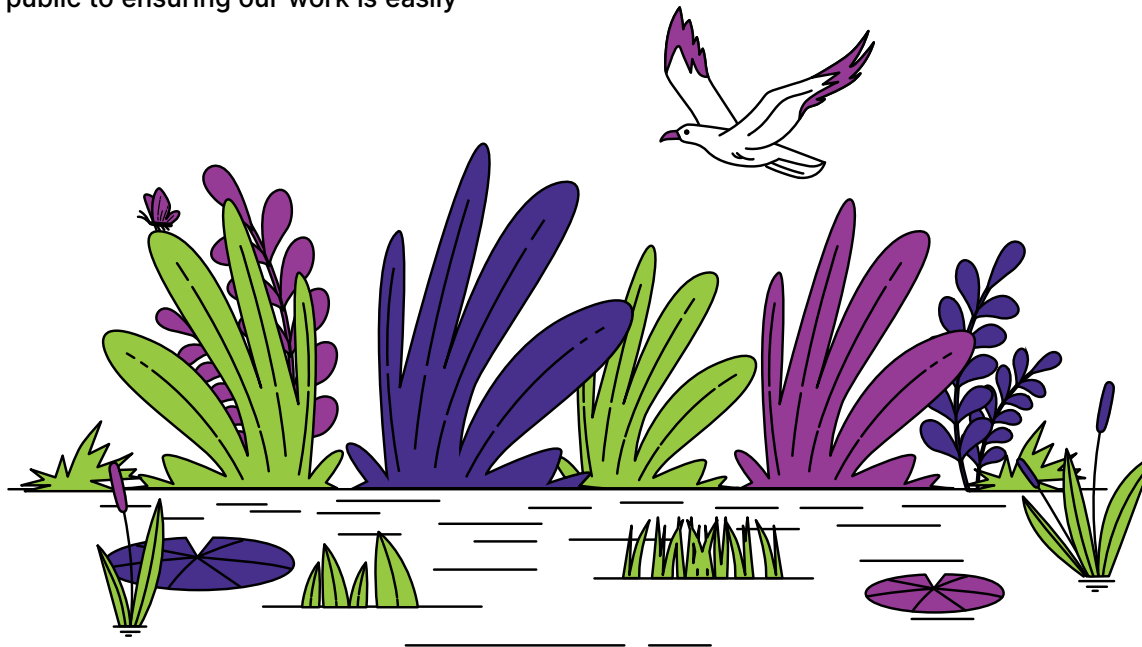
Some ways in which we adopt these principles are:

- Open issues and shepherd discussions to solicit feedback and input on any significant efforts or pieces of work.
- Actively seek code reviews or alternative feedback mechanisms from contributors outside Quansight Labs.
- Follow a project's decision-making processes and enhancement proposal processes.
- Prioritizing work and approaches that help build a healthy, inclusive, and equitable culture, such as seeking collaborations with community members outside Quansight Labs.
- Dedicate substantial time and effort to day-to-day maintenance tasks such as triaging, reviewing contributions, welcoming new contributors, fixing bugs, and updating documentation.
- Share attribution and acknowledge others' efforts and contributions.
- Encourage community-driven projects to develop and maintain roadmaps and support them in this effort.

Funding for open source projects

Funding opportunities should align clearly with the project and its community's needs and goals. To ensure all our funding-seeking and execution tasks align with this principle, we have adopted the following guidelines:

- Engage with the appropriate leadership groups before seeking funding. Such an approach ensures alignment on the scope, deliverables, and ownership. Also, it allows for the inclusion of other community members — outside of Labs — to collaborate on funding proposals and execution of tasks.
- Treat collaborators as equal partners.
- Our default way of working is transparent and open; this helps build trust with other interested parties and makes it easier to hold ourselves accountable. We meet this principle in many ways, from making submitted grant proposals public to ensuring our work is easily accessible by the community by leveraging community spaces, open tools and processes for project management, and regularly presenting updates to community stakeholders.
- Routinely assess our progress and allow for and invite feedback from community members.
- Prioritize knowledge sharing and long-term maintainability.
- Prioritize resource sharing with individuals and organizations serving historically marginalized or disenfranchised groups.



