Successful proposal to the Turing Online Training Grant 2020 - Awarded to The Turing Way and Open Life Science

Cite this as: *Sharan, M., Whitaker, K., Yehudi, Y., The Turing Online Training Gran 2020 awarded to The Turing Way and Open Life Science teams. DOI: 10.5281/zenodo.6974060*

APPLICANT DETAILS

Title: Dr

Name Surname: Malvika Sharan

Email (Work) msharan@turing.ac.uk

Please state the names of other training collaborators, roles and organisations.

Collaborators

- Kirstie Whitaker: The Alan Turing Institute, London, UK (ORCID: https://orcid.org/0000-0001-8498-4059)

- Yo Yehudi: Open Life Science Ltd, UK (ORCID: https://orcid.org/0000-0003-2705-1724)

Training proposal

If you selected other, please provide further details here.

Not applicable

Proposed number of total participants if your proposal falls under the Dynamic Learning track.

100 participants in 2 cohorts of OLS in 2021 including members from Turing and its partner organisation

Please state the key training objectives and expected learning outcomes of the proposed activity or resource.

Background:

The Turing Way is an open source book project that involves and supports a diverse research community in ensuring that reproducible and ethical data science is accessible and comprehensible for everyone. This project is dedicated to fostering gold-standard reproducible

research and promoting a culture of collaboration that can transform data science worldwide. Started less than 2 years ago in The Alan Turing Institute, it has become a global community of researchers, engineers, educators, policymakers, data librarians, industry professionals and experts from various research domains, who share and promote research best-practices. As the national institute for data science and artificial intelligence, different stakeholders of Turing, specifically, its researchers and engineers, are encouraged to share their recommendations in data science through The Turing Way. However, they also need to be trained and mentored to share their skills and knowledge with researchers at the national and international levels. To deliver such an opportunity for training, we have collaborated with Open Life Science (OLS).

Objectives:

OLS is an online cohort-based program that supports individuals working in research organisations to learn about open research practices, apply them in their work, develop community-oriented resources and share them openly. OLS's goal is to enable individuals and teams in research institutes and organisations to lead open research projects and become Open Science ambassadors. This training program is highly aligned with the mission of the Tools, Practices, and Systems Research Programme that aims to democratise access to information and the human-first approach for research and development. One of the main objectives is to help our researchers develop understanding of open research practices, open leadership and open communication -- skills that are transferable across different domains beyond scientific research. This will ultimately help them to involve contributors from their research fields and organisations and ensure validity and sustainability of their project in the long term. In this program, experienced open science practitioners will be invited to join the mentors group.

Learning outcomes:

Through this training program, we will support members from Turing and its partner organisations to participate in the OLS program in 2021. Participants will join a 16-week long mentoring and cohort-based training program to achieve the following learning outcomes:
- gain knowledge essential to create, lead, and sustain an Open Science project
- connect with members across different projects, communities, backgrounds, and identities from different organisations and research areas
- empower each other to become effective Open Science ambassadors in their communities
- share their expertise with each other through their work and guidance developed within The Turing Way

Particularly, the project leads (mentees) will:
- receive one-to-one support in their projects from their mentors, who will guide them to develop a clear vision of their projects
- create a roadmap by defining their project development plan and community engagement guidelines
- set up an online repository (if they do not have one already) to start building their project openly and add an appropriate open license to their projects
- gain an understanding of what open research approaches and best practices exist and what tools they can apply in their research, organisation and community

- potentially launch their projects (if the project is new) and involve other members to promote open practices in the data science community
- graduate from the Open Life Science program by presenting their work, progress, and learnings over the 16 week and sharing plans for the long-term sustainability of their projects

In parallel, the mentors will:

- receive training on mentorship from professional instructors
- exchange ideas and tips with other mentors and organisers
- learn how to guide people through an open project
- learn how to give feedback and provide support to early stage researchers

Please outline a preliminary programme of activities; high-level summaries would be acceptable, if you do not have a detailed schedule.

