

## Lessons for Librarians in Open Science Principles and Methods

The University of California, Los Angeles (UCLA) Library seeks funding from the Laura Bush 21st Century Librarian Program for an Implementation Project supporting the Program's Goal 3, Objective 3.4. Specifically, this project would use a peer review method to solicit and fund modular lesson creation in training librarians to support researchers in open science methods and digital processes. The UCLA Library seeks \$237,839 in funding and offers \$46,598 in cost-share to establish and coordinate a lesson development process for 14 open science lessons for librarians. This process would include 1) creating a review committee of researchers, librarians, and data science professionals, 2) developing and issuing a call for lesson proposals, 3) selection of lessons, and 4) two virtual summer schools to support selected lesson designers in curricular and modular lesson development life cycle. The result will be 14 open source and modular lessons supporting continued skills development for librarians in open science to effectively provide instruction to new researchers and serve as effective collaborators with faculty and researchers in science projects using open science practices. Scholarly and research practices change rapidly, and library and information professionals are vital members of the research activities at universities and research institutes. Librarians who can advocate for and support open science will enable researchers and their data to reach a wider community. Though librarians play an essential role in promoting open science, few systematic and pragmatic open-source courses are tailored to train librarians as practitioners of open science. This project will develop, implement, and refine a reusable curriculum to ensure library and information professionals have the skills to participate in an open science engaged research lifecycle.

### Project Justification

Open science is the movement to make scientific research, data, and their dissemination available to any member of an inquiring society, from professionals to citizens. Open science goals are to bolster scientific research, in part, via reproducibility and replicability of findings. Open science is increasingly recognized as a critical approach to improving and sustaining scientific inquiry and research as a new practice. For example, the United Nations Educational, Scientific, and Cultural Organization's (UNESCO) Recommendation on open science was recently unanimously adopted on 23 November 2021 by all 193 UNESCO Member States, including the US.<sup>1</sup> A recent report from the European Union had similar recommendations for coordinated training on open science and integration at the university level.<sup>2</sup>

In the US, the National Academies of Sciences, Engineering, and Medicine (NASEM) formed a Roundtable on Aligning Incentives for open science in 2019 and recently published a toolkit for fostering open science.<sup>3</sup> NASA recently announced an initiative to transform to open science (TOPS)<sup>4</sup> and has

---

<sup>1</sup> <https://en.unesco.org/science-sustainable-future/open-science/recommendation>

<sup>2</sup> <https://op.europa.eu/en/publication-detail/-/publication/af7f7807-6ce1-11eb-aeb5-01aa75ed71a1/>

<sup>3</sup> <https://www.nationalacademies.org/our-work/roundtable-on-aligning-incentives-for-open-science>

<sup>4</sup> <https://github.com/nasa/Transform-to-Open-Science>

designated 2023 as the year of open science (YOOS). Planning for an open science training curriculum and programs is already underway as part of this initiative. American Geophysical Union (AGU) and a broad consortium of organizations work with NASA to reach the most comprehensive community of researchers. Libraries are an essential aspect of the NASA TOPS program to reach and include research communities from various backgrounds. Should this proposal be successful, it would complement and directly support these programs. Indeed, NASA TOPS members are connected to the library open science training program proposed here and intend to link the programs in various ways. NASA TOPS members also have library backgrounds, networks, and perspectives that will aid the program.

Despite the recent recognition of a need for open science, the practices of open science remain highly fragmented and awareness low.<sup>5</sup> An initial review by NASA reached a similar conclusion, some of which is reflected via the current NASA TOPS public website.<sup>6</sup> There is a pronounced need for explicitly targeting training at early career researchers. Funder mandates and policy decisions do not always translate to awareness or compliance in the lab. Advocacy for these principles and practices is critical, and that is where this program leads and contributes to open science training at the campus support level.

Librarians and information professionals serve as essential educators and collaborators within the research life cycle. They will be critical to ensuring the practice of open science is promoted and supported. However, there is a dearth of understanding in the profession, and librarian training in open science is insufficiently addressed from the formal library school curriculum, and integrated, targeted, and reusable curricula do not exist<sup>7</sup>. While recent efforts in library training programs have focused on research data management<sup>8</sup>, data science<sup>9</sup>, and even looked at the intersection of scholarly communication and information literacy<sup>10</sup>, a holistic approach to open science training is needed to formalize the work started by efforts such as Foster Open Science<sup>11</sup>, and to respond to national<sup>12</sup> and international<sup>13</sup> calls.