Co-Ownership Governance Model

Quansight Labs follows a [Feminist co-leadership](#) model of governance founded on principles of shared ownership, inclusiveness, collaboration, participation, and empowerment. Such a model can help decentralize power dynamics further and support collective leadership at all organizational levels. This has allowed us to apply this shared leadership model across multiple levels within Labs, including Labs' strategic areas. Some of the benefits we've experienced through adopting models of co-leadership for strategic areas are:

- ☑ Robust, decentralized, and effective decision-making.
- ☑ Diversification of ideas and perspectives at strategic levels.
- ☑ Increased transparency and enhanced delivery.
- ☑ Continuous assessment of and iteration of strategic priorities and goals.

This year our strategic areas of Packaging and Multiple Array operated in a co-ownership model, and we plan to extend this to other strategic areas over the next couple of years.

2023 Focus Areas

Two of the biggest strengths of Labs are its wide range of expertise across many areas within the scientific computing open source ecosystem and our deep roots and involvement within its community. As such, we are uniquely positioned to identify and work on ecosystem-wide pressing challenges and opportunities, thus maximizing our organizational impact. We reflect on and set strategic goals each year around some focus areas that align with those key community needs or challenges and mission and vision.

In the following sections, we present some highlights of three of our strategic areas for 2023: Python packaging, access-centered practices, and multiple tensor / array library support.

Python Packaging

Packaging is a topic that is relevant to every package author and end user. Despite significant improvements across the Python packaging landscape over the past years, it will probably be no surprise to readers that we chose this as a focus area. Our efforts this year can be categorized into three related but distinct areas with their own tools, standards, and communities: the PyPI ecosystem, the Conda ecosystem, and build tools. The first two center around package managers and repositories, and the third around Python projects and their (needs for) build systems.

Contributions focused on the PyPI ecosystem included:

- ✓ Active participation from several Labs team members in the big-picture discussions on the Python Packaging Discourse around a common vision and governance for Python packaging at the start of the year.
- ✓ Improvements around the handling of external (non-Python) dependencies for Python packages (e.g., [PEP 725](#)).
- ✓ Improved BLAS/LAPACK handling (e.g., OpenBLAS distribution guidelines and [wheels](#), and [a new set of CI jobs for NumPy](#) for BLAS configs).
- ✓ [Pypackaging-native](#) updates (e.g., more accurate content on GPU support).

Contributions focused on the conda ecosystem included:

- ✓ Updated Fortran support via Flang in conda-forge (which also enabled Python 3.12 support for SciPy, as described in [this compelling blog post](#)).

- ✓ Providing [a GPU-enabled server](#) including appropriate access controls for integration into the conda-forge CI infrastructure.
- ✓ Improving [conda-libmamba-solver](#), now the default solver for Conda.
- ✓ [Menuinst v2](#) - cross-platform menu item installation, relevant to application distribution.
- ✓ Collaborating on the [new conda-forge website](#).
- ✓ The large amount of work around environment management with [conda-store](#) (see ["stories: Nebari and conda-store"](#)).

Contributions focused on build systems and tooling included:

- ✓ A large amount of [meson-python](#) features, releases and maintenance work (with NumPy, pandas, and Matplotlib prominent new adopters).
- ✓ Helping many scientific Python projects with migrating away from numpy.distutils before the Python 3.12.0 release.
- ✓ Work towards a new [sysconfig API](#) maintenance & new features.
- ✓ Better cross-compilation support in Python packaging (e.g., [PEP 720](#), [PEP 739](#)).

Collectively we spent over **4,000 hours** dedicated to these build & packaging topics. About 60% of that was work funded by grants from NASA and the Chan Zuckerberg Institute and contracts from Anaconda, the Sovereign Tech Fund, and Intel. We did not have external funding for the remaining 40%.

Access-Centered Practices

As an organization with a core mission of building sustainable and inclusive communities and projects, we are committed to approach our work through a lens of diversity, equity, and inclusion. We truly believe that ensuring digital tools and content are accessible to everyone, including people with disabilities, is everyone's responsibility and an imperative within open source. That is why developing access-centered practices for and with the open source scientific community has been one of our strategic areas for various years.

Achieving disability inclusion requires a dedicated effort and ongoing commitment to:

- Consider accessibility and access-centered practices at every stage of the design and development process.
- Continuously assess and improve accessibility.
- Adhere to accessibility guidelines and standards.
- Including people with disabilities.

