

Building Capacity for NumFOCUS Open Source Projects to Develop and Sustain Industry Relationships

Narrative for the proposal to the Alfred P. Sloan Foundation from NumFOCUS

Approved for funding: June 23, 2016

Requested Start Date: July 1, 2016

Overview

The goal of this project is to support and promote the sustainability of Open Source Projects (OSPs) by providing training and infrastructure for OSPs to develop business and financial plans, improve communication and build industry relationships and broadly develop strategies for fostering these relationships.

NumFOCUS is a 501(c)3 that supports and promotes world-class, innovative, open source scientific software and is the Fiscal Sponsor of 15 open source projects, providing a context for Projects to coordinate activities. This project aims to 1) connect the NumFOCUS Projects to each other to jointly develop strategies for partnering with industry, 2) provide training on business and financial planning and marketing strategies, including effective communication, 3) connect Project leads with people with relevant expertise and networks, including through training workshops and an Advisory Board and 4) support infrastructure that would help the Projects more effectively manage finances and client and business relationships.

Major activities that would be involved in this project are training and coordination workshops for OSP leads, the establishment of an Advisory Board for the NumFOCUS OSPs and the development of advisory program for the projects by the NumFOCUS Projects Director.

Expected products are two workshops providing training in business development practices to Project Leads, a jumpstart guide to sustainable project maintenance - an Open Source Nurturing and Actionable Practices Kit (OSNAP Kit), and a publication based on the activities that evaluates the effectiveness of the strategies for establishing OSP/business relationships.

Expected outcomes are OSPs that have developed and implemented business development plans and built relationships with industry partners who can help support their longer term sustainability and new knowledge on the factors important for developing and maintaining sustainable open source projects

What is the main issue, problem, or subject and why is it important?

Open Source infrastructure is at the foundation of much industry and scientific software and research efforts. This infrastructure includes the tools that help developers and researchers build software and conduct analyses - programming languages, libraries, frameworks and even training materials. These are projects that have great value now; they're being used on a daily basis by thousands of people, but it's not clear how to determine or find the necessary resources to support and sustain Open Source Projects (OSPs).

One challenge comes from the projects themselves. Most project leads are developers with deep technical knowledge, but little training or experience in the business administration required to develop and run a sustainable project. It is difficult to seek funding without knowing current or future needs or being able to provide business or strategic plans. Another challenge comes from a lack of awareness in industry and academic environments of the need for support. Open source software is often considered 'free' or is thought to already be well-supported. Additionally there are not clear pathways on how to interact with the projects to provide resources. There are often not strong ties between open source and industry communities, relationships that if developed, could help both the OSPs and the industry or academic organizations. These relationships can provide better communication between the OSP developers and the communities that are using the tools, highlighting opportunities where they can work more effectively together and how OSPs could be more directly supported.

There are individual OSPs and businesses that have built effective relationships and business strategies, but these approaches are being developed independently without shared information or resources, repeating efforts and wasting energy.

NumFOCUS is a 501(c)3 that supports and promotes world-class, innovative, open source scientific software and is the Fiscal Sponsor of 14 open source projects that exist in different sectors of the scientific software ecosystem. The NumFOCUS name comes from NumPy, the first sponsored project, and FOCUS as the acronym for Foundation for Open Code for Usable Science. Projects range from NumPy (fundamental infrastructure for array-oriented computing in Python) to domain-specific packages such as AstroPy, and inclusive of new technologies such as Julia (language for scientific computing) and educational communities such as Software Carpentry and Data Carpentry, including many that the Sloan Foundation does or has supported. The scopes of these projects touch different areas in the ecosystem, but all struggle with issues such as sustainability and viability. While some of these projects have industry sponsorship (pandas, Jupyter), the majority of these projects have either little to no support or have support from private foundations or grants that will last 1 to 2 years. The grants and foundation funding are typically bootstrap funding to help support initial development of the software or project. They are not designed to support the projects over the long term, but instead to get them started. In the case of many government grants, for instance from the National Science Foundation or the National Institute for Health, the projects are included as

Education and Outreach or Products components and there is no attention to the sustainability of the project after funding ends. Private foundation grants often include requirements that the project discuss its plans for after the grant ends, with the idea that a model will be developed and implemented during the funding time period and will become a self-sustaining organization. It is therefore left to each of the projects to build not only the product they're developing, but a sustainability model during this time of initial funding. However the PIs on these grants are typically the technology leads and have an academic background, more CTO than CEO, and lack the knowledge and experience to build these plans.

