

Advancing an inclusive culture in the scientific Python ecosystem

Essential Open Source Software for Science Diversity and Inclusion Grants

Proposal Details

Grant lead Melissa Mendonça

Open Source Projects NumPy, SciPy, Matplotlib, Pandas

Proposal Purpose

To support the onboarding, inclusion, and retention of people from historically marginalized groups on scientific Python projects and structurally improve the community dynamics of NumPy, SciPy, Matplotlib, and Pandas.

Proposal Summary/Scope of Work

The lack of diverse representation among contributors and maintainers of open-source projects is a well-known problem. For projects in the core of the scientific Python ecosystem, this is arguably even worse. While hiring maintainers from diverse backgrounds would address this problem temporarily, it's not a long-term solution. It also takes real effort to diversify contribution paths inside these projects and build spaces where diverse newcomers can thrive.

This joint proposal from NumPy, SciPy, Matplotlib, and Pandas focuses on creating a DEI action plan and dedicated "Contributor Experience Lead" positions to identify, document, and implement practices to foster inclusive open-source communities.

We'll create two cross-project positions dedicated to contributor experience and a project lead position (Melissa Mendonça, NumPy) to mentor and man-

age them. Introducing "Contributor Experience Lead" positions will allow us to:

- take a focused approach to engaging, onboarding, and supporting new contributors, especially those from historically underrepresented groups,
- promote retention and leadership progression for current members of our projects.

The scope of these new positions is a mix of technical and community work, with a focus on DEI. Having technical responsibilities will help our hires to further expand their technical knowledge and experience and provide them with better employment prospects in the future. These roles would be best served by people from historically underrepresented groups. However, the responsibility for DEI work mustn't fall solely on their shoulders. Additional mentorship and guidance will be provided by Ralf Gommers (NumPy, SciPy), Hannah Aizenman and Thomas Caswell (Matplotlib), Matt Haberland (SciPy), and Joris Van den Bossche (Pandas).

Key responsibilities for each Contributor Experience Lead role will include:

- Improving the experience of contributors and maintainers of two projects.
- Building a DEI-focused cross-project team.
- Improving cross-project integration.

Technical activities for this role will include onboarding new contributors, monitoring first-time contributors' PRs and ensuring they receive feedback, acting as maintainer for the projects, and fixing common developer pain points.

Organizational activities will include organizing regular community meetings, sprints and mentoring activities (possibly including multiple projects), and maintaining a public calendar of events.

Other activities will include aligning documentation across projects, including themes and layouts where possible, coordinating the internationalization of websites and improving translation workflow.

Proposed structure:

- PI, 8 hrs/wk. Responsible for mentoring and managing new hires, grant management, and aligning with contact(s) for each project. Coordinate the execution of a cross-project DEI action plan and document challenges and strategies to be shared with others.
- Contributor Experience Lead 1 (full-time, 40 hrs/wk)
- Contributor Experience Lead 2 (part-time, 20 hrs/wk)

At least one of the Contributor Experience Lead positions should be focused on a senior/mentor profile, increasing the chances they are more independent and more likely to succeed. Inessa Pawson, an experienced member of the NumPy community, is an ideal candidate for this position. To better understand actions that will promote inclusion and have a measurable impact on increasing diversity in our projects, we'll develop a DEI action plan in collaboration with an outside consultant familiar with open source in the first six months of this work.

Work Plan

To expand on the proposal summary, the work proposed by this group will be executed by two Contributor Experience Leads, the newly created cross-project positions, and a Project Lead who will be mentoring and managing them. The scope of these new positions will be a mix of technical and community work, focusing on actions that promote diversity, equity, and inclusion. Having technical responsibilities will help our hires to expand their technical knowledge and experience further and provide them with better career prospects in the future. Additionally, it will ensure new contributors have a dedicated point of contact regarding technical issues.

These roles would be best served by individuals from historically underrepresented groups. However, the responsibility for DEI work mustn't fall solely on their shoulders. Experienced maintainers of these projects, namely Ralf Gommers (NumPy, SciPy), Hannah Aizenman and Thomas Caswell (Matplotlib), Matt Haberland (SciPy), and Joris Van den Bossche (Pandas) will provide additional mentorship.