Some examples of access-centered work carried out in the past year are:

- JupyterLab and Jupyter notebook accessibility audits, remediation and enhancements (see [Stories: Jupyter accessibility](#)).
- Significant accessibility improvements to the [PyData Sphinx Theme](#) and the [Scientific Python Hugo web theme](#).
- Knowledge sharing through workshops, conferences and content focused on accessibility, accessible content generation, accessibility testing, and disability inclusion in open source.
- Continue to lead the Jupyter Accessibility fortnightly community calls.
- Supporting and providing access-centred advice to other projects and maintainers in the scientific community.
- Labs funded design and accessibility contributions to projects like napari, conda-store, Jupyter notebook, JupyterLab, nebari, conda-forge, and more (which accounted for around 250 hours this year).

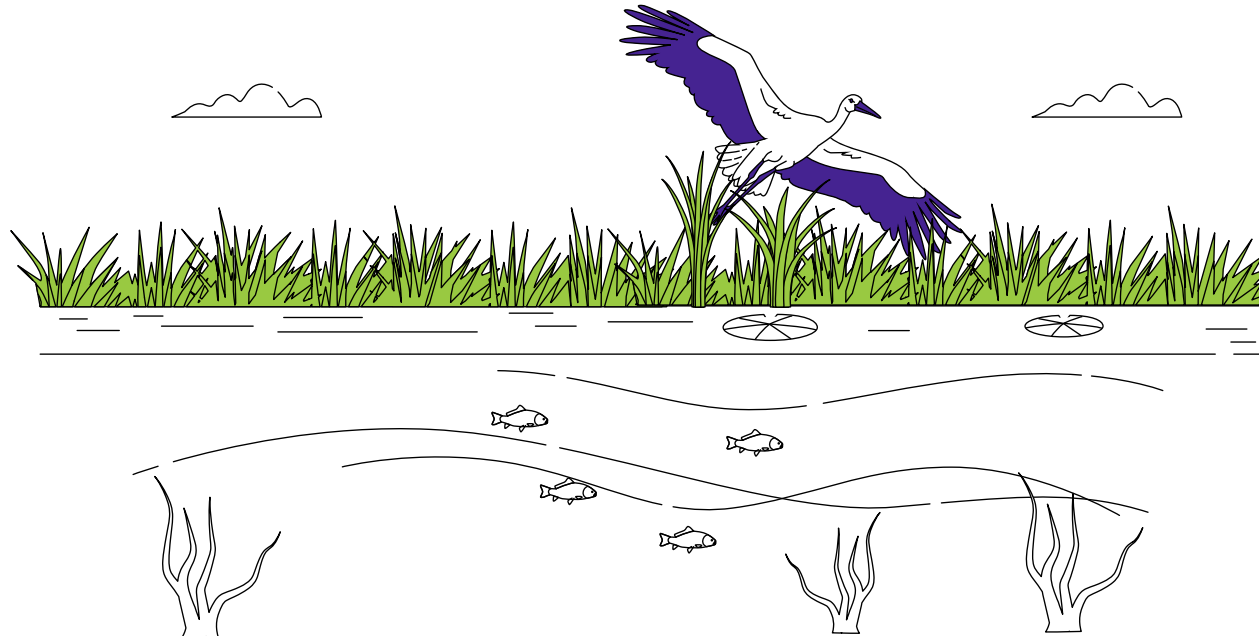
Collectively we spent around 8000 hours this year working on access-centred initiatives with about three-quarters of that time funded through CZI EOSS grants.

Multiple Array / Tensor Library Support

Standardization efforts are often long, multi-year efforts, and the Array and Dataframe API Standards are no exception. Thankfully, we at Labs are fortunate to have industry partners, library maintainers, and other stakeholders who believe in our vision and who are committed to helping reduce fragmentation and further the goal of array and dataframe interoperability across the scientific Python ecosystem. As a result, we believe the future is bright and are excited for the year ahead, where Quansight Labs will continue leading in coordinating standardization activities and accelerating adoption among downstream libraries throughout the PyData ecosystem and beyond.

Our activities around multiple array/tensor library support topics are described more deeply in [“Stories: Consortium for Python Data API Standards”](#).

Collectively, we spent over 4,200 hours on this focus area. Almost half of that time was funded by contracts with Meta and Intel and a grant from NASA. We did not have external funding for the other half.

