

# The Mozilla Foundation

## Teaching Scientists to Think Like the Web

Accelerating Scientific Discovery Through the Effective Use of Technology

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

Sharing information on the web, automating common processes, managing large volumes of data, and similar tasks are no longer the sole preserve of professional programmers. Increasingly, journalists, filmmakers, educators, artists, and other “end user programmers” find it necessary not just to use, but to create new software. This is especially true for scientists, yet the training available is usually outmoded, overly complex, or focused on the wrong skills. Mozilla and Software Carpentry hope to change all this.

A decade after Udell's seminal report *Internet Groupware for Scientific Collaboration*, only a small minority of scientists use computers and the web to their full potential. The hidden costs of this are painful: tasks that should take minutes wind up taking hours, insights are missed, and collaboration is impeded. A 2008 survey found scientists spend 30% of their time wrestling with software, and most expect this figure to increase.

The root cause is a lack of the basic skills that allow scientists to create and customize software or use the web as more than a publishing medium. But it does not have to be like this: open source software and browsers' ubiquitous “View Source” allow everyone to “look under the hood” to see how things are done. Scripting languages, HTML5, GitHub, and the like permit a Lego-like approach to programming that can allow scientists to manipulate data sets, crowdsource solutions, and share findings — *if* they know how.

## 2. What is the major related work in this field?

---

A number of studies on how scientists use computers and the web have appeared in the past decade [HB08, SK08, Seg08, Seg09]. The largest, [HLM+09], found that most scientist learn what they know about using computers and the web through osmosis, which leads to crippling gaps in their skills. On the education and training side, only a handful of the more than one hundred papers presented at [SIGC11] focused specifically on scientists. Those that did invariably asked, "How can we use computers to teach science?" rather than, "How can we teach scientists to make technology do what they want?"

Many recognize that this lack of skills is slowing scientists down, but most existing training meant to address the problem is flawed in one or more ways:

- **Does not target scientists' specific needs.** Most "Computing 101" courses are run for students from a range of disciplines, so applications and examples often fail to engage students from the sciences.
- **Too much emphasis on programming.** Programming is only one part of building useful software and using the web. Scientists almost always have to figure out "the other 90%" (discussed below) on their own.
- **Too much emphasis on calculation.** Number crunching is also only one part of how scientists use computers today. Managing data and sharing ideas with colleagues are already key to effective practice, and becoming more so every day.
- **Too much emphasis on "big iron".** Scientific computing is often identified with high-performance computing, which skews discussion and training toward the concerns of a small (but vocal) minority.

Dr. Wilson's Software Carpentry project first started working to address these shortcomings in 1997. Now in its fourth major revision, the Software Carpentry web site has an active user base of 350 to 1000 individuals per day and its content regularly appears in courses delivered at laboratories and universities. Dr. Wilson's experience indicates that a modest investment in training can increase scientists' productivity significantly, while making their technology-based work more reliable and shareable. Hard data is difficult to obtain, but follow-on surveys, qualitative feedback, and testimonials often report "order[s] of magnitude" improvement in productivity; even the most conservative of these are typically phrased as "saving [me] a day a week".

The key to increasing productivity is to focus on fundamental skills such as version control, testing, task automation, data management, and program design. While these are not as exciting as things with "cloud" and "peta" in their name, they are what actually empower scientists to solve today's problems efficiently and tackle new ones tomorrow. Our initiative will build on Software Carpentry's experience, and on Mozilla's efforts to make technology easy to understand and use through its work on Firefox, standards-based computing, and the Mozilla Developer Network.

### **3. Why is the proposer qualified to address the issue or subject for which funds are being sought?**

---

#### **Mozilla Foundation**

Mozilla is a non-profit, 501(c)3 organization dedicated to promoting openness, innovation, and opportunity online. Best known as the maker of Firefox, we work to empower individuals to use and shape technology to their own ends. The principal activity of the Mozilla Foundation is teaching "webmaking" to non-programmers, such as

scientists. The activities detailed in this proposal will build on three, existing Mozilla programs that support this objective:

- **School of Webcraft (SoW):** A partnership with P2PU and an online platform to support self-, peer-, and expert-led instruction and study groups. The program will house the online training developed and delivered under this initiative.
- **Mozilla Developer Network:** A large repository of DIY resources and best practices on how to build and create with the technology and tools of the open web. More than 10 million people visited the MDN web site in the last year.
- **Open Badges:** A distributed accreditation framework to support the award and display of badges by peers, experts, and institutions. Badges are a key mechanism through which to incent, recognize, and expand participation in distributed, peer-led, and other forms of learning.