Furthermore, by developing a common and familiar train-the-trainer approach, libraries can turn STEM, data, or public librarians into open science advocates who can teach best practices to early career researchers and introduce methods to practitioners. As mentioned above, clear guidance and recommendations from international, national, and institutional organizations on open science have been lacking to this point. However, standards and practices initiatives are starting to emerge. Aligning training around these initiatives in library workforce development is needed. Ambiguous funder

---

<sup>5</sup> <https://datascience.codata.org/article/10.5334/dsj-2022-002/>

<sup>6</sup> <https://github.com/nasa/Transform-to-Open-Science>

<sup>7</sup> <https://osf.io/uycax/>

<sup>8</sup> <https://rdmla.github.io/>

<sup>9</sup> <https://d-scholarship.pitt.edu/33891/>

<sup>10</sup> <https://www.ala.org/acrl/standards/ilframework>

<sup>11</sup> <https://www.fosteropenscience.eu/>

<sup>12</sup> <https://www.nap.edu/catalog/26308/developing-a-toolkit-for-fostering-open-science-practices-proceedings-of>

<sup>13</sup> <https://en.unesco.org/science-sustainable-future/open-science/recommendation>

mandates and scholarly community advocacy must be coupled with established support mechanisms. This program will enlighten the conversation with practitioner training via one of the best-known continuing education (CE) programs in libraries today, the Carpentries.

To address this gap in librarianship educational opportunities, we propose developing an open science lesson program using an evidence-based design process employed by the Carpentries<sup>14 15 16 17</sup>, a global teaching community with over 2,800 volunteer instructors that have delivered 2,700 workshops in 71 different countries since 2012. This design process foregrounds learner objectives and focuses lesson design on authentic tasks that avoid cognitive overload that often impedes learning.

This project will work with experts in information, data, and open science to develop modular curricula that libraries and MLIS programs can adapt and re-use to teach open science principles to librarians in workshops throughout the country. This proposal will complement the curricula created and offered through the Library Carpentry lesson program of The Carpentries organization and constitute its own lesson collection. It will also employ a collaborative lesson development model that will enable a community of contributors to engage actively and improve on the materials over time.<sup>18</sup> These lessons will be organized into a lesson program, such as Data Carpentry for Ecology<sup>19</sup>, that cohere as a connected set of modular episodes with overarching objectives and a distinct librarian-focused audience.

This design process leads to more flexible and modular lessons, allowing the curriculum to be responsive as international and national standards around open science emerge and mature. In the early stages of these developments, academic support institutions must cultivate targeted awareness and common support structures, including human resources. The lesson program will align with this work through committee members participating/developing both. A generation of librarians advocating for these principles, at this time when major STEM funding agencies like the NIH and NSF are advocating for open, are poised to make a generational shift in support and thinking for open science.

One key lesson learned from The Carpentries, particularly Library Carpentry, is that the contextual and cultural aspects of open science, including FAIR data and software workflows and best practices, are not easily covered in the workshops and lessons that are taught using a live coding pedagogy to promote hands-on data and coding skills. For instance, librarians can learn a tool and approach, but without overarching context regarding how it fits within the research lifecycle and aligns with library services, it is difficult for librarians to understand the impact of the training and the role of open science advocacy in their outreach and liaising work. These curricula, and the underlying lesson modules, will emphasize the contextual aspects of the open science domain beyond a tools-focused approach, providing

---

<sup>14</sup> <https://cdh.carpentries.org/>

<sup>15</sup> <https://teachtogether.tech/en/index.html>

<sup>16</sup> <https://f1000research.com/documents/9-1377>

<sup>17</sup> <https://carpentries.github.io/lesson-development-training/>

<sup>18</sup> <https://doi.org/10.1371/journal.pcbi.1005963>

<sup>19</sup> <https://datacarpentry.org/ecology-workshop/>

librarians with open science, research ecosystem-wide view, and helping librarians better understand how they can contribute and connect to active initiatives, e.g., NASA TOPS. The curriculum will also be framed by work at multiple levels (community, national, and international) that aims to advance research culture towards open science, aligning the libraries' efforts to this broad array of actions. The curricula will provide a coordinated measure and clear pathways for libraries in open science when high-level reports are calling for it, e.g., EU<sup>20</sup>, National Academies<sup>21</sup>, and UNESCO<sup>22</sup>.

