Grant to The Carpentries

Sloan Foundation Grant No: G-2018-11120

Title: Development of data skills curriculum and scalable open and collaborative lesson development infrastructure

Narrative Annual Report

In year 1 of our 'Development of data skills curriculum and scalable open and collaborative lesson development infrastructure' grant, we focused on laying the foundation for the domain specific curriculum development, hiring a Curriculum Development Lead, and updating the social and technological lesson development infrastructure around both lesson development and maintenance.

In our Narrative we've outlined progress on grant goals and deliverables, including areas where we have updated our strategic approaches.

Grant deliverables:

1. Lesson templates and guidelines for open and collaborative lesson development in general
2. Open curricula for two-day workshops focused on the foundational skills for data analysis and management in Economics and Image Analysis
    - A minimum of 6 pilot workshops taught in different countries for each curriculum
    - Development of strong networks in each domain for future workshops
    - An assessment instrument for each of these curricula
    - An assessment report on outcomes of workshops
3. An outline and set of milestones for the development of a curriculum in Chemistry

Key milestones achieved in Year 1:

- Established teams and committees, with clear roles and responsibilities, to oversee, lead and conduct the work in this proposal.
    - Staff Curriculum Team that includes Dr. François Michonneau (Curriculum Development Lead) and Dr. Erin Becker (Associate Director) to focus on curriculum development infrastructure and specific lessons.
    - Lesson Infrastructure Committee that oversees the technical development of the lesson infrastructure and templates.

- Established lesson development roadmap that includes release stages and updated approaches to development based on our experiences in year 1.
  - Established lesson development milestones - alpha, beta, and stable - that better guide people through the development process and the community in the use of the lessons.
  - Updated our approach to the development of the alpha version of lessons, focusing the first stages on a few core contributors rather than a broad community effort.
- Initiated a human-centered design approach to the update of lesson templates
  - Created 'personas' to understand how lesson templates are being used and by whom in our community, to guide a human-centered design approach to the lesson template.
  - Evaluating technical aspects and usability of lesson templates.
- Initiated a pathway and mechanism for community contributed lessons
  - Started 'Carpentries lab' to create a pathway and mechanism for community contributed lessons, complementing core lessons, inspired by rOpenSci's model for package development.
- Developed curriculum on 'how to develop curriculum' to provide guidance for the lessons in this proposal, as well as other Carpentries lessons
  - Released an alpha version of a [Curriculum Development Handbook](#) to teach people how to develop curriculum.
- Developed infrastructure for the maintenance of lessons and for building an Instructor community around new lessons
  - Established Maintainer community and guidelines including maintainer onboardings, maintainer meetings, and clear roles and responsibilities
  - Established [Curriculum Advisory Committees](#) that oversee and provide guidance on content-level decisions for domain-specific sets of curriculum.
  - Established [instructor onboarding](#) to introduce instructors to new curriculum.
- Initiated of an alpha version of Economics curriculum
  - Working with Dr. Miklos Koren, an associate professor in the department of Economics at European Central University and a senior research fellow at the Institute of Economics.
  - Planning initial pilot workshop for June
- Almost complete alpha version of [Image curriculum](#)
  - Working with Dr. Tessa Durham Brooks and Dr. Mark Meysenburg at Doane College on curriculum for undergraduates and graduate students, supported by an NSF IUSE grant.
  - Two pilot workshops complete, with a third planned in May
- Initiated training of instructors with backgrounds in economics and chemistry

Below we include more information on each of these milestones as they relate to our three primary deliverables.

1. Lesson templates and guidelines for open and collaborative lesson development in general

The foundation of this project is to improve and develop our lesson development and maintenance infrastructure. This infrastructure allows for more effective development of the specific curriculum in this proposal, and also establishes processes for the development and maintenance of all of our future and current lessons. Across Software Carpentry, Data Carpentry and Library Carpentry we have more than 40 active lessons, with the ones in this proposal, along with others, under active development. Interest and momentum for open, collaboratively developed curriculum continues to increase, and we want to ensure that we can support and foster that interest and demand.