The Contributor Experience Lead will be a role that's 50% technical and 50% organizational. Key responsibilities for each Contributor Experience Lead will include:

- Improving the experience of contributors and maintainers of two projects.
- Building a cross-project team focused on improving diversity, equity, and inclusion of contributors from historically underrepresented groups in open source.
- Improving cross-project integration and collaboration.

Technical activities for this role will include onboarding new contributors, monitoring their PRs to ensure they receive appropriate feedback and feel encouraged to become regular contributors, and fixing common developer pain points. We expect up to 25% of the Contributor Experience Lead time will be spent on maintenance and small features development on one or two projects. The goals of this activity are (a) to build up an understanding of

project maintenance, (b) to develop new technical skills, and (c) to enhance the quality or sustainability of the project. The candidates can choose the topic(s) according to their interests.

Organizational activities will comprise hosting regular community meetings, sprints, and various mentoring activities, including "Newcomers' calls" (on-line meetings dedicated to onboarding new contributors), and maintaining a public calendar of events. Some of these activities will be executed in coordination with groups focused on the inclusion of historically underrepresented groups in open source. The key personnel of the proposal has already hosted and collaborated on a number of such events with Data Umbrella, PyLadies, and WiMLDS.

Other activities to be performed by these candidates will include:

- improving contributor onboarding guides for each project and implementing governance practices that promote participation, empowerment, retention, and leadership progression of contributors from groups that are underrepresented in scientific open source,
- aligning documentation across projects, including themes and layouts where possible,
- coordinating the localization of the websites and improving the translations workflow.

Since these activities include coordinating four open-source projects, we are proposing two roles, one full-time (40 hrs/wk) and one part-time (20 hrs/wk). At least one of the Contributor Experience Lead positions should be focused on a senior/mentor profile, increasing the chances they are more independent and more likely to succeed. Inessa Pawson, an experienced member of the NumPy community, is an ideal candidate for this position.

Key responsibilities for the Project Lead role (8hrs/wk) will include:

1. Manage execution of the grant, including recruiting, managing, and mentoring the two Contributor Experience Leads.
2. In the first six months, develop a DEI action plan in close collaboration with the leadership of the projects included in this application and an outside DEI consultant familiar with open source. After that, for the rest of the grant period, coordinate the execution of the plan and document challenges and successful strategies to overcome them.
3. Mentor new maintainers from systemically underrepresented groups in scientific open source.

We propose Melissa Mendonça for the position of Project Lead.

As an example of activities that in recent months have advanced the participation, retention, and leadership progression of contributors that are under-represented in scientific open source, we'd like to highlight the work being done in the NumPy project. The first Cycle of the CZI EOSS grant has supported the creation and maintenance of the Documentation Team for NumPy. The team continues to facilitate onboarding of new NumPy contributors and maintainers as well as hosts, coordinates, and participates in a number of outreach activities (e.g., Google Season of Docs, mentored sprints, Newcomers' calls, etc.). This grant has also created a leadership opportunity for Melissa Mendonça, who is now a NumPy maintainer and member of its Steering Committee.

Milestones and Deliverables

We expect the work supported by this grant to generate increased participation and community engagement within our projects, with a specific focus on increasing gender, race and geographical location diversity of our contributors. Given our projects' unique position in the core of the scientific Python ecosystem, we believe our actions towards diversity, equity, and inclusion will model practices and behavior that bring awareness and concrete improvements to a great number of open-source communities.

For the first two months of the proposed work, the main responsibility of the Project Lead will be to onboard the personnel in the Contributor Experience positions.

The deliverables for the Contributor Experience Leads will be:

- Improve the contributor onboarding guides for each of the four projects. (first 3 months)
- Organize calendars, set up community meetings for each project, and sync best practices for community engagement across projects. (first 3 months)
- Start regular mentoring sessions with newcomers and document specific needs or common pain points for new contributors of these projects. (first 6 months)
- Become comfortable maintaining two projects (according to interest).