### **Dr. Greg Wilson**

Dr. Wilson is a 25-year veteran of the software industry who received ComputerWorld Canada's "IT Educator of the Year" award in 2010 and was co-winner of the Jolt Award for Best General Technical Book in 2008. His Software Carpentry initiative, which began as a training course at Los Alamos National Laboratory, has been accessed by more than 100,000 visitors since May 2007. The materials are freely available and have been used in courses at over a dozen universities and labs in six different countries.

#### 4. What is the approach being taken?

---

##### A Five-Point Approach

To turn scientists from passive consumers of software into empowered users and makers, a successful approach must:

1. **Target graduate students.** Their time is more flexible than that of undergraduates, but they are still focused on learning. They are often face-to-face with the challenge of making computers and the web work for them, instead of the other way around.
2. **Provide peer- and institutionally-recognized rewards** to encourage students to make acquiring these skills, and passing them on, a priority.
3. **Solve immediate problems.** Scientists always have pressing deadlines, so any training must be seen by them to solve problems that *they* realize they have.
4. **Use face-to-face instruction** as a complement to online learning. A 2010 report from the US Department of Education found that combining the two produced better results than either on its own, which is consistent with our experience.
5. **Engage scientists in a larger learning community**, so that they will pass their skills on to an ever-larger circle of colleagues.

##### **Institutional Engagement**

Much of Mozilla's work seeks to challenge and transform established practices within various fields. Borrowing from agile development methodology, Mozilla builds, launches, and tests new programs in short, iterative cycles; projects are allowed to experiment, fail, and regroup without significant up-front planning. Traditionally, this innovation takes place outside of institutional contexts to avoid potential pitfalls: burdensome process requirements, never-ending calls for "more research", and

continual re-design at the planning stage. However, the nature of this project – working with graduate students – provides an impetus and opportunity for Mozilla to explore how to overcome these obstacles and engage with formal institutions. We have allocated a significant portion of the budget to engage computer science faculty in the delivery of the training content. Once we have established a successful framework, we will work with existing faculties and institutions to see the program become a standard component of their scientific training. Our hope is that the resulting exchange will provide a model to facilitate future collaboration between Mozilla and academic institutions.

### **The Resulting Program Framework**

Under the experienced leadership of Dr. Wilson, Mozilla will:

1. Migrate existing and produce new training materials that directly address the technology learning needs of scientists;
2. Design and launch the first iteration of a self- and peer-led learning and badge program through the SoW;
3. Organize and document the results of 4 in-person workshops; 2 through grassroots, peer-based instruction and 2 through conventional coursework at universities;
4. Work with at least 4 institutions to examine their requirements for formal engagement and the necessity of this engagement to affecting the desired learning outcomes; and
5. Gather and document the project findings to underpin future program iterations.

*Training Materials:* The materials available through MDN and Software Carpentry will be framed to the specific needs of scientists. Mozilla community members will produce

screencasts showing how to perform tasks of use to scientists. The resulting videos will be enriched with instructional and reference metadata using Mozilla's Popcorn.js technology, which allows for the integration of web content and video.

*Online Learning:* Self- and peer-led courses will be offered through the School of Webcraft. Participants will complete learning challenges, find support, ask questions, and connect with the broader community of scientists across disciplines and institutions. A badge program will provide near-term incentives for both learning and mentoring; a framework to support viral, peer-driven engagement with the program; and facilitate recognition by partner institutions and potential employers.

*In-person Workshops:* We will run 4 in-person workshops at colleges and universities across the United States, Canada, and the United Kingdom. (Letters from universities and colleges indicating their support for this are appended to this proposal.) The workshops will be hands-on, interleaving short tutorials with live coding sessions. Two workshops will follow a peer-led, grassroots model and take place at universities but outside of formal, faculty engagement. Two additional workshops will be offered in concert with faculty at universities. The resulting baseline will facilitate a comparison to inform future efforts. We presently have strong expressions of interest in hosting such workshops from:

- University of Wisconsin – Madison
- Michigan State University
- University of California at Berkeley
- Georgia Tech
- University of British Columbia
- Utah State University
- Indiana University
- Queen Mary University London

A sample workshop outline is provided at <http://software-carpentry.org/sloan2011.html>.