Milestone: Established teams and committees, with clear roles and responsibilities, to oversee, lead and conduct the work in this proposal.

To support this work, in year 1 we hired Dr. François Michonneau and created a staff Curriculum Development Team. François is a long time community member and contributor to Software and Data Carpentry, helping to develop and maintain the first set of Data Carpentry lessons and serving on the infrastructure subcommittee. He also has extensive programming experience and is a domain scientist himself, with a PhD and postdoctoral research in biodiversity and informatics. Dr. Erin Becker, our Associate Director, is the other member of our Curriculum Development Team. She started as the Data Carpentry Associate Director in 2016, and has experience in community engagement, curriculum development, and R development and a background in domain research and educational pedagogy, with a PhD in Microbiology and postdoctoral research in Biology Education. Including both François and Erin on the team has brought together complementary experience and expertise and allowed them each to focus on particular areas of this work. François works with and oversees the work of the Lesson Infrastructure Committee and Erin works with the Maintainer and Curriculum Advisory Committees and communities.

Milestone: Established a lesson development roadmap that includes release stages and updated approaches to development based on our experiences in year 1.

Lesson Development Roadmap

To scale our lesson development process, we have been working on clarifying the development milestones. Matching the terminology used in software development, we have introduced the [alpha, beta and stable milestones](#) to the lesson development process. This process emphasises the iterative and collaborative nature of our approach to lesson development. Naming these three stages allows lesson creators to communicate what has been done and what remains to be done to others interested in contributing to the content or interested in teaching it. It is also

helpful to communicate to lesson creators our expectations for lesson content before lessons can be taught globally when they embark in creating a new lesson with us, and therefore facilitate planning of their development.

The alpha stage focuses on having a first complete draft of the lesson content. It is how the lesson will look when it is first taught (usually by the lesson creators). The lesson creators can focus on choosing an appropriate dataset for the lesson, identifying the skills that should be taught, and determining the type of exercises that will be included in the lesson.

The development leading to the beta release of the lesson is when the community starts to embrace the lesson. At this time the number of contributors to the lesson increases beyond the initial set of lesson creators. This community-review process is useful to ensure that the flow of the lesson is logical, that the content of the lesson is inclusive, and that bugs and mistakes get addressed. This period is also when the "instructor guide" takes shape. This document provides the context and backstory of the lesson. It is what allows instructors who did not follow its initial development to know how to teach the lesson and the design decisions behind how it was assembled. During this period, workshops where this material is taught are "beta pilots" and we collect additional feedback from learners and instructors beyond our regular pre- and post-workshop surveys. We use this feedback to inform lesson design decisions until the lesson is ready for release and to be made available as part of our core offerings.

Updated approach to development

A vital part of our pedagogical model is the importance of limiting cognitive load by focusing on data types that are familiar and skills that are immediately useful for our learners. This focus helps ensure that learners get the maximum possible benefit from our workshops, as they don't need to waste cognitive capacity trying to become familiar with foreign data types. A downside of the domain-specificity of our approach, however, is that our lessons can't be used indiscriminately with learners from a variety of domain backgrounds. We continuously receive requests and proposals for development of new curricula to serve learners in different domains who are working with different types of data.

Dr. Erin Becker has provided extensive support to new curricula throughout the development lifecycle. This support has included advising on content development, organization of community members to serve curriculum development and support roles, and organizing and assessing feedback from pilot workshops. Although we have expanded our curricular offerings within the last year to include full two-day workshops for working with tabular survey data in the social sciences, and geospatial data, this approach to curriculum development does not scale with the growth of our community into new domains.

We have therefore learned how to approach effective collaborative lesson development from our experience with our first set of lessons. For these first lessons, we gathered specialists in a hackathon-style event to produce a first draft. We have found that this process created

heterogeneous and incohesive content that took a lot of effort (most often from our staff) to clean up before the lessons were ready to be taught. More recently, we have experimented with a model where a smaller group of people identify the skills that need to be taught and create an outline for the lesson. Then, a few more people can be added to the group, and with guidance from our staff members and the documents we have been developing, they work together to generate a first draft for the lesson. Once the first draft is complete, we open up contributions to the community at large who can provide input on the content based on their experience teaching these lessons. This approach still requires significant dedicated staff time to produce high-quality lessons because the polishing of the details is tedious and technical expertise is needed to use the template effectively. However, this process produces lessons that are more cohesive and easier to use for the community broadly.