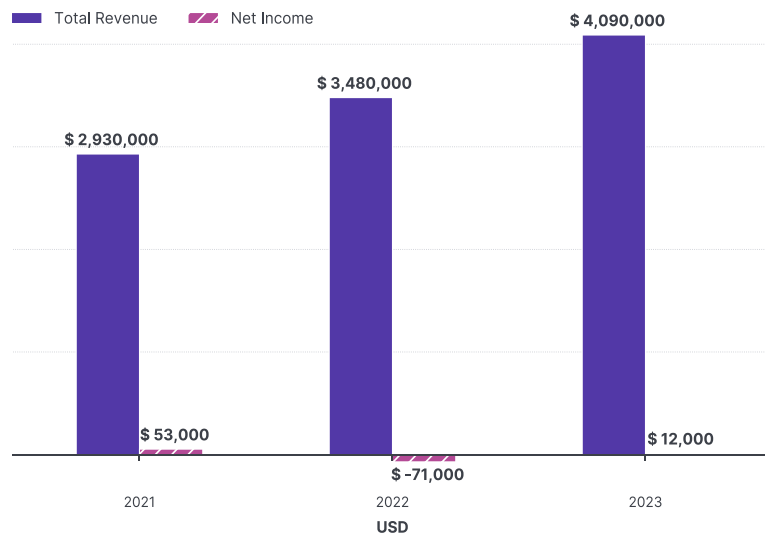


Financials

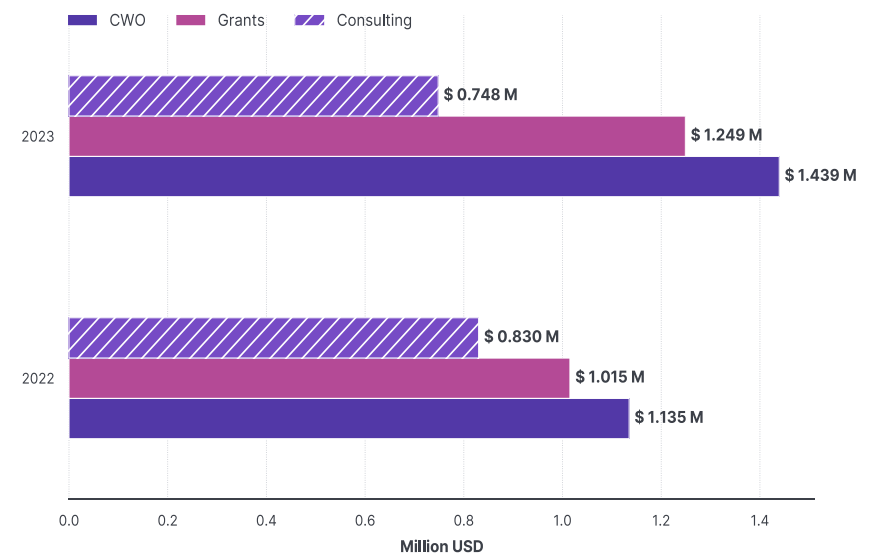
We run Quansight Labs as a public benefit organization — we don't intend to make a profit, but rather work at cost on community-driven projects and on topics we think the community will benefit from.

We aim to support our work through a diverse set of income streams. In 2023, our revenue grew slightly to just a little over \$4,000,000, yielding a Net income of \$12,000. This income came from three types of sources: [community work orders](#) (CWO), grants, and consulting activities.

Our revenue and income



Funding sources



This year, our CWO funding grew as we received support from the Sovereign Tech Fund [to work on OpenBLAS](#) through their [General Investment program](#) and [con-da-store through their Contribute Back Challenges program](#).

Most of our grant funding came from the Chan Zuckerberg Institute, through its [Essential Open Source for Science](#) program and [a grant for Scientific Python](#), and NASA, through the software component of its [Research Opportunities in Space and Earth Sciences \(ROSES\)](#) program.