By the end of the first year, we expect that the Project Lead and Contributor Experience Leads together will:

- Identify appropriate metrics and responsible ways of gathering information about the projects' contributors to study the demographic make up and plan for improvements.

- Develop and make public a DEI action plan in collaboration with an external consultant familiar with open source.
- Organize sprints and participate in other outreach activities such as talks, live coding events, or events targeting historically underrepresented groups in open source.
- Produce a report for the activities executed in the first year, to be shared publicly.

While continuing with the set of activities set up during the first year, at the end of the second year, we expect to see an increase in community engagement, measured by 1) the number of individuals attending community meetings and 2) an increase in the number of contributions per contributor in each project (code and otherwise).

Specific deliverables will be:

- An increase in contributors, maintainers, and leaders from underrepresented groups in all four projects.
- Establishing and documenting a pathway for new contributors to progress to a maintainer role or join a subcommittee on each project.
- Providing mentorship within each project dedicated to a specific area of a project's maintenance (e.g., documentation, website support, issue triaging, translations, fundraising, etc.)
- A public report detailing the work done during the grant period, highlighting the impact of the set of actions developed as part of the project.

Expected Outcomes

We aim to produce lasting changes in the processes and culture of our projects. By creating and implementing a DEI action plan and introducing paid roles dedicated to improving the contributor experience, we anticipate the following outcomes:

- a more diverse pool of contributors on each project,
- greater engagement with contributors from historically underrepresented groups, promotion of active open-source contributors from historically underrepresented groups to leadership roles.

The DEI action plan will help guide decisions in all our proposed activities and beyond. It will include applying practices that have proven to be effective in one project to other projects. Concrete examples are the community meetings and newcomer-friendly Slack workspace and office hours from NumPy,

the way the Matplotlib Twitter account is used for outreach, and the PyData mentored sprints format.

By dedicating time and resources to onboarding and mentoring, we will create a better environment for newcomers, especially from underrepresented groups. Also, we will be able to diversify contribution paths inside each project which in turn will allow greater participation of contributors with diverse backgrounds and expertise.

The contributors in paid positions will gain valuable work experience that will allow them to further their careers in tech. This position will also give them visibility and credibility in the open-source community and beyond and present them as role models for the larger community, specifically potential and new contributors from historically underrepresented groups.

Evaluation and Learning Goals

The scope of work for this proposal includes activities that should structurally improve the community dynamics of our projects. We expect to develop a clearer picture of what works and what doesn't in our projects for engagement and retention of new contributors, especially from historically underrepresented groups. Furthermore, establishing cross-project roles will introduce a new collaboration model to the Scientific Python communities, allowing community-building work within the ecosystem to be done more efficiently and with greater outcomes. In all, the work performed during the grant period will have a lasting impact on our communities and many others given the central place our projects hold in scientific open source.

To assess and monitor our progress towards the desired outcomes, we will adopt the DEI metrics introduced by the CHAOSS Working Group. These include diversity in leadership, the effectiveness of mentoring and how it supports diversity and inclusion in each project, and how to structure sponsorship by long-term members of the projects to support a wider representation of gender, race and geographical location in our projects. Specifically, we will focus on communication inclusivity, diversity of the contributions within each project, contributor recognition, documentation usability and accessibility, issue label inclusivity, communication channels, response times and quality, event diversity, diversity in leadership and governance. To evaluate our efforts, we intend to employ a variety of qualitative and quantitative methods: case studies, interviews, surveys, analysis of GitHub data, website traffic analysis, and the number of contributors in regular community calls. This work is to be done in partnership with an external DEI consultant to inform the action plan we intend to carry out over the two years of this proposed work.

While improving onboarding and outreach should increase the diversity of

our projects' demographics, to improve both equity and inclusion we will focus on diversification of pathways for contributions and contributor retention. Our proposal aims to do that by lowering barriers to contribution (through mentoring, documentation, dedication to accessibility and internationalization) and creating welcoming and safe environments for all.

Finally, we plan on producing detailed reports on the actions executed and if they have impacted our projects in terms of representation and interaction with our communities. These reports are meant to both internally guide the actions of each project and externally serve as benchmarks to maintainers in other projects and communities that can benefit from our work.