Milestone: Initiated a human-centered design approach to the update of lesson templates

Our lesson template and guidelines for its use are managed by our Lesson infrastructure committee and our Curriculum Development Team member, Dr. François Michonneau. This committee is composed of community members who are familiar with the tools and technical infrastructure The Carpentries uses to manage our online lessons. In year 1, we have restructured the Lesson Infrastructure committee and developed a document that outlines its roles and responsibilities. This committee meets once a month to discuss, provide guidance, and make decisions on the lesson infrastructure. The focus of this committee has been around decisions that make maintenance of the lesson infrastructure easier and more scalable. For instance, we have been implementing continuous integration and deployment for our lessons. This approach replaced the tedious and error-prone manual process that lesson maintainers have been doing. This tool provides direct feedback to maintainers when something is broken in the lesson, allowing them to spot and fix mistakes more effectively. In parallel, we have been developing additional tools that integrate these automated tests to check the integrity of our lessons such as validating the URLs and the images that are included in the lessons.

Our current lesson template is very technology-centered. To use the template effectively, lesson creators have a lot to learn. As we plan the next iteration of our lesson template, we are using a human-centered design approach. To this end, we have been working with our lesson maintainers to develop "personas". This design approach helps us to identify who are the different people who use our template, their needs and goals, and thus the features the lesson template should have to respond to their needs. This approach will create a better experience for all users of the template. This new version of the template will also use best practices in web-accessibility to ensure that our lessons can be used by everyone. We received funding from the R Consortium Infrastructure Subcommittee to create this template as an R package, so that it can be used easily and broadly by the R community, as well as by the Python community or others who work with Markdown.

Milestone: Initiated a pathway and mechanism for community contributed lessons

To spread our approach to lesson design and lesson maintenance, and to allow our staff to focus our efforts on our core lessons, we have developed guidelines to clarify lesson development pathways. In addition of the process we have developed for lesson creators to write lessons that will become part of our core offering, we are putting in place the infrastructure needed for community members to develop lessons on topics that are too specialised to be part of our core offering. By differentiating these two types of lessons, it is easier to communicate our commitment and expectations during the development process.

By organising a space for community-contributed lessons (The Carpentries Lab), with little intervention from our staff, we create professional development opportunities for members of our community to conduct editorial and technical reviews of the lessons. We provide guidelines (see the curriculum development handbook below), lesson templates, and a venue for lesson creators to publish their lessons. This is based on the rOpenSci model for review and integration of R packages into the rOpenSci ecosystem.

Milestone: Developed curriculum on 'how to develop curriculum' to provide guidance for the lessons in this proposal, as well as other Carpentries lessons

To more sustainably support community members in the initial content development phase of developing new lessons, we have produced a Curriculum Development Handbook. This pre-release focuses on early stages in curriculum development (e.g. defining a target audience, designing content). The full Handbook will include materials around recruiting and training lesson maintainers, running pilot workshops, and building a broad community of instructors to teach the lessons. A brief overview of the Curriculum Development Handbook can be found below.

- Chapter 1: Conceptual elements - an overview of The Carpentries approach to curriculum development including the stages of development (extended in chapters 2-5) and the components of a Carpentries curriculum.
- Chapter 2: Deciding what to teach - a guide to the process of defining your target audience and establishing a "skill set" that can be used for the rest of the development process
- Chapter 3: Designing Challenges - covers the practicalities of the second stage of the curriculum development process, creating exercises. Starts with a guide to how to select a dataset. Introduces several different types of exercises and discusses when each is appropriate.