The Project Leads recognize that they "don't know what they don't know" and are challenged to figure out where to get started on how and what to learn. Project Leads have consistently been coming to NumFOCUS staff for guidance and advice. We therefore conducted an informal survey of the projects, and found that they are interested in increased communication between the projects, opportunities for training and direction on what to learn, support in building business plans and help with networking outside of academic and open source communities.

In the same way that organizations such as YCombinator provide training and mentorship for early startups and now non-profits, we plan to build a program and a pipeline for supporting the growth and sustainability of OSPs. However, this is an area where NumFOCUS does not currently have adequate staffing and enough in background, experience, or capacity to provide leadership. Most funding coming to NumFOCUS is supporting the core staff that administers and runs NumFOCUS. The two primary sources of revenue are administration fees collected from projects and the PyData conferences. This grant will provide support for staff that goes beyond supporting NumFOCUS core activities and will bring on board a Projects Director with the experience and vision to help shape this program. This program will also bring additional revenue to NumFOCUS through project administration fees and additional industry opportunities, which will ultimately continue to support the Projects Director role and aid in the sustainability of NumFOCUS.

The goal of this project is to support and promote the sustainability of OSPs by providing guidance for and bringing together the Projects to provide training and infrastructure for them to be able to develop business and financial plans, improve communication and build industry relationships. More broadly this project will develop a roadmap of strategies for fostering these relationships. NumFOCUS, as a Fiscal Sponsor of 14 open source projects, provides an ideal context for Projects to coordinate these activities. ***This project aims to 1) connect the NumFOCUS Projects to each other to jointly develop and share information on sustainability strategies, 2) connect Project leads with people with relevant expertise and networks, including a Projects Director to be hired on this project and an Advisory Board, 3) provide training on business and financial planning, marketing strategies and effective communication, 4) develop and disseminate an Open Source Projects Guide to fundraising and project management and 5) support infrastructure that would help the Projects more effectively manage finances and client and business relationships.*** Another primary goal will be to make this program itself self-supporting, so that it will continue past the end of grant-funding, as we see the need for these activities only growing over time.

In working to achieve these goals, the Projects will also be generating data and information on the factors that are important in developing and continuing industry partnerships as a part of robustly sustaining Open Source projects. Information generated from this project can help guide other Open Sources Projects and industry groups on this path. As OSPs differ in size, scope, objective and community, there will be no 'one size fits all' approach, and the variety of OSPs within NumFOCUS will provide a perspective on these different needs.

What is the major related work in this field?

Several organizations have served as inspiration for both NumFOCUS as a whole and the proposed work in this document. These include the Apache Software Foundation (ASF) and the Software Freedom Conservancy (SFC). ASF specifically includes an incubation process for developing a self-sustaining project, wherein accepted projects undergo review, mentoring, and potentially graduation to a fully-fledged ASF project.

A number of programs have identified issues with sustainability of software from academic sectors, and worked to identify causes, strategies, and trajectories for such projects. These include the NSF SI2 principal investigator meetings as well as the Workshop for Sustainable Software for Science: Practices and Experiences. Both of these programs are much more focused on the mechanisms for sustainability within the academic sector; how do career trajectories work and how can they be improved, how can software be designed for sustainability within academic environments, and how to provide sustained investment of intellectual and monetary resources. We are taking a different approach, identifying mechanisms that align with industry practices. The projects under the NumFOCUS umbrella, in many cases, provide infrastructure for other, external projects; in that way, the means of sustainability must be developed differently from purely academic projects. Furthermore, we aim to identify trajectories for these projects that correspond with trajectories found in industry and that can learn explicitly from industrial practices in project management. In some cases, this will differ quite extensively from those found in academic environments.