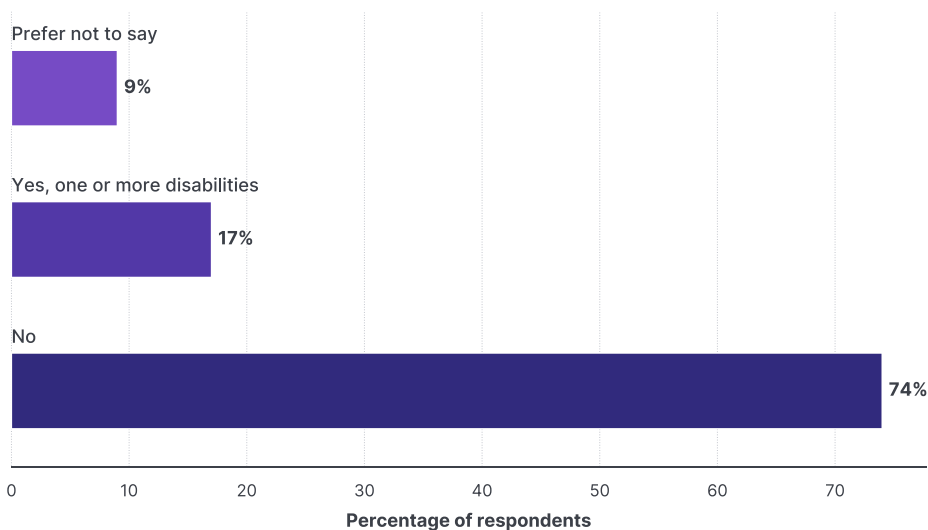
Consulting income comes from projects done with a variety of industry clients. This income is used to support key unfunded work on community projects. The consulting activities themselves are an opportunity for open source maintainers in Labs to gain experience with how their projects are used in the industry.

Diversity Within Labs

Quansight Labs is committed to building a diverse and inclusive organization that reflects our world. This principle guides how we do things at Quansight Labs, from our internal policies to community initiatives we support or spearhead. Our open source internship allows us to support individuals from historically underrepresented backgrounds in tech and provide them with the tools, training, mentorship, and support they need to join and thrive in open source.

Like many other organizations, Labs had limited growth this past year (in terms of hiring), nevertheless, we remained committed to building a diverse and inclusive remote-first culture where every member experiences a sense of belonging and growth.

Disability status

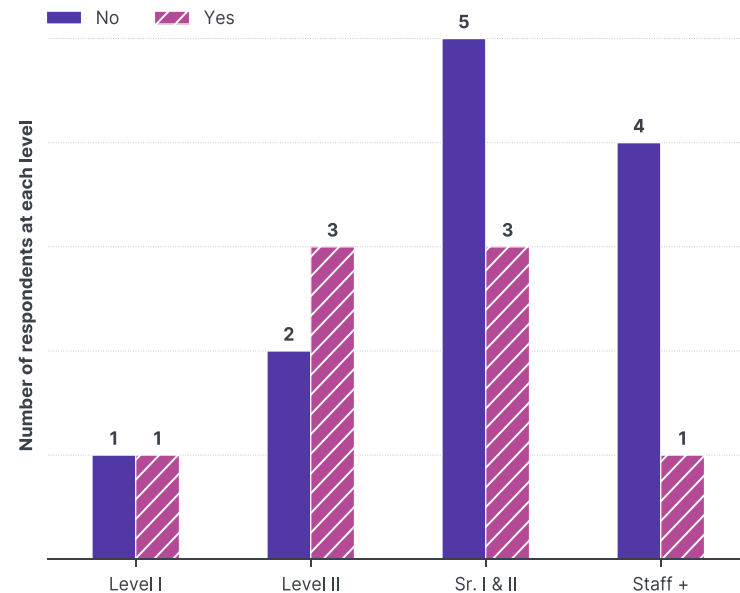


This data reflects the state of Quansight Labs as of January 19th, 2023. The data was collected through an internal voluntary survey with a participation rate of 92%. Plots are presented in terms of the percentage or total number of respondents to each question on the survey.

We saw a steady representation of disabled individuals within Labs' members in 2023.

We also observed a relatively steady representation of Labs' employees belonging to marginalized groups across all role levels.