- [Chapter 4: Developing content](#) - builds on Chapters 2 and 3. Walks lesson authors through the process of creating code chunks, narrative sections, and other supporting elements of the curriculum.
- [Chapter 5: Community development](#) - lays out the framework for community support of a lesson, including the roles played by various community segments, training needed for those roles, and when each role should be filled during the lesson development process.
- [Chapter 6: Technological introductions](#) - goes into some detail about how our templates work and how we structure our lessons on GitHub.

We are now using this Curriculum Development Handbook to guide projects that are in the early stages of design, including a new curriculum in Economics. Having this resource available reduces the staff load in communicating about the curriculum development process, and will increase the consistency of lesson quality for community developed lessons. Based on the level of community interest in creating new lessons for particular domains and data types, we are excited about transforming this handbook from a written resource to a curriculum in the future. This would be analogous to our [Instructor Training curriculum](#), and would enable us to actively scale our evidence-based approach to curricular design. Dr. François Michonneau will be leading a pilot of this Curriculum Development workshop in June at a [regional Carpentries conference in Manchester, UK](#). In the upcoming year, we will make this 2-day workshop part of our offering for members.

The initial creation of new lessons is a major step, but it is not the full picture. Lessons are never perfect in their first iteration, and even if they were, any lesson around software or technical skills can become outdated as software and best practices evolve. We've found that it is essential to create a community around a lesson so that it has continuing support for modifications, improvements, and updates and can remain viable in the long run. Our lesson support structure includes two major components, lesson Maintainers - who take care of the day to day process of implementing small-scale improvements to the lesson based on feedback from instructors and learners - and Curriculum Advisors, who provide high-level oversight, vision, and leadership for a curriculum and guide large-scale updates.

Milestone: Developed an infrastructure for the maintenance of lessons and teaching from the Carpentries community

Maintainer community

The Carpentries have had the [Maintainer](#) role in our community for several years, but until recently, Maintainers did not receive any training or guidance in their role. At the end of 2017, [we conducted individual interviews](#) with our Maintainer community, and identified a strong desire for more formal training and guidance in the role. In response to this feedback, we developed and piloted a [Maintainer onboarding process](#) to introduce new and existing Maintainers to the social, curricular, and technical aspects of being a Maintainer. We also

introduced [Maintainer guidelines](#) to provide a clear set of responsibilities for those interested in joining the Maintainer team. In early 2018, we onboarded 27 new Maintainers to our lessons using these materials.

Through Maintainer interviews, we also identified a desire for greater interaction within the Maintainer community. Many Maintainers stated that they had little to no contact with their co-Maintainers, making it difficult to coordinate work on their lessons. In 2018, we re-started monthly Maintainer meetings to provide a space for these community members to come together and discuss issues relevant to their role. Starting in early 2019, we are testing out a new meeting format, which includes a mixture of co-working time ("Office Hours") and skill-building sessions ("Skill-Ups") led by Maintainers or Maintainer teams. Skill-Ups are designed to facilitate skill-sharing within our highly talented Maintainer group and help them level-up their skills for open-source project maintenance - including both technical and social aspects. Office Hours can be either free-form co-working time (using breakout rooms to group Maintainers working on similar lessons) or time for individual Maintainers to get feedback on particularly difficult PRs/Issues in their repos. We have also introduced a Slack channel for Maintainers, as well as for individual lessons, to provide a platform for real-time discussion of lesson changes, bugs, and other issues of relevance to this community. By bringing the Maintainer team together regularly and providing opportunities to co-work and share ideas and skills, we are working to strengthen this community so that it continues to provide a strong backbone for our lessons.