Outline of course content:

The full syllabus and schedule are available online at https://openlifesci.org/ols-2/schedule/

Week	Meeting type	Length	Topic	Agenda
Week 01	Mentor-Mentee	30 min	**Meet your mentor!**	Meet each other and discuss your personal motivation, expectations, working practices and project goals
	Mentors	90 min	**Mentor training**	
Week 02	Cohort	90 min	**Welcome to Open Life Science!**	Meet other members of your cohort, Share project vision, Intro to working openly (open canvas)
Week 03	Mentor-Mentee	30 min	**Meet your mentor!**	Discuss assignments from the cohort call & concrete implementations
	Cohort	90 min	**Social Call & Coworking session**	
	Mentors	180 min	**Mentor training**	
Week 04	Cohort	90 min	**Tooling and roadmapping for Open projects**	Working with GitHub as a community hub: Markdown as a tool to make websites, Licence, Goals and Roadmap, Contributors,

				Code of Conduct
Week 05	Mentor-Mentee	30 min	**Meet your mentor!**	
	Cohort	60 min	**GitHub Tutorial**	Skill-up session/workshop
Week 06	Cohort	90 min	**Open Science I: Project Development**	Developing Open Projects: Iterative and agile Project management, Open -Source, Software, Hardware, Data for reproducibility
Week 07	Mentor-Mentee	30 min	**Meet your mentor!**	
	Mentors	60 min	**Mentor training**	
	Coworking	60 min	**Coworking session**	
Week 08	Cohort	90 min	**Open Science II: Knowledge Dissemination**	Sharing Open Project: Preprint publications, DOI and citation, Open protocols, Open Education & Training for transparency
Week 09	Mentor-Mentee	30 min	**Meet your mentor!**	
	Cohort	60 min	**Personal Ecology, self care and Ally Skills**	Skill-up session/workshop for understanding personal biases and removing them from their projects
Week 10	Cohort	90 min	**Open Science III: Next steps**	FAIRification of existing or mature projects to maintain reproducibility
Week 11	Mentor-Mentee	30 min	**Meet your mentor!**	
	Coworking	60 min	**Coworking session**	
	Mentors	60 min	**Mentor training**	
Week 12	Cohort	90 min	**Career Guidance Call**	Healthy Leadership
Week 13	Mentor-Mentee	30 min	**Meet your mentor!**	
	Cohort	90 min	**Creative writing and presentation skills**	Skill-up session/workshop
Week 14	Cohort	90 min	**Designing &**	Personas and pathways for

			Empowering for inclusivity	contributors, Implicit bias & mental health care, Community interactions & Ally-skill
Week 15	Mentor-Mentee	30 min	**Meet your mentor!**	Preparation for the final demos
	Cohort	90 min	**Final presentation rehearsal**	Test of the final demos
Week 16	Cohort	90 min	**Final presentations & Graduation!**	5-minute demos of projects (Audience: entire community & public, Open and recorded call)
	Mentors	60 min	**Mentor wrap up**	

How do you expect this activity or resource to be delivered?

This training program will run online twice in 2021, each will be delivered over the course of 16 weeks. Participants will join the program with their research idea that they propose to work on or develop during the program. They will receive training on different topics in Open Science, reproducibility, collaboration and leadership over the online cohort (training) calls. These topics will include methods such as open data, open hardware, open-source software, citizen science, and technical training. Additionally, they will learn about open science dissemination methods such as preprints, preregistration, open peer review, open access scientific publications, and archiving of research components. They will also be introduced to hands-on skills such as collaboration on GitHub, open communication of their work and community building that encourages bringing in new contributors from their local communities, learning about ally-skills and creating inclusive pathways for them to engage with the project in the long term.

Mentors and mentees will meet in alternate weeks, creating a rhythm of cohort call one week, and mentor call the other. Mentors also receive training to deliver their mentorship effectively and are supported in their work by the program organisers. Training on each principle is complemented by assignments that help participants apply the learned skills to their projects with guidance from their mentor in mentor-mentee calls, and other participants during coworking sessions. The training format allows our participants to systematically implement “open by design” principles in developing and leading collaborative and community-oriented projects.

Please explain how this training activity or resource will align with the Turing’s equality, diversity and inclusion principles**.**

The Turing Way and OLS are established as globally inclusive initiatives and community-oriented data science and capacity-building projects for researchers from diverse backgrounds and identities.

The Turing Way involves its members from around the world in developing, maintaining and sharing its resources. We have over 200 international contributors who have together co-authored more than 100 subchapters and community resources that are read and used by thousands of readers worldwide. We facilitate mentored contributions to the project via GitHub and through virtual coworking calls every week. We have a detailed Code of Conduct, contribution guideline and pathways for engagement.

