

Scope of Work

Organizational Mycology, Turing Way and JupyterHub Collaboration

This is a collab doc meant to create the SOW for coordination of the next phase of the EOSS grant for JupyterHub, administered through The Turing Way project at the Alan Turing Institute. This work builds on Phase 1 of the grant, delivered by Sarah Gibson. More information on the first phase of the grant can be found here: [EOSS D&I Progress Report](#).

Overview

The JupyterHub community is looking to enhance their capacity building, community growth and development. They are collaborating with The Turing Way project at the Alan Turing Institute to coordinate and generalise learnings across open source communities of practice. The work from Organizational Mycology is meant to support these efforts by 1) helping to develop a community engagement strategy by identifying blockers to contributor participation, 2) supporting the refresh of the Team Compass, 3) working on training for new contributors, and 4) embedding inclusive practices in the JupyterHub community.

Parties to this work package:

OM - Organizational Mycology - Jonah Duckles, Beth Duckles, Dan Sholler

TW - Turing Way - Kirstie Whitaker, Arielle Bennett

JH - Jupyter Hub - Min Ragan-Kelley, Chris Holdgraf

Deliverables

1) **"Voices of the JupyterHub Community" Research and Report**

- The goal of this portion of the work is to identify successes, challenges and opportunities to improve team dynamics within the JupyterHub community
- This portion will include a scoping phase where we discuss with the leadership and community what the research questions are and how to best address those questions. Possible research questions include:
 - What are the past and existing successes of the JupyterHub community?
 - What are the activities and actions that underpinned the past successes?
 - What opportunities has the JupyterHub community missed, or not been able to leverage for impact, in the past?
 - How can community development be improved?
 - What does the developer community need?

- What are the blockers to engagement with the JupyterHub community?
- OM will conduct two research projects, including:
 - 1) Interviews of around 20 people for the Voices of JupyterHub report. Responses to these interviews will also provide qualitative data to support other deliverables, specifically: developer on ramp pathways; event guides; and supporting team members in mentorship roles.
 - 2) Interviews with around 6-8 members of leadership for facilitating team process and workflow improvements.
 - The ideal interviewee participant profile and the number per group may change depending on the research questions, but we anticipate the work to include 26 to 28 interviews total.
- **Deliverable:** OM will analyze the interview data and write a draft report with a collaborative final report process to make the findings and approach openly available on the JupyterHub, Turing Way and Org Mycology's websites with a citable DOI on Zenodo.
- Appendix to include research instruments (survey questions, interview questions)
 - OM will make this information available for Turing Way to add to their materials.

2) Leadership Workshops

- JupyterHub and Turing Way seek to hold leadership workshops. OM will facilitate these workshops to help the JupyterHub Council focus on 1) challenges, 2) opportunities, and 3) actions related to building a diverse community.
- OM will create preparatory documents for the meeting
- OM will need prep time with project leads to shape the facilitations to be targeted and impactful for participants.
- OM will conduct the facilitation of a leadership group between 10 and 20 individuals.
- Time, date and length of time of workshop is to be determined.
- Deliverable: OM will write and share summary notes from the meeting and will work with the team to talk about openly shared documents that result from the conversations at the meeting.
 - At a minimum, we expect that a summary blog post will be published on the Jupyter blog.

3) Develop Four Community Based Initiatives

- We see four projects for OM to carry out in connection with the community. These projects are likely to be informed by the research we conduct. Our goal in these is to

enable and encourage community co-creation of these projects so that they meet the needs of the community.

- Each of these four will have different approaches and different emergent deliverables depending on what is needed.
- The four initiatives are:
 - Facilitate team process and workflow improvement conversations by:
 - Interviewing leaders about what process and workflow “best-practices” exist (e.g., in their heads or in written materials) but have yet to be implemented
 - Asking developers if they are aware of these best-practices and/or why these practices have yet to be implemented
 - Asking team about how workflow is going and how it could be improved
 - Identifying barriers to following set processes
 - Identifying missing policies and processes
 - **Deliverable:** written document from facilitation and conversations. This data includes interviews beyond the Voices of Jupyter interview data.
 - Further work for JH, 2i2c, TW which is out of scope for this SOW: Updating team compass to reflect these conversations
 - Developer on ramp pathways:
 - What are the GitHub Issue labels, milestones, and task management activities (for example) that make it easier for others to participate?
 - **Deliverable:** OM will provide a written document and/or other materials drawn from the Voices of Jupyter interview data. We will supplement the interview data with archival data from community interaction points (e.g., GitHub, Slack, forums and email lists).
 - Event guides (co-created with community):
 - Develop events guides for running team events that broaden participation (e.g., issue tackling parties), potentially in partnership with other members of the Jupyter community (like the Jupyter events manager).
 - **Deliverable:** OM will provide a written report with material ready to be integrated into the Turing Way and JupyterHub Team Compass materials. These materials will reference Voices of JupyterHub interview data, existing materials that have yet to be included, and any findings relevant to the Compass learned across the project.