Curriculum Advisory Committees

Other feedback we received from Maintainer interviews was that Maintainers did not feel comfortable introducing large-scale changes into their lessons and were concerned that they did not have the authority to do so. Maintainers signed on for their role with the understanding that they would be responsible for day-to-day maintenance and small improvements, but were often being asked by contributors to implement major overhauls to the curriculum (including introducing new tools or packages or completely changing the lesson narrative). This feedback enabled us to identify a missing layer of our lesson support structure, which is now filled by Curriculum Advisory Committees for each curriculum. Curriculum Advisors provide high-level oversight, vision, and leadership for a curriculum and guide large-scale updates. Unlike Maintainers, who are responsible for the day-to-day work of keeping lessons stable and teachable, Curriculum Advisors maintain a broader perspective on the state of the field and make strategic decisions about major changes to a lesson, for example, updating the technology being taught to take into account major advances in the field or changing the dataset used in the lessons to appeal to a broader group of learners. We now have [Curriculum Advisory Committees](#) for our Social Science, Geospatial, and Genomics lessons.

A Curriculum Advisory Committee (CAC) is composed of 5-8 people with significant domain expertise who represent the breadth of the field that a curriculum is intended to reach. For example, the [Data Carpentry Geospatial CAC](#) includes researchers in ecology, limnology,

environmental sciences, and sociology, along with university staff leading institutional GIS education efforts. Multiple career levels are represented, from PhD candidates at the end of their graduate work, to mid-late career professionals. At least one member of the CAC should be actively teaching in the field, so they can bring a practical perspective about what skills students and early-career researchers need. Curriculum Advisors commit to a minimum one-year term, but may serve multiple terms. A CAC should include members from multiple geographic regions and cultural and linguistic contexts to ensure that the curriculum meets the needs of our global community.

A Curriculum Advisory Committee meets virtually approximately twice a year to discuss and make decisions about proposed large-scale changes to the lessons within their curriculum. These proposals may be initiated by community members, including Maintainers, or by members of the CAC. The CAC communicates their recommendations back to the Lesson Maintainers and provides consultation and support to Maintainers in implementing proposed changes.

This model has been effective in guiding a major plotting system update prior to publication of our [new Geospatial curriculum](#), and in modernizing the data and toolkit used in our [Genomics curriculum](#).

Onboarding instructors to new curriculum

Having a formal structure for maintenance and broad level developmental guidance for each curriculum has been very useful for supporting the lessons and making sure they stay relevant with changes in the field. In addition to supporting the lessons, we've found it important also to support the Instructor community around teaching the lessons and giving feedback to guide the Maintainers and Curriculum Advisors in their work. Our [Instructor Training curriculum](#) prepares Instructors to teach a workshop from the broad perspective of pedagogical principles, evidence-based teaching, and The Carpentries philosophy and community structure, but it does not teach Instructors how to teach particular lessons. With over 40 lessons in The Carpentries curriculum stack, it would be impossible to prepare Instructors to teach all of the lessons within the structure of a two day workshop. However, as we expand our curricular offerings, we recognize the importance of onboarding Instructors to teach new lessons so that Instructors feel prepared and confident and so that our workshops continue to be delivered to a consistent high-standard.

Our approach to this has been two-fold. First, we have leveraged the strengths of our existing Instructor community, who already have a strong background in pedagogy as it applies to teaching our workshops and who are already connected in The Carpentries community. We have begun to offer curricula-specific Instructor onboarding webinars, targeted towards badged Carpentries Instructors. These one hour webinars cover the structure of the new curriculum from the perspective of Instructors who are already familiar with some of our existing curricula. These trainings focus on introducing the data that is used in the lessons and drawing attention to

things that differ from other curricula, where Instructors will need to adjust their normal approach to teaching that tool or skill. These webinars are open to others who are interested in teaching the lesson, but are not yet Carpentries Instructors and have served as an on-ramp to getting peripheral community members interested in and motivated to become Instructors.

We have now piloted this approach for our Geospatial and Social Sciences workshops, leading to a pool of 38 instructors who are certified for teaching the geospatial workshop, and 21 for social sciences. This number provides a nucleus of highly motivated and engaged Instructors, many of whom have already actively begun teaching and planning workshops. To make this approach scalable, we have also made our onboarding videos available through our YouTube channel ([1](#), [2](#)), and direct new Instructors who are interested in teaching these curricula to begin their preparation by watching the recorded webinar. This approach helps us to maintain the high quality of our workshops as we continue to expand into new domains. This process has received positive feedback from participants, and we are looking forward to expanding this as we bring on board new curricula in Image Analysis, Economics, and Chemistry.