Similarly, with the “open by design” principle at its core, the OLS training modules teach how to develop and share research work openly, and thoughtfully build inclusive projects. Notably, we accept projects with both technical and non-technical focuses (such as community-oriented projects or events) in data science to bring diversity of skills in each cohort. In addition to technical knowledge on reproducibility and open research, our participants learn about inclusion and diversity, implicit bias, ally skills, self-care, and community building. OLS has a welcoming and open channel for community participation -- we use a Slack channel for our cohort participants and have a public facing Gitter channel (details: https://openlifesci.org/ols-2#resources). There are detailed Code of Conduct, participation, and reporting guidelines that apply to all training, resources, and communication platforms that help us establish a welcoming and safe environment for all our participants. We strive to avoid jargon and other non-inclusive language that can alienate, and make underrepresented people feel excluded. We use simple and jargon-free English in our communications, website and training resources. We have provided ways for everyone to share their pronouns, and respect those in our conversations and communication.

In the first round, OLS-1, we supported different projects from Kenya, Netherlands, Brazil, Canada, Thailand, Spain, UK, Japan, Russian, India, USA, Norway, Germany, and Nepal. Our mentors represented research communities from China, Greece, UK, USA, South Africa, Germany, Kenya, Netherlands, and Brazil. The projects they started in OLS-1 ranged from developing Machine Learning software, to establishing networks of bioinformatics learners, creating open source databases, developing a citizen science project and enhancing accessibility of computing resources in the global South. Several of these projects have made substantial impact in their communities, which would have not been possible without the mentorship, time and efforts of the 3 organisers (founders), 22 mentors and 50 experts. Our mentors and experts also gain immensely from the mutual support and exchange among OLS members and appreciate the opportunity to give back to the community.

In the second round, OLS-2, we attracted mentees and mentors from equally diverse interests and backgrounds from both developed and developing countries (blog: https://openlifesci.org/posts/2020/09/01/ols2-announcement/). The demand for our training has extended beyond life science, which is evident from the number and diversity of the topics proposed in OLS-2 from all around the world. We accepted 50% more applications and participants than before and started in September with 31 projects, 48 mentees, 35 mentors, and 67 experts from countries spread across 6 continents. We also have options for our graduates to return as mentors and experts and continue to gain support from OLS even after they have

finished their training. By bringing in successful OLS graduates to become mentors and experts, we avoid gatekeeping and share a sense of community. It also broadens access to Open Science resources in their communities as they continue to lead their projects and support others.

Keeping up with our commitment towards diversity, equity and inclusion within The Turing Way and OLS, participants will be invited from different backgrounds and domain knowledge. They will be mentored and supported in the program based on their project proposals without discriminating by nationality, career-stage, institutional affiliation, gender identity or expression, disability, or sexual orientation. We understand that the passive advocacy of diversity and inclusion is not sufficient, therefore we will actively deliver workshops, public resources and seminars before launching the call for application in order to reach out to researchers who are traditionally underrepresented in open science. We are currently testing different training models to make our cohort calls more accessible to people who require live captioning, people who don't use English as their primary language and those joining calls only with audio (limited equipment set-up, internet access or visual-blindness accessibility requirements). In the next cohorts, we aim to reach out to an even wider community by offering similar workshops, public seminars on Open Science and improving the accessibility aspects in our program.

What approaches will be taken to enable learning and development during the activity or resource, particularly in an online setting?

Participants are trained through the online cohort calls, which feature collaborative group discussions, networking, and a range of speakers from our expert community. In an application-driven format, each online cohort call is supplemented with assignments and references to deepen their understanding in a self-paced manner. Mentees are supported by their mentors in leading their projects. Mentors are matched with their mentees based on their specific requirements of skills, research, preferred languages, and time zones (wherever applicable). Experts in the program are experienced open science practitioners who share their insights from working on open science topics. OLS provides a platform for these inspiring individuals from diverse identities, nationalities, and backgrounds to present relatable examples from their work, discuss challenges, and share resources with our participants.

We strive to encourage and recognize the quietest voices, and not just those with the most confidence, and volume. During the cohort calls, we facilitate breakout discussions and encourage silent note-taking in shared notes. The one-to-one mentor call ensures that everyone feels supported in the program and gets an opportunity to discuss project-specific concerns. Cohort calls are live-transcribed using Otter.ai, and some breakout rooms are conducted in a quiet written format, to ensure that people with varying language skills and/or hard of hearing participants can still get an interactive experience. Mentors and experts will receive support, training and small grants to be supported in their roles. All participants of this program learn "open by design" principles and implement them in their work and share their research widely.

Why is this an important topic for data science researchers at this moment in time? In what ways does the workshop align with the Institute's aims and research challenges**?**

Data science can be most impactful when the research objects are freely shared to enable access to information, re-use of the underlying technology, inclusion of diverse perspectives, and collective advancement of knowledge. It is particularly crucial now when most research projects are being managed online by involving collaborators from different research domains. In the current pandemic situation, individuals in the data science fields who can conduct their work entirely online need to adapt their work to be effective, efficient, and impactful. However, academic institutions rarely teach researchers how they can use open science and reproducibility for tooling and road-mapping of projects, plan reproducible workflows, reuse open data, share research objects, involve others, and lead a truly inclusive project team. As a result, researchers avoid sharing their work openly and wrongly assume that their work will be criticised, scooped, or not acknowledged.