What is the approach being taken?

To achieve these objectives we will **convene an Advisory Board** of industry and open source leaders to advise projects on business, financial and community matters. We will **recruit and hire a Projects Director** who will lead, develop and implement the outcomes and outputs described and act as an advisor to the project leads on business development strategies. We will **develop and convene meetings of project leads**. In year 1 this meeting will train them on the important elements of running a non-profit organization, including financial projections, business plans and connecting with industry partners. This will be akin to a 'business bootcamp' with outside trainers coming in to speak. In year 2 this meeting will continue training based on identified needs from year 1 and allow the projects to share outcomes and effective, or

ineffective, strategies employed. Outcomes from these meetings will be used to **create and disseminate an Open Source Projects guide to fundraising and project management**. This project will also **support infrastructure** that can be shared across projects, such as CRM or accounting software and support for web development and the development of other communication materials. Specific actions are outlined below for the two years of the grant.

Convene an Advisory Board

We have already begun the process of convening an Advisory Board for this project. Initial members are being identified by the NumFOCUS Board of Directors and the NumFOCUS community. The Advisory Board will include a diversity of members from industry, academic and non-profit backgrounds with various expertise, also with attention to gender and ethnic diversity. Ms. Silen will draft a document of the opportunities and expectations for Advisory Board members and invite initial candidates to be on the Board. In the first month a subcommittee of the Advisory Board and Ms. Silen will be involved in the hiring of the Projects Director. The Advisory Board's expertise will be invaluable in the final drafting of the job ad and recruiting and hiring for the position. Once hired, the Projects Director will be involved in additional recruiting for the Board. By month 3 of the grant the first meeting of the full Advisory Board will occur, with the meeting led by the Projects Director until an Advisory Board president is selected. The Projects Director will be the liaison between the projects and the Advisory Board. The Projects Director will also establish with the Advisory Board the best communication approaches for Advisory Board members to interact with and advise projects. For the duration of the project, the Advisory Board will hold quarterly meetings and work with individual projects as needed, or for a maximum number of hours.

Hire a Projects Director

The Projects Director, to be hired as a part of this project, will be the lead on these initiatives. This will be a key role, as there is not currently staff within NumFOCUS to focus on the business development of the projects or coordinate activities and efforts. The hiring process will be initiated upon hearing the outcome of the grant. A subcommittee of the Advisory Board and Ms. Silen for the project will conduct the search. The Projects Director will have a strong background in business development, marketing and/or operations and experience in a management or advisory role (job description provided in the Appendix). A draft timeline of activities for the Projects Director is outlined below, but he/she will be the one with the expertise to best set milestones and timelines and will do so in coordination with the Advisory Board and NumFOCUS staff.

Once hired, in the first month, the Projects Director will conduct interviews with the project leads from each of the NumFOCUS OSPs to identify areas of concern and needs for training and support and will work with NumFOCUS staff to determine workshop dates and location. He/she, in consultation with the NumFOCUS Board of Directors, will identify and recruit additional people for the Advisory Board for the project. In the second month, he/she will identify skills or perspectives that would be fundamental concepts to be included in the 'business bootcamp' workshop, and speakers who could present or provide tutorials on these topics. They will

coordinate the speakers for the workshop. NumFOCUS staff will handle workshop logistics, including travel for project leads. In the second month, the Projects Director will work with the projects leads to inform them of useful metrics and information to have for the 'business bootcamp' and advise them on how to gather that information. They will also review the various projects web sites and resources to have a more complete picture of the projects' needs and areas for potential improvement on communication. They will consult with the projects as needed on these topics. In the fourth month the 'business bootcamp' will occur. The Projects Director will focus on this workshop and developing the online infrastructure for communicating outcomes with workshop participants and more broadly online. After the workshop, the PD will assemble and disseminate materials, include summaries and videos of presentations. The PD will begin to develop the topics that will be included in an Open Source Projects guide to fundraising and project management, an Open Source Nurturing and Actionable Practices Kit (OSNAP kit), (more details below) and work to identify or develop necessary materials. At the third quarter of year 1, the PD will begin planning the year 2 workshop, based on experiences from the first three quarters of the grant, especially identifying gaps in training. In year 2, the PD will organize sessions for the year 2 workshop and run the workshop. After the workshop he/she will again assemble and disseminate materials. By the end of year 2, the OSNAP kit will be completed and made available on FigShare. Materials will be publicly developed, so available even while under development.