The second prong of our approach to creating strong community around new lessons has been to intentionally grow our pool of certified Instructors with expertise in the lesson domain. Historically, our community has been heavily weighted towards certain domain communities, with a [majority of our Instructors](#) coming from backgrounds in biology, computer sciences, or engineering. To grow our community in new domains, we are fast-tracking applicants through our Instructor Training program who have expertise in target growth domains (including economics or social sciences, arts, humanities, or library science).

2. Open curricula for two-day workshops focused on the foundational skills for data analysis and management in Economics and Image Analysis

Much of our focus in year 1 was on the general infrastructure and framework. We also have worked with the community on alpha versions of the Economics and Image curriculum. For both sets of curriculum, rather than starting with a hackathon, a smaller group, which is connected with and has surveyed their domain community, is initiating development. Hackathons will instead be used in the phase of moving lessons from alpha to beta versions and through the pilot process. Information on the development in each of the areas is included below.

Economics Curriculum

Milestone: Initiation of an alpha version of Economics curriculum

Because we are in the early stages of development of an Economics curriculum, we are able to use this curriculum to pilot the approach we've taken in our new Curriculum Development Handbook. This process focuses on starting with a single (or small group of) dedicated lesson authors, who are well-placed to have a broad perspective on learner needs in their field and are

proactive about consulting and collaborating with others in their field from the very beginning stages of designing a new lesson. For Economics, we are working with Dr. Miklos Koren, an associate professor in the department of Economics at European Central University and a senior research fellow at the Institute of Economics. Dr. Koren is an active member of the European Economic Association and an editor for the *Journal of International Economics* and the *Review of Economic Studies*. He is working with our Associate Director and Curriculum Team team, Dr. Erin Becker, to develop an initial draft of a two-day curriculum for Economics, which will include lessons on Stata and the Unix shell and will be organized around an open data set from The World Bank. Working together closely from the beginning will ensure that the foundation is laid for lesson content to be consistent with The Carpentries pedagogical model and community values. Repositories for developing and hosting this curriculum have been created (1, 2), and the lesson author will be actively supported in the technological aspects of lesson creation. Please note that these repositories are not yet populated with content, as we are in the pre-content phases of backwards design curricular development (see chapter 2 of our Curriculum Development Handbook).

Although Stata is not an open-source tool (as are the other technologies that we teach), it is widely used in the Economics research community and learners will likely have some exposure to this software before attending a workshop. By teaching a tool that learners have some familiarity with (and that other members of their research group likely use), we are keeping to The Carpentries principle of "meeting learners where they are". Rather than introducing a high cognitive load by trying to teach new skills at the same time as a new tool interface, we can help learners develop immediately useful skills that will slot into their current workflows and start them on the path towards developing confidence to support future learning. However, The Carpentries highly values democratization of data skills, which requires making our lessons available and teachable in a variety of contexts, including for learners who do not have institutional access to subscription-based tools like Stata. To this end, we plan to complete an initial version of the lesson materials using Stata and then translate those materials into R. A colleague of Dr. Koren has used this approach in the past to convert lesson materials between R and Stata, so we have a model to build on and guide our efforts.

An alpha version of this curriculum (using Stata), will be piloted in late June 2019 at CarpentryConnect Manchester, taught by Dr. Koren. Dr. François Michonneau and Dr. Tracy Teal will be in attendance and able to collect feedback from the pilot for future improvements to the lessons. After this alpha pilot in June, we will use a combination of community contributions and dedicated staff time to prepare a beta version of the curriculum for broader teaching. The beta version is targeted for release in September 2019, at which point it will be available to request for an official coordinated workshop. Starting in September 2019, we will actively promote these materials and arrange for beta pilot workshops at a variety of institutions. This effort will be coordinated with recruitment and onboarding of Instructors for this curriculum, discussed above.