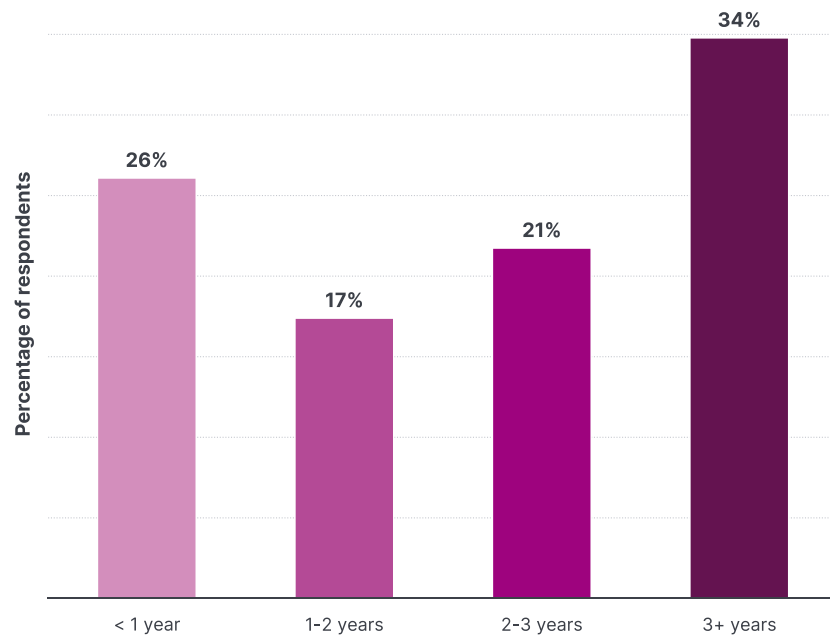
Self-identification as belonging to a marginalized group by level



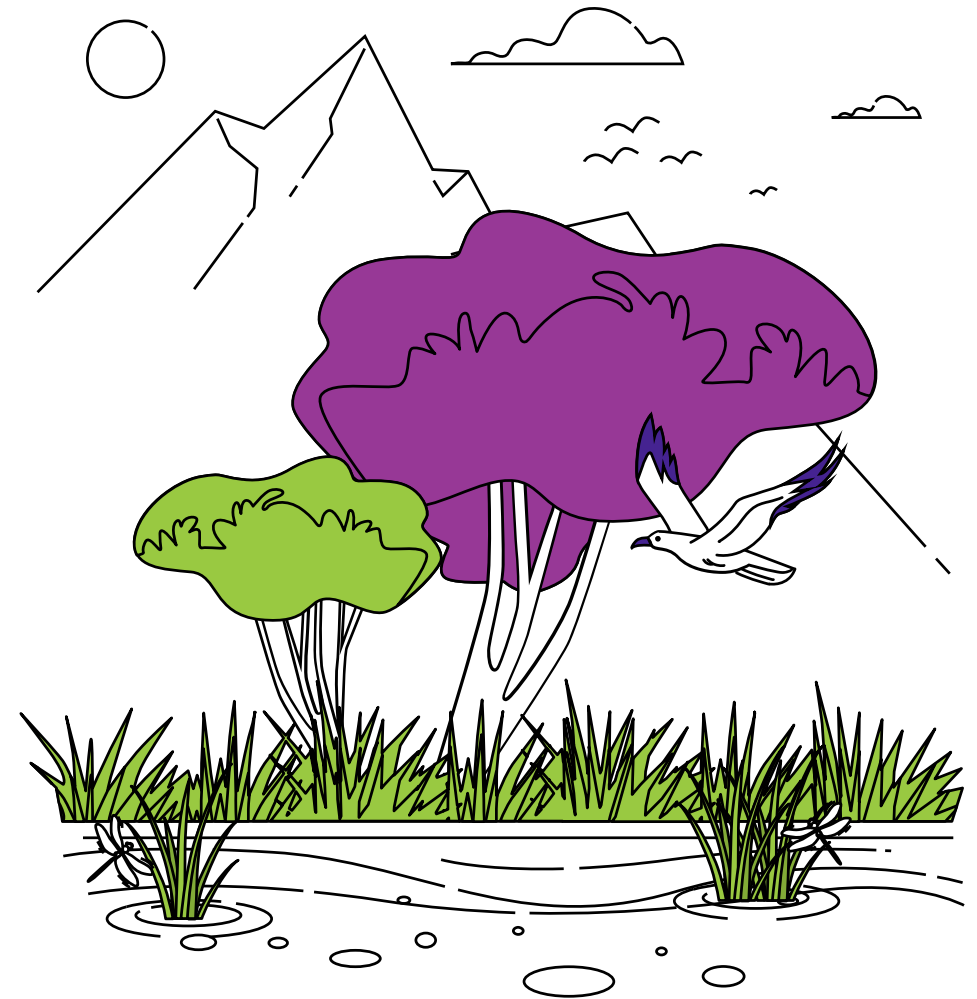
This data reflects the state of Quansight Labs as of January 19th, 2023. The data was collected through an internal voluntary survey with a participation rate of 92%. Plots are presented in terms of the percentage or total number of respondents to each question on the survey.

55% of our employees have been part of Quansight Labs for 2 or more years.