In addition to the continuing education aspect of the program, the outcome will be reflected as open science practices and principles gain wider adoption across STEM professionals and labs. As transformative agreements work to solidify the publishing industries' stranglehold on academic discourse, the advocates for open access to data and software are gaining ground. With funding agencies and scholarly communities advocating for open principles, it is incumbent upon the librarian profession to provide services and tools to our patrons. We believe this lesson program will be most directly applicable to academic librarians engaged in supporting confirmatory science and STEM fields and for librarians promoting research transparency practices in their work and liaising. We further feel this program has a role in promoting available and sound science in public-facing libraries supporting the citizen science movement by providing access to well described and organized open science resources.<sup>23</sup> <sup>24</sup> Finally, we envision modules from this educational plan to be formally and informally adopted by library and information science programs in the US. This proposal provides a model for a robust educational program for scientists, career information professionals and the next generation of LIS graduates.

## Project Work Plan

Our work will establish a review committee of nine members to develop program goals, curriculum limits, call for proposals, and serve as a review board to select lesson proposals. The following individuals or entities have already volunteered to serve on the review committee should this project receive funding: Chris Erdmann (Assistant Director, Data Stewardship, American Geophysical Union, RDA Libraries for Research Data co-chair, Library Carpentry Advisory Group), Elaine Westbrooks (Carl A. Kroch University Librarian, Cornell University Library), Caroline Coward (Information Science Manager and Library Group Supervisor, Jet Propulsion Lab, National Aeronautics and Space Administration), Matt Mayernik (Project Scientist and Research Data Services Specialist, National Center for Atmospheric Research Library), Jane Greenberg (Alice B. Kroeger Professor of Information Science, Drexel University), a representative from the NASA TOPS program, Jason Williams (Assistant Director, External

---

<sup>20</sup> <https://op.europa.eu/en/publication-detail/-/publication/af7f7807-6ce1-11eb-aeb5-01aa75ed71a1/language-en/format-PDF/source-190694287>

<sup>21</sup> <https://www.nap.edu/catalog/26308/developing-a-toolkit-for-fostering-open-science-practices-proceedings-of>

<sup>22</sup> <https://en.unesco.org/science-sustainable-future/open-science/recommendation>

<sup>23</sup> [https://www.academia.edu/64817910/Citizen\\_Science\\_in\\_Libraries\\_Results\\_and\\_Insights\\_from\\_a\\_Unique\\_NAS\\_A\\_Collaboration](https://www.academia.edu/64817910/Citizen_Science_in_Libraries_Results_and_Insights_from_a_Unique_NAS_A_Collaboration)

<sup>24</sup> <https://doi.org/10.1016/j.lisr.2021.101090>

Collaborations, Cold Spring Harbor Laboratory, DNA Learning Center), Cody Hennessey, (Journalism & Digital Media Librarian, University of Minnesota/Library Carpentry Advisory Group), and UCLA iSchool faculty (to be identified). Ex Officio members, individuals of the community known to the project core team, will be invited to provide different perspectives and diverse backgrounds (see Supportingdoc3 for a complete list of participants and affiliations).

Tim Dennis, UCLA Library Data Science Center, will be the Director for this project. He is an internationally recognized instructor in foundational computational and data science skills. He is also actively involved in teaching and promoting open science through membership in the Berkeley Initiative for Transparency in the Social Sciences (BITSS)<sup>25</sup> and the Center for Open Science (COS). He was on the teaching faculty for three years at BITSS' Research Transparency and Reproducibility Training<sup>26</sup>. He also has extensive experience in providing oversight, vision, and leadership for Library Carpentry lessons as a Curriculum Advisor and Advisory Board member. He trains potential instructors as part of the Carpentries. Dr. Zhiyuan Yao, Spatial Data Science Librarian, UCLA Library Data Science Center, will serve as the Coordinator for this proposal. Zhiyuan has actively participated in activities, events, and conferences to promote open science. She taught open sources Carpentries workshops quarterly, is Carpentries community facilitator<sup>27</sup>, was one of the core committees of University of California (UC) GIS Week<sup>28</sup>, UC Libraries Forum<sup>29</sup>, and UC Love Data Week<sup>30</sup> and attended conferences that promoted open science, such as North American QGIS conference, Free and Open Source Software for Geospatial (FOSSG) conference, and Research Reproducibility conference. UC GIS Week has been held in 2020 and 2021. Each year, over eight hundred participants from ten Universities of California attended over eighty presentations covering ten topics. UC Libraries Forum had over four hundred registrations, 300 attendees coming for sixty proposals. It was a UC systemwide collaboration and an annual celebration and professionalization opportunity for librarians and library staff. UC Love Data Week was also a UC-wide collaboration with over eight hundred attendees from worldwide to join the data-related workshops. In addition to teaching and organizing conferences, Zhiyuan will bring her experience in project management, data management, and programming to execute the proposed project.