Throughout the duration of the project, the PD will consult with projects, and act as a liaison with industry leaders throughout the duration of the project. He/she will also attend Advisory Board meetings and act as a liaison between the Advisory Board and the projects throughout the project, and establish communication strategies for Advisory Board members and projects to best interact.

Through the project, the PD will also be building continued support for the project and the PD position. The Sloan funding is specifically bootstrap funding that would get this project started, but the need for these efforts will continue as more projects come on board and the program expands. By the end of year 2, more projects will have come on board at NumFOCUS. This training and support will be crucial to them, and the sustained support and attention of Projects Director for all projects will help the projects grow and support their sustainability.

Develop and convene meetings of project leads

In year 1 the workshop planning will commence once the Project Director (PD) is on board, anticipated to be within the first month of the grant start date. The PD will lead the efforts to set an agenda and invite speakers. Leah Silen and NumFOCUS staff will handle workshop logistics, including venue and travel details. During the workshop, anticipated to be in the first 4-6 months of the grant, the PD will run the workshop, with the support of the Executive Director. After the workshop the PD will be responsible for assembling and disseminating output, including videos of presentations and summaries of the sessions. The PD will also integrate these materials and training information into an Open Source Projects guide to fundraising and project management, an Open Source Nurturing and Actionable Practices Kit (OSNAP kit). Planning for the workshop

in year 2, will commence 9 months into the first year, approximately 6 months after the first workshop. Dates will be set and topics begin to be identified for the sessions, with particular attention to the gaps being identified in year 1. The PD will run the workshop in year 2, with logistical support from other NumFOCUS staff. After the workshop the PD will again be responsible for assembling and disseminating materials and working to complete the OSNAP kit based on outcomes from the workshop.

Open Source Projects guide to fundraising and project management

A guide, the Open Source Nurturing and Actionable Practices Kit (OSNAP kit), will be designed as a “kickstart” for developing or building sustainability, and will be accessible to both projects that are directly engaged with through this grant as well as publicly available. This will be a living, evolving resource for OSPs to foster better practices, including both socially-relevant and technically-actionable materials (such as templates for project governance documents, user and developer engagement, and devops recipes and scripts). This will provide guidance to projects on what it is they need to learn and where and how to get started, to overcome the current challenge of project leads to find necessary information and resources.

Support infrastructure

As a part of the workshop training and guidance from the Projects Director, necessary software and hardware resources will be identified to support business development. Types of software will be evaluated and selected for reasonable costs and good usability within the context of the projects. Marketing and graphic designer consultants will be identified once the Projects Director is hired, and will be able to be consulted, up to a fixed number of hours for each project, and for NumFOCUS for the duration of the project.

What will be the output from the project?

Output

We divide the output from this project into three primary categories.

1. **Workshops** for project leads: At these workshops, we will be building social connections between projects themselves, between projects and the external participants (such as speakers), and through skill and knowledge transfer at the workshops themselves. The workshops will have direct and indirect impact to facilitate sustainability of the projects and to the development of the OSNAP Kit (see below).
2. The Open Source Nurturing and Actionable Practices Kit (**OSNAP Kit**). This collection of practices, guidance, and materials for fostering sustainability will be accessible in online form, downloadable form, and will be maintained through a version controlled repository. This set of resources will be designed as a “kickstart” for developing or building sustainability, and will be accessible to both projects that are directly engaged with through this grant as well as publicly available. This will be a living, evolving resource for OSPs to foster better practices, including both socially-relevant and

technically-actionable materials (such as templates for project governance documents, user and developer engagement, and devops recipes and scripts).