Time working at Labs



This data reflects the state of Quansight Labs as of January 19th, 2023. The data was collected through an internal voluntary survey with a participation rate of 92%. Plots are presented in terms of the percentage or total number of respondents to each question on the survey.



Building a Sense of Belonging

As a remote-first organization, we find it very valuable to get together at least once yearly. It allows us to build relationships, align on our strategic goals, and find opportunities for collaboration and team building.

In June, we gathered for “Camp Quansight Island Edition”, our annual company retreat, held at Monte Isola, Italy, this year.

Labs members work really hard for the open source community, and as a small token of recognition, we start our quarterly all-hands meetings with the PyPie Raffle — we spin our wheel of names, and the winner gets a pie, or a savory treat, or a small present!



How Far We've Come And Looking Ahead

We are thrilled to share some of the highlights and learnings from 2023 in this annual report. Quansight Labs' projects and initiatives accomplishments reflect our organizational strengths and our commitment to the open source ecosystem. In 2023, we continued to demonstrate steady growth and progress toward our goals of building sustainable and inclusive open source projects and communities. We are also incredibly proud of the team we have built at Labs — an extraordinary group of talented individuals who bring diverse perspectives to our organization and communities — and all their achievements.

Every year, we take some time to reflect and ask ourselves: Are we making a difference/challenging the status quo? What can we do better or differently? We can undoubtedly say that we did make a difference in 2023. This coming year — 2024 — will continue to present challenges at the global and industry levels. Still, it will also bring new opportunities to continue working toward our mission based on community-led and equity practices within our organization and in our communities and projects.

We will continue to focus on strategic impact areas that reflect our community's most significant challenges and needs. We will also continue to support leaders' emergence within the organization. Based on what we have learned this year, we will formalize a strategic Documentation area following a co-leadership model. At the same time, we will sustain and ideally extend our focus on other already established strategic areas — access-centered practices, multiple array/tensor library support, and packaging — in ways that align with the community's needs and direction. There will also be a place for innovation and exploration.

On the economic side, we will continue exploring funding models and identifying diverse funding sources to ensure our organizational sustainability. We expect modest growth. Our focus and commitment to serve our community through transparent, equitable, open, and collaborative initiatives and projects will remain strong.

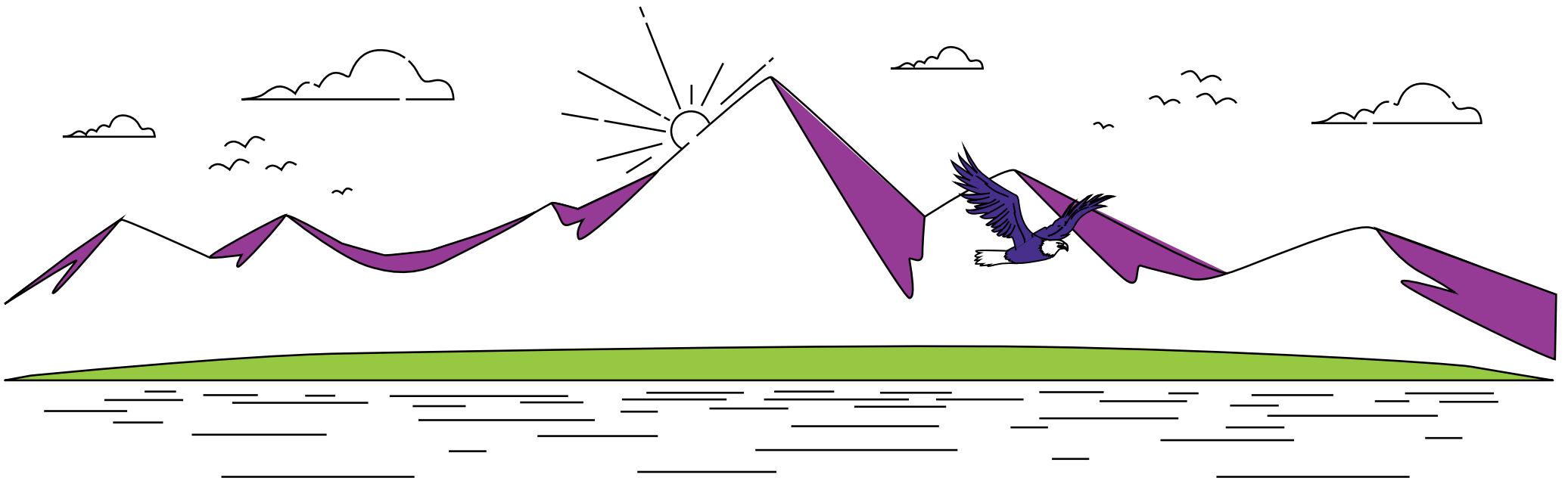
Here is to a successful coming year.