## Roles

The project team will include the Project Director, Coordinator, Lesson Infrastructure Technology Developer (LITD), and Review Committee. The Project Director and Coordinator will coordinate curriculum development to encourage inclusion within The Carpentries lesson development methodology and enable beta-teaching workshops to test lessons with potential learners. The Project Director will design the Summer Seminar curriculum development workshop. The Coordinator will

---

<sup>25</sup> <https://www.bitss.org/>

<sup>26</sup> <https://www.bitss.org/education/rt2/>

<sup>27</sup> Community Facilitators in The Carpentries are people who are empowered to advocate for others in the community and guide community members as they navigate The Carpentries global community.

<sup>28</sup> <https://uc-gis-ucop.hub.arcgis.com/pages/gis-week>

<sup>29</sup> <https://wiki.library.ucsf.edu/display/UCLF>

<sup>30</sup> <https://uc-love-data-week.github.io/>

develop a call for proposals based on guidance from the review committee and the Program Director and run a competitive proposal process each year. The Coordinator will also be responsible for planning, organizing, and hosting the Kick-Off Meeting and Summer Seminar. The LITD will develop the repository for storing lessons and set up templates to modularize the lessons. This role will also provide technical support to authors. This role will also identify lesson integration platforms (learning management systems, website, external partner's content management systems). The Review Committee will develop clear guidelines and review and select lesson proposals to fund. These guidelines will ensure funded projects are well-conceived, budgeted, and achievable.

## Work plan stages

Project initiation (Project team established, members recruited, and review committee established, developed, and published a call for lesson modules) (August - October 2022)

Round 1 - Lesson proposals received and selected for Round 1 (November 2022 - March 2023)

Round 1 - Virtual Summer Institute planning (April - June 2023)

Round 1 - Virtual Summer Institute (July 2023)

Round 1 - Lesson Development and Instruction piloting (August 2023 - March 2024)

Round 2 - Module proposals received and selected for Round 2 (November 2023 - March 2024)

Round 2 - Virtual Summer Institute planning (April - May 2024)

Round 2 - Virtual Summer Institute (June 2024)

Round 2 - Lesson Development and Instruction piloting (July 2024)

## Project initiation

### *August - October 2022*

Initial tasks for the Coordinator and Director include formulating a project charter and confirming members for the Review Committee. The Coordinator and Director will also hire the LITD to help us plan and implement the project's technical infrastructure. This phase will include selecting tools that will aid in our collaboration and mirror the iterative, reflective process we want to foster in the overall project.

With Review Committee members, Coordinator, Director, and LITD confirmed, we will organize a kickoff meeting to initiate the project. This meeting will finish crafting a call for lesson proposal and identify networks we want to target for our appeal (i.e., the application website/form disseminating to many channels, lists, social media). We have already identified distribution networks where project members and participants are active (see Supportingdoc4 for a list of these networks). Other networks will be identified at the kickoff. We will also define the essential elements of a lesson proposal, utilizing norms established by successful programs like the Carpentries. The Coordinator, with feedback and input from the Review Committee and Project Director, will create rubrics and criteria, drawing on resources and

material from The Carpentries, Turing Way<sup>31</sup>, ADDIE<sup>32</sup>, CodeRefinery<sup>33</sup>, and OpenScapes<sup>34</sup>, for assessing and selecting winning lessons and developing any artifacts and processes needed to ensure a fair and transparent review process.