To formally train researchers to work and lead “open” by applying Open Science best practices in their research, the Open Life Science (OLS) project was designed. Under the collaboration name OLS-2 for Turing, in its second round, Open Life Science partnered with The Turing Way, a project within the Tools, Practices, and Systems (TPS) Research Program in The Alan Turing Institute. This partnership has been offering training and mentoring to 6 projects with 12 members from The Alan Turing Institute and The Turing Way (https://github.com/alan-turing-institute/the-turing-way/blob/master/open-life-science-mentoring/README.md), in addition to 38 non-Turing participants from 25 projects from around the world.

We would like to extend the collaboration of OLS with The Alan Turing Institute in 2021 to support the participation of researchers from the institute and its partner organisations. This program strongly aligns with the aim and challenge of institutes to make algorithmic systems fair, transparent, and ethical. This collaboration specifically aligns with the TPS Research Programme: “which seeks to build open source infrastructure that is accessible to all, and to empower a global, decentralised network of people who connect data with domain experts.”

Training conducted by OLS teaches open practices for maintaining transparency in development and reporting, understanding personal biases and designing projects for diverse membership, creating healthy leadership, developing inclusive pathways for engagement at all levels and sharing research objects with an aim to benefit the society. OLS training and mentoring programs will help develop open, reproducible and collaborative research aspects in the projects that participants are either already working on, or want to develop in the near future within the institute and beyond. Mentors and experts will also be invited to join the program where they will be given training for skill-building, the opportunity for networking, and supporting projects in the broader cohort.

If your proposed training is an activity (i.e. Dynamic Learning), would you consider turning the materials in a resource (i.e. Self-Paced Learning)? If so, could you provide details to illustrate how you would do that?

The OLS program is delivered using resources that include training materials, shared notes from the cohort calls, post-cohort assignments and website contents. All the training resources developed and used in OLS are made available online via our website (https://openlifesci.org), GitHub repositories (https://github.com/open-life-science) and Zenodo (https://zenodo.org/communities/openlifesci) under

CC-By 4.0. All our training videos are uploaded on YouTube to share with our participants under CC-BY-SA 4.0 license (https://www.youtube.com/c/OpenLifeSci/videos). we actively encourage our participants to reuse, share, adapt, and remix resources used in the program, which itself is built upon open-source resources.

Even though these materials are specifically intended for our cohort members who can't attend the calls in real-time, these are freely accessible also for people outside the program for self-paced learning. Furthermore, we aim to draft a manuscript capturing the insights and outcome along with our materials, methods, and lessons learned from our cohorts of open science training. This will be made openly available so that others can replicate a program like OLS in different communities and conduct them in a format that is useful for them. The Turing members who participate in OLS will be supported to document their learning outcome and data science recommendations in The Turing Way as a part of their training, which will be widely shared with others in the organisation.

What prior knowledge and skills would the learners require to be able to benefit from this training experience?

Each participant will join the program with a project they will develop individually or in a team using the Open Science principles and best practices. Prior experience with open source will be advantageous for them to progress faster but it will not be mandatory. We will offer webinars before the application period to help interested applicants prepare their applications with a project they intend to develop and lead openly.

If the learners do not already possess this knowledge can you please identify online materials that would adequately prepare them?

This program is designed for both novice learners and experienced open science practitioners. Ideally, these participants will propose a project in a team that they intend to build together as a part of this training and mentoring program. If they don't already possess the knowledge of Open Science, they will be able to start from scratch in the program itself with the training, assignments, and mentoring offered to them.

Research mapping

Please tick the research areas under the heading of 'Social data science' that are most applicable to this event.

Cognitive science

Data science of government & politics Developmental psychology

Ethics for data science & AI

Linguistics

Management science

Networks

Research methods <----------

Social media

Social psychology

Please select from the available list below the priority domain that is more relevant to your training proposal. This could be direct relevance or indirect, through application domain(s).

Healthy world

Engineered world

Natural world

Secure world

Economic world

Social and cultural world <----------

Please select from the available list below the cross-cutting theme that is more relevant to your training proposal. This could be direct relevance or indirect, through application domain(s).

Policy and governance

Tools, practices and systems Safe and ethical AI <----------

Theoretical foundations

END OF THE DOCUMENT