Acknowledgements

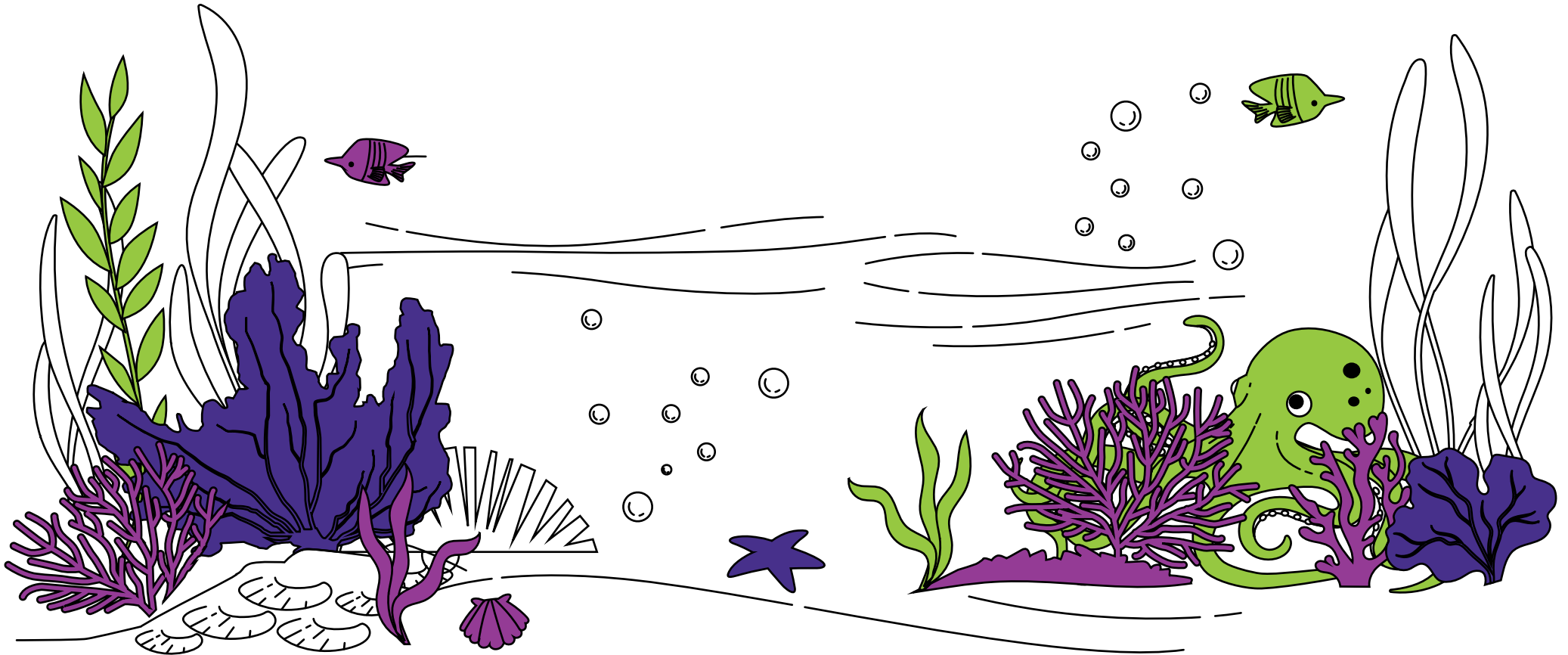
This year's report is presented by Tania Allard, Pavithra Eswaramoorthy, and Ralf Gommers, with contributions from Athan Reines, Melissa Weber Mendonça, Jaime Rodríguez-Guerra, and Gabriel Fouason.

All the beautiful graphics and the report were designed by Irina Fumarel and Mariana Di Giovanni.

Special thanks to Kate Brack for project management, Pierre Simonard for data collection, and Ashley Baal for facilitating information sharing within the team.



See you next year!



Visit labs.quansight.org to learn more about us, and reach out to us at connect@quansight.com