Image Curriculum

Milestone: Almost complete alpha version of Image curriculum

With support of an NSF iUSE grant as part of the DIVAS project at Doane (https://www.doane.edu/divas-project) Dr. Tessa Durham Brooks and Dr. Mark Meysenburg at Doane College have been developing and piloting a workshop on Image analysis. Doane is an undergraduate teaching college and they have developed an intro to Python via image processing that would be attractive to undergraduates in the program. They have developed a workshop introducing Python and image processing skills in a rich and visual way that helps learners connect the way images are stored on a computer to computational concepts in Python. We have been coordinating with their team throughout the development, and the lessons are formatted and taught using Carpentries templates and approaches. They have been developed in Doane repositories, and we are now working together on the process of moving the work over to Data Carpentry repositories for further development. Tessa and Mark are enthusiastic about this transition to bring broader visibility to the curriculum, and scale its impact through broader teaching and a larger group of developers and maintainers, as exists for our other core lessons.

The Doane team have now piloted this curriculum two times in workshops that include undergraduates and graduate students and received very positive feedback, with high value responses to "How much of the workshop contents do I feel I have mastered" and responses to the question "Rate on 1-10. 1: I wish I never came 10: I can't wait to come tomorrow."

The next pilot is planned for May 20, 2019, and for that workshop, Erin Becker will participate and we will conduct our standard assessment, and Erin will evaluate with a qualitative approach during her participation. Identified updates will be made to the lessons after this workshop, and the curriculum finalized for an alpha release. As this curriculum now moves to the beta stage, we will coordinate the teaching of this curriculum in other locations, and incorporate feedback through teaching and hackathons.

These lessons are based on images of growth of plants. We recognize that other image sets may be more broadly applicable, and that there is opportunity for more domain specific images. The curriculum as written is generalizable to allow for future adaptations. Aspects such as spot counting would be easily modifiable to applications in astronomy and chemistry, for instance. The Doane team is also applying for an IUSE renewal to continue to develop and support this curriculum.

One challenge for this set of curriculum is the computing environment. One of the packages, Opencv, is often difficult to install. The first pilot workshops have been taught using virtual environments, but we have found this is extra cognitive overhead for learners and that it then doesn't apply as easily for their own work. We will be exploring other options for the computing

environment for these workshops, including Amazon AWS, which we already use in our [Genomics workshops](#), having students access the lessons through a web environment, such as using MyBinder or JupyterHub, or potentially using Docker.

3. An outline and set of milestones for the development of a curriculum in Chemistry

While there is interest in this curriculum, we have not yet begun work or community development for the Chemistry curriculum. All of the general infrastructure and community work that we have done will allow us to more easily work with the community to develop a roadmap for this curriculum in year 2.

Milestone: Initiated training of instructors with backgrounds in economics and chemistry

In the past year, we have badged 16 Instructors with expertise in Economics, with another 33 currently going through the certification process. In Chemistry, we have 20 badged Instructors, with another 23 currently certifying. As these lessons come closer to completion, we will increase our recruitment efforts to build instructional capacity in the relevant domains.

Budget changes

As outlined above, much of year 1 work went towards the development of general infrastructure, guidelines, community and 'how to' guides. This lays the groundwork for the release and piloting of Economics and Image curriculum and pathway for Chemistry curriculum in year 2. We realized that spending more time on the infrastructure would significantly improve the process, in that it would scale more effectively and overall take less staff time, and be a more positive experience for the community. Additionally, as described, we updated our alpha development approach, moving to having just a few people work on the alpha version and using the hackathons for refinement and teaching. We realized that this infrastructure work was more than time for one person. Therefore, we reallocated some of the budget from hackathons to Dr. Erin Becker's time on the Curriculum Development Handbook and Maintainer and Curriculum Advisory Committee guidelines and framework. We therefore overspent on staff salary, adding part of Dr. Becker's salary, and underspent on hackathons. Dr. François Michonneau's salary is also higher than planned, because in our original proposal we had a different fiscal sponsor with different fringe overhead. The remaining hackathon funds will be spent in year 2 and the hackathon events conducted in year 2, to move the lessons from alpha to beta stage and conduct pilot workshops.