## Round 1

### Lesson proposals received and selected

#### *November - March 2023*

The Coordinator will receive lesson proposals and share them with the Review Committee. The Review Committee will conduct initial reviews using the rubric and criteria articulated in the Project Initiation above. In addition to the criteria mentioned in the previous stage, the project team will work with the application and review each proposal for viability and feasibility to ensure each funded application has set realistic and achievable goals and a plan for completion. The Coordinator will then call a meeting of the Review Committee to discuss merits and feedback on lesson proposals. Modules with the highest scores from the Review Committee ranking will be selected as the lesson modules for the first round. The Coordinator will notify the authors of their successful submission. Confirmation of participation by lesson developers will also happen in this round.

### Virtual Summer Institute planning

#### *April - June 2023*

The Coordinator, Director, and LITD will plan for the Virtual Summer Institute with input and feedback from the Review Committee. This planning will involve adapting and modifying the Carpentries Collaborative Lesson Development Training materials for the seminar. The Coordinator will work with the LITD to set up a workspace and instantiate lesson templates for selected lessons. The workspace will be in a distributed version control system like GitHub or GitLab. We will also define other collaboration tools we will need to work with during this phase. We will establish best practices to ground norms of collaboration and communication during this phase as well<sup>35</sup>. Finally, we will plan the timing, structure, and content of the Virtual Summer Institute.

### Virtual Summer Institute

#### *July 2023*

This Virtual Seminar will be organized as a curriculum development school where lesson authors participate in a four-week program to support lesson design. Each week will include three to four hours of interactive sessions that provide authors with background information and guidance on writing and

---

<sup>31</sup> <https://the-turing-way.netlify.app/welcome>

<sup>32</sup> [https://en.wikipedia.org/wiki/ADDIE\\_Model](https://en.wikipedia.org/wiki/ADDIE_Model)

<sup>33</sup> <https://coderefinery.org/>

<sup>34</sup> <https://www.openscapes.org/>

<sup>35</sup> <https://doi.org/10.1371/journal.pcbi.1005963>

improving their lessons. The training, designed and delivered by the Director and Coordinator, and modified from existing curricula<sup>36</sup>, will introduce the steps that authors need to take to design and develop a lesson that meets their target audience's goals. The seminar will also incorporate exercises and give feedback to facilitate lesson development over the month of the seminar. The Director has extensive experience in lesson development, has led global lesson development sprints, is a member of the Curriculum Advisory Committee on Library Carpentries, and is a published lesson author on six Carpentries lessons. The Coordinator is a Carpentries instructor and has extensive experience developing lessons as well.

Components of the lesson development school will include:

- Identifying and designing for the target audience
- Defining lesson objectives
- Building a lesson website in our infrastructure (GitHub or GitLab)
- Designing formative assessments

We will structure a period for the lessons to be piloted to provide further feedback following the seminar.

### Lesson Development and Instruction piloting

#### *August 2023 - March 2024*

After the summer institute, we will provide venues for beta-testing lesson modules in the larger library community. The Coordinator and Director will organize a Fall 2023 workshop series including the lesson modules offered to UCLA iSchool students and LA area Librarians. In Winter 2024, we will offer a select set of the curricula as part of the UC Love Data Week 2024. We will collect feedback from learners and share actionable lesson improvements with module developers throughout. This phase will also prepare the lessons available in multiple venues, including The Carpentries, NASA TOPS, and Open Science MOOC. We highlight these three but will draw on our network, which we listed in Supportingdoc4, to refine and broaden our outreach in the US. Additionally, the committee and Ex Officio group are well situated to promote the lessons internationally (EU via Ligue des Bibliothèques Européennes de Recherche – Association of European Research Libraries (LIBER)<sup>37</sup>, Australia via Australian Research Data Commons (ARDC)<sup>38</sup>, Africa via AfricArXiv<sup>39</sup>, and The Carpentries<sup>40</sup>). The goal is to make this course program more broadly available. For the Carpentries, the Project Director and Coordinator will facilitate adding the lessons to The Carpentries Incubator<sup>41</sup>. The Carpentries Incubator is a repository for lessons in whatever state of development and is intended to increase visibility and community during lesson

---

<sup>36</sup> <https://carpentries.github.io/lesson-development-training/>

<sup>37</sup> <https://libereurope.eu/>

<sup>38</sup> <https://ardc.edu.au/>

<sup>39</sup> <https://info.africarxiv.org/>

<sup>40</sup> <https://carpentries.org/>

<sup>41</sup> <https://github.com/carpentries-incubator/proposals#readme>



development. The lessons in the Carpentries Incubator will follow a defined lesson lifecycle to ensure “they are sufficiently documented to be teachable by instructors outside of the initial author group”<sup>42</sup>. Once incubated, the lessons will be included into the Library Carpentries curricula.