3. A whitepaper or other position paper coalescing the results of the workshops, the discussions between projects, and including information about the motivation and development of the OSNAP Kit. In contrast to the OSNAP Kit itself, this will be more inline with a traditional academic paper, designed to add to the scholarly study of sustainability of OSPs. This will include process information and reflections on the mechanisms (successful or otherwise) used to engage projects during the course of the workshops and the long-term sustainability.

Outcomes

Outcomes are also divided into three categories.

1. OSPs will be able to directly connect with business partners relevant to their interests or potential sustainability plans. For instance, connecting data analysis projects with “big data” partners in industry, connecting infrastructure projects with product-focused industry partners that rely on those projects, and connecting industry partners that provide consulting services with projects that benefit from such services. This will take the specific form of genuine relationships (likely mediated through NumFOCUS) between industry partners, either through fee or service-based consulting or through memoranda of understanding describing open source contribution policies.
2. Identify and acquire revenue sources for projects. The specific revenue sources are likely to vary considerably by project, but this may take the form of donations of specific fiscal amounts (i.e., “\$50,000 for development of new date time support” or “\$250,000 to implement OpenMP array operations”) or for longer-term revenue sources (i.e., “To support NumPy in your NSF grant, use this text and SLA to provide a portion of your grant funds to support ongoing development through fixed-price consulting services.”) Constructing the fiscal, legal and financial structure will be as important as developing the social structure for such relationships, particularly as the relationships interface with project governance structure.
3. Disseminate information about both challenges and opportunities for funding and support for OSS projects. As an example, support for commercial products is straightforward, as it can be a line item in academic grants, but development of the challenges of OSS development and developing a structure for financial support similar to commercial projects will be important for some paths of sustainability. By increasing awareness of the challenges of sustainability, socializing individuals and projects that rely on OSPs to contribute back in some way (consulting fees, line item support, software as a service, etc) can increase the likelihood of longer term sustainability paths being successful.

Attention to Diversity

There are broad challenges in Open Source Software that we are looking to address as a component of this project. Only approximately 11% of Free/Libre/Open Source Software developers are women (FLOSS Survey 2013)[1]. There is not a clear figure on percentages of underrepresented racial groups, but there are similarly low numbers. While still not fully understood, there are many factors that are likely playing a role in this lack of diversity. Some of these factors include communication style, a lack of opportunity for paid work on open source projects and a lack of clear hiring processes. The training efforts in this project and opportunities to build relationships with industry are components of addressing these issues. Training on communication and hiring is included in the workshops and will be in the Open Source Nurturing and Actionable Practices Kit (OSNAP Kit), and the Projects Director will assist in developing hiring policies for the Open Source Projects (OSPs). The strategies to generate revenue and build industry relationships can provide paid opportunities for people to work on OSPs. While not explicitly included in this project, opportunities for OSPs to interact at workshops and through the Projects Director, allow them to share practices on building and supporting inclusive communities, including Codes of Conduct.

Information products

Information products created in the course of this project include scholarly papers, an Open Source Nurturing and Actionable Practices Kit (OSNAP Kit), workshop agendas and materials and videos of presentations from the workshops.

Scholarly papers will be published in open access journals and any data used in the papers deposited in FigShare with an ODC-By license, which makes the data freely available for reuse with attribution.

Workshop agendas and materials and videos of presentations from the workshops will be made available on the NumFOCUS web site and in public repositories. NumFOCUS commits to maintaining these resources on its web site, to aid in discoverability, as long as it continues to be in operation. However, they will also be available in public repositories. Workshop organization will occur on GitHub, so all materials will be available through a public repository. Presentation videos will be posted on YouTube.

The OSNAP kit resources will be available on the NumFOCUS web site and will also be on FigShare. The OSNAP kit materials will have a CC-BY license.

[1] FLOSS Survey 2013 (<http://floss2013.libresoft.es/results.en.html>)