2021 BSSw Fellowship Application Form

Name: **Amy Roberts**
Organization: **UC Denver**
Department: **Physics**
Position: **Assistant Professor**
Research: **Direct-detection dark matter searches**

Personal web page, LinkedIn, Github:
https://clas.ucdenver.edu/physics/amy-roberts-phd

How did you hear about the BSSw Fellowship?
Searching for US Software Sustainability Institute

Describe your work relevant to scientific software.
I have worked with data acquisition, simulation, and analysis software throughout my career. Currently, I serve as the software working group chair of the Cryogenic Dark Matter Search collaboration. Among other efforts, I lead the effort to build a web-accessible analysis environment using the JupyterHub platform. This analysis environment allows scientists who are new to our collaboration to make progress on dark matter analysis in weeks rather than months.

My groupâ€™s role in the new analysis platform has been to â€œcontainerizeâ€ our analysis software stack into a Docker image. This containerization started with our reconstruction software, which is notoriously difficult to compile due to its dependencies on particular versions of ROOT, BLAS, and boost libraries. Building this software successfully in a container allowed my group to run the first-ever data-reconstruction workshops and to containerize other software that relies on this code.

I maintain the collaborationâ€™s git repositories and am currently leading the migration from an in-house host machine to GitLab. I am the primary contact for git questions within my collaboration and continuously work with my students to improve contribution instructions for common collaboration software. I also work with students and faculty in the Denver area through informal â€œGit Togethersâ€ (a term coined by student Stephanie Dehlin) where we focus on specific version control issues faced by attendants.

Describe your background and experience relevant to being a BSSw Fellow.
I believe that complete, usable software infrastructure - together with accessible training and documentation - are necessary for a successful and equitable scientific community. My efforts towards this future are a cornerstone of my research efforts. I work to address systemic computing issues in the dark matter community, my collaboration, and on campus.

At the community level, I serve as an editor for the Journal of Open Source Software, which provides peer-reviewed publication credit for software work. In addition, I am part of a cross-experiment team that has received funding to address common training issues across the dark matter field.

Within my collaboration, I serve as the Software Working Group chair. I lead efforts to improve our analysis environment; the prototype allowed my students to create our first interactive analysis tutorials.

Collaborations with the XSEDE Extended Collaboration Support Services and the Open Storage Network will enable re-analysis and experimental data analysis that would be impossible without this platform.

On campus, I serve as our first XSEDE Campus Champion and am a co-organizer of the annual CU Denver Data Science Symposium, a focal point for the Denver data science community. This symposium led to a collaboration between PIs that secured funding for our first all-campus computing cluster. I successfully advocated for a dedicated educational queue on this cluster and am eager to incorporate this resource into my courses.

What would you do as a BSSw Fellow?
Often, my scientific collaborators do not know how to use version control. And when science requires software collaboration, they struggle to participate.

I plan to improve functional knowledge of version control in the scientific community by making version-control training materials that focus on user stories. I will (1) develop story-based, interactive lessons, (2) commission art that maps common programming scenarios onto version control solutions, and (3) improve these materials with group testing.

Existing version control training materials focus either on the underlying theory or a comprehensive overview of version control. Such material can be inaccessible to new developers if they cannot identify what information is applicable to their specific problem.

Software Carpentry community lessons focused on specific version-control situations would make it easier for scientists to find step-by-step information for their specific needs. I will also commission posters that are companions to the step-by-step tutorials. Posters will illustrate basic properties of version control software and help new users translate everyday concepts into the equivalent version control terminology.

Finally, I will recruit a small group of researchers who want and need version control support. A programmer in my lab will provide rapid support for any issues they encounter, and I will work with my programmer and the research groups to improve the materials.

What impact do you foresee from your efforts?
Version control is a foundational skill that opens the door to effective scientific software collaboration. It is a prerequisite for scientific work in the Cosmic Frontier and required for many best practices like packaging code for reproducible analyses.

For new developers, infrequent developers, and analyzers whose development is restricted to analysis code, sorting through the many tutorials on the web to find answers requires hours and sometimes days. Given the extraordinary time constraints many scientists face, it is not surprising that many clone a repository and then never interact with version control again, leading to lost or unshareable code.

This work will build a foundation across communities that use scientific software. Scientists working on individual or small-group software will be able to find documentation tailored to community use-cases. Developers working on larger-scale community projects can write robust contribution instructions by pointing to tutorials that already exist. And early-career scientists will benefit from the improved training ecosystem

and from projects that have approachable contribution instructions.

Which BSSw focus areas and subtopics apply to your proposed work? Specify all that apply.
Focus Area: Reliability; Subtopic: Reproducibility

Focus Area: Better Planning; Subtopic: Software Process Improvement

Focus Area: Better Development; Subtopic: Documentation

Focus Area: Better Development; Subtopic: Revision Control

Focus Area: Better Skills; Subtopic: Personal Productivity and Sustainability