## Round 2

Round 2 will repeat Round 1 with similar activities described above. Because lesson development is intensive, we want to keep the cohorts small for each round. A small cohort will enable us to provide better support, guidance, and feedback during the design process.

## Diversity Plan

Attention to a diversity of backgrounds will be threaded throughout the roles involved in the project, from module creators to program committee members and collaborators. A distinguishing component of our funding request is that lesson developers will be compensated for their contributions, which assists us in being able to recruit a broader range of lesson developers who might be graduate students, postdoctoral researchers, or early-career faculty. The development of inclusive mechanisms for participating in the project and the curriculum (e.g., code of conduct) will be at the project's core and central to all outreach activities. In addition, criteria for the module material will include diversity, equity, and inclusion components. UCLA and partnering institutions commit to diversity and inclusion, and this project will leverage and involve the people and resources from the organizations in this project.

## Project Results

This project will result in fourteen open science course modules (see a list of proposed course modules in Supportingdocs1) for librarians at any stage of their career. These curricula will be published online in an unrestricted and freely available manner for reuse and remixing. These lesson modules will be developed and supported through two virtual summer institutes, and feedback from lesson piloting will be used to test the lessons with librarian learners.

They will address a missing set of curricula that will provide a holistic overview of important components of open science and how to engage in larger international initiatives leading on open science practices and standards. Because the lessons will be designed in a modular and flexible way and available in open formats, they will be embeddable, informally or formally, in numerous venues. We will facilitate the submission of these lessons through the Carpentries Incubator<sup>43</sup>, a repository to share Carpentries-style teaching materials at all stages of development, incorporate them into the Library Carpentry curriculum, and promote them institutionally.

---

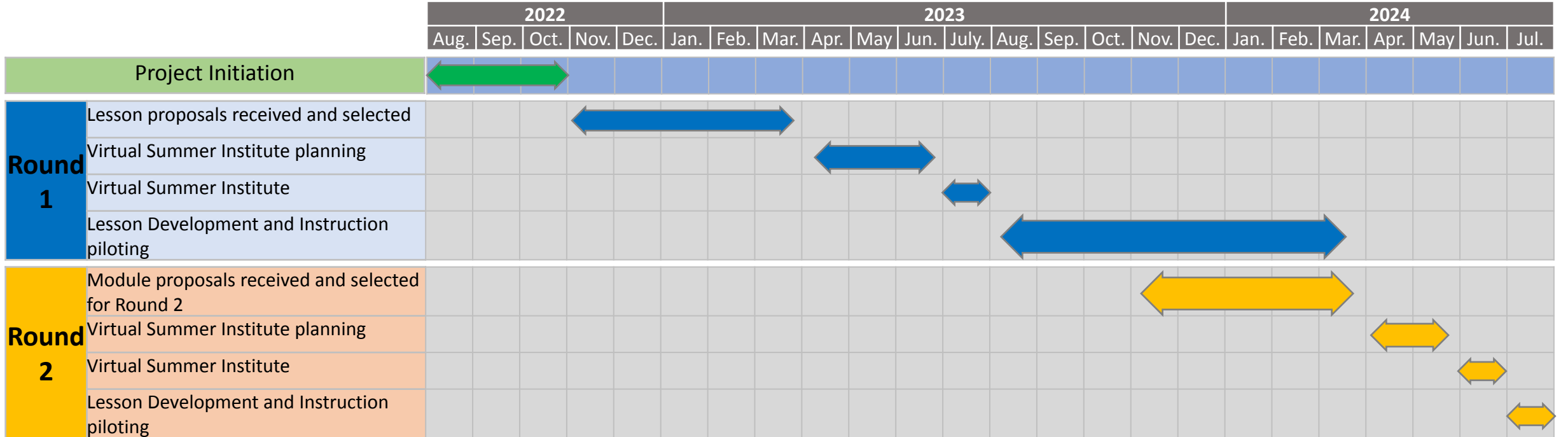
<sup>42</sup> <https://cdh.carpentries.org/the-lesson-life-cycle.html>

<sup>43</sup> <https://github.com/carpentries-incubator/proposals#readme>

These lessons will be featured in the University of California's multicampus events (Love Data Week, GIS Day, Libraries Forum) and the AGU at their annual event for immediate visibility to hundreds of data support professionals and thousands of AGU member scientists. Furthermore, the distribution through the Carpentries network will ensure the lessons have dynamic growth potential and impact open scientific discourse for years to come.