## 5. What will be the output from the project?

---

1. **New, tailored technology training materials for scientists.** Workshop curriculum and instructional materials. 15 hours of “how-to” video content produced by the Mozilla community and enriched with Popcorn.js enabled metadata.
2. **Online training in ‘webmaking for scientists’.** A minimum of 80 students completing 10 self-led learning challenges, and a badge program offered through the SoW and supported by the Mozilla community.
3. **Pilot implementations of in-person workshops.** Two workshops delivered through a grassroots, peer-led approach, and two additional workshops delivered in concert with university faculty, both continued online through the self-led learning challenges mentioned above. Documented analysis of the relative success of each approach, as well as comparisons to the online-only workshops.
4. **Recommendations and plan for institutional engagement.** Results and feedback from institutions and computer science faculty regarding their interest and requirements to engage in future iterations of the project. Evidence-based recommendations regarding the importance of this approach.

### Metrics

The ultimate measure of our success will be whether scientists can do more research in less time and tackle problems they could not have tackled before. Both are difficult to measure, especially in the short term, so we will use several proxy metrics to gauge the project's success:



1. **Repeat participation and peer recruitment.** The percentage of online course participants that offer or plan to offer in-person workshops and study groups on their own, as well as recommend the online courses to their peers.
2. **Badge display and associated reputation.** The number of participants showcasing their badges through social media and other web sites will provide insight into the real and perceived reputation of the program.
3. **Institutional engagement.** The number of universities who choose to experiment with and/or integrate the materials into their overall scientific training.
4. **Former students become makers.** Mark Surman, our Executive Director, recently wrote, “Everything we’re doing is about learning through making and collaborating on the web.” Early participants in this program creating content for others to use will be the surest possible sign of success.

## **6. What is the justification for the amount of money being requested?**

---

### **Project Direction**

Dr. Wilson will manage the project with support from Mozilla’s Executive Director, Mark Surman, and Learning Program Officer, Michelle Levesque. Dr. Barbara Hueppe of Mozilla’s communications team will work on media and community outreach. A post-graduate student will be contracted to establish, monitor, and report on performance metrics.

### **Project Implementation**

*Content Development:* Members of the Mozilla community will produce fifteen, 1-hour screencasts to serve as training materials for the workshops.

*Online Course & Badge Development:* Existing and new training materials will be converted into SoW self- and peer-led learning challenges, as well as an associated badge program. The work is expected to take 3 months. An additional amount will be required for technical production of the badge media assets.

*Workshop Delivery:* Local experts recruited from the participating institutions will deliver 4 in-person workshops with support from Dr. Wilson.

*Institutional Engagement:* We will hire an individual to lead our outreach to and conversations with formal institutions. This individual will also prepare a report with recommendations on how to proceed with future iterations of the program within an institutional context. The work is expected to be a 3-month, part-time commitment for a total cost of \$9,375.

## **7. What other sources of support does the proposer have in hand or has applied for?**

---

*Mozilla Developer Network:* Launched in 2003, MDN currently has 40,000 members, and has received over \$2M in investment from Mozilla since its inception.

*School of Webcraft:* Since 2009, the SoW has received \$250K from the Mozilla Foundation, as well as benefitted from a \$400K grant to P2PU from the Hewlett Foundation; part of the Hewlett grant has been used to pilot the use of Open Badges within P2PU.

*Open Badges:* This new program has received \$1,150,000 from MacArthur since 2010.

*Software Carpentry:* has received a total of \$230,000 in the past two years from Indiana University, MathWorks, Michigan State University, Microsoft, MITACS, Queen Mary University of London, SciNet, SHARCNET, and the UK Met Office. Follow-on proposals

to QMUL (for workshops at 4-6 UK universities) and Microsoft (for updating Software Carpentry's Windows-specific materials) are currently being prepared.

#### **8. What is the status and output of current and/or previous Sloan grants?**

---

Neither Mozilla nor Dr. Wilson have been involved in previous Sloan grants.

References and a sample workshop outline can be found at <http://software-carpentry.org/sloan2011.html>