## Schedule of Completion

Project Title: Lessons for Librarians in Open Science Principles and Methods



## Digital Products Plan

- **Type:** What types of digital products will you create?

The products will be 14 open source and modular lessons supporting continued skills development for librarians in open science to effectively provide instruction to new researchers and serve as effective collaborators with faculty and researchers in science projects using open science practices.

- **Availability:** How will you make your digital products openly available (as appropriate)?

The lessons will be established by adopting Carpentries lesson design method and be published as webpages and maintained on Data Science Center's github which is accessible to all users.

- **Access:** What rights will you assert over your digital products, and what limitations, if any, will you place on their use? Will your products implicate privacy concerns or cultural sensitivities, and if so, how will you address them?

These lessons and related materials are made available under the Creative Commons 4.0<sup>1</sup> license. The lessons are free to use, but require attribution. Users can share and adapt the lessons for any purpose, even commercial purposes. These lessons will not implicate privacy concerns or cultural sensitivities.

- **Sustainability:** How will you ensure the sustainability of your digital products?

The Data Science Center is a dynamic and evolving part of UCLA Library. We have librarians, Carpentries facilitators, and data scientists who can update, maintain, and archive the lessons periodically. DSC has an active github account where the lessons will be published and made accessible to worldwide users. We also manage an online repository, Dataverse, at UCLA which we use to archive the lessons if needed.

---

<sup>1</sup> Creative Commons (CC) licenses are public licenses. You can use them to indicate what other people are allowed to do with your work.

## Organizational Profile

The Data Science Center's mission is to foster a welcoming research community by developing data literacy and foundational coding skills through consulting and education. We support researchers throughout the full data life cycle of activities: planning research projects and data management plans; data acquisition, storage, cleaning, and usage; analysis; publication; curation; and preservation<sup>1</sup>. Established in 2018, Data Science Center works as a part of an effort in UCLA Library to expand access to its expertise in data management and curation while modernizing its services to support current research innovation.

The Data Science Center (DSC) is part of the UCLA Library. It supports the UCLA Library's mission to empower and inspire communities of scholars and learners to discover, access, create, share, and preserve knowledge. Inside the UCLA Library, the DSC is part of the Digital Initiatives and Information Technology Department. It actively fulfills the team's mission to provide sustainable, scalable, and unique opportunities to create new knowledge.

The DSC primarily serves UCLA's community of approximately 44,900 registered students. For undergraduate students, underrepresented groups make up 34% of all admitted California-resident first-year students for 2021-2022 — the highest proportion at UCLA in over 30 years. In terms of the socioeconomic diversity of prospective freshmen from California, with admissions of 32% first-generation students and admissions of 36% of those from low-income families<sup>2</sup>. In addition to the UCLA community, the DSC's workshops, tutorials, and data management service are open educational resources available to all learners and have a state-wide audience (around 50% attendees of workshops and over 80% of attendees are from other Universities of California).

### Background information

Created in 1977 as part of the UCLA Survey Research Center, the Social Science Data Archive (SSDA) maintained and preserved data produced by and purchased for UCLA researchers<sup>3</sup>. By the 2010s, the SSDA's service profile of supporting survey research data collection, management, and preservation had become too narrow for researchers' emerging needs. Given the complexity, scale, variety of sources that generate data, and the advent of data science, the Social Science Archive joined the UCLA Library, redesigned its missions and goals, and renamed the Data Science Center in 2018. With six full-time staff, the DSC promotes data literacy through educational instruction (including programming skills, data management, etc.), consulting, project participation and advocating Open Science via promoting good data management practices, advertising research transparency initiatives, and sharing resources on open source platforms (e.g. github).

---

<sup>1</sup> <https://www.library.ucla.edu/about-0>, last revised 2019-02-10

<sup>2</sup> <https://newsroom.ucla.edu/releases/ucla-admits-most-accomplished-class-2021>

<sup>3</sup> Some early studies include the Los Angeles Metropolitan Area Surveys 1-10 (1970s), Southern California Social Survey (1980s) and Los Angeles County Social Survey (1990-2002). All of these continue to be available in the Social Science Data Archive collection.