- Supporting team members in mentorship roles
 - Create strategy around supporting team members in mentorship roles based on research from Voices of JupyterHub interviews. This may include how to give the team the proper training they need to support others, and defining structure and mentorship for them to succeed.
 - **Deliverable:** OM will provide a written document referencing the Voices of JupyterHub interview data.
- OM would like to have more conversations about how to do this best, but one option is to consider ways to convene folks through open working calls or other avenues for engagement that are flexibly-synchronous or asynchronous.
 - The Turing Way collaboration cafe calls (<https://book.the-turing-way.org/community-handbook/coworking/coworking-collabcafe.html>) and the JupyterHub collaboration cafe calls (<https://jupyterhub-team-compass.readthedocs.io/en/latest/meetings>) are good options for inclusive synchronous working.
 - The Turing Way Slack (<https://book.the-turing-way.org/community-handbook/communication-channels/slack-welcome-guide>) and Jupyter Discourse with a JupyterHub tag (<https://discourse.jupyter.org/c/jupyterhub/10>) are good options for asynchronous working
- Note: For each of these initiatives, OM plans to have at least documentation if not something written specifically, but we are keen to pay attention to how you work and do our work alongside what you already do. We don't want to produce reports just to produce reports.

Project Plan

Overall we'd like the main research efforts for the project to fit in the July - December 2024 timeframe with final report releases / collaborative review in mid January into February.

Scoping Phase (2 to 4 weeks)

- Develop research questions for the Voices of Jupyter/Community Assessment that will inform the Community Based Initiatives
- Discussion and creation of two research instruments to find answers to these research questions. The methods may change but they are likely to include interviews and a survey.

- Interviews with approximately 26 to 28 members split by groupings as needed.
 - Survey of the community.
- Discussion of Strategy Needs and any aspects of the community projects that are highest priority for the work.
- Submit research ethics approval to the Alan Turing Institute
 - Kirstie and Arielle will be accountable for submitting this form.
 - There are panel reviews every 6 weeks so submission at the end of July should lead to approval at the start of September.

Data Collection Phase (1 to 2 months)

- Research projects will be conducted and data will be gathered.
- Data will be summarized and shared.
 - Initial Assessment deliverable.
- OM will use this initial deliverable to work with the community on how best to do the community co-creation phase of research
 - Folks from JupyterHub and Turing Way teaming OM up to work with the community/get feedback would help to generate buy-in during this phase.

Community Co-creation Phase (4-5 months)

- Based on the research from the Data Collection phase, OM will share the information with other parties in this project and spearhead the four efforts. This may include running workshops to solicit feedback/information/dialogue, writing up short reports and drawing from data for the specific need to focus on these four areas.
- These projects will likely be accomplished in asynchronous modes, but OM will anchor these with synchronous conversations (e.g. working group calls etc)
 - Developer on-ramps and pathways
 - Team process and workflow improvements
 - Event guides
 - Supporting team members in mentorship roles.

OM anticipates the project could be accomplished within 6 to 8 months with the research phases up front and the community phase perhaps extending longer. OM is eager to get started on this project, getting the initial scoping and data collection started as soon as possible, with an eye to putting that work into practice in the final phase.

Changes to the timeline of the SOW will be agreed in writing between OM, TW, and JH.

Budget

Current budget for contracting is \$60,000

Additional facilitation budget for the in-person meetings is \$7,500 per meeting for the 2 in person meetings planned + travel and accommodation costs.

Other expenses relating to the delivery of the SOW will be approved on an ad hoc basis from overall project budget allocation in writing by the PI, Chris Holdgraf.

Ways of Working

- **Working document storage:** [location, accessibility] Working documents will be stored in a designated folder in OM's Google drive, which Chris Holdgraf, Kirstie Whitaker, Arielle Bennett will have access to. Additional people may be given access to working documents with the agreement of the project team.
- **Asynchronous:**
 - Slack on Turing Way
- **Project meeting schedule:**
 - **Regular project meeting**
 - **Frequency:** Monthly
 - **Format:** Zoom
 - **Length:** 60 mins
 - **Timing:** to be agreed based on team availability
 - **Agenda:** standing items to be agreed in the initial project kick off meeting
 - **Ad hoc project meetings**
 - The team will plan a series of alignment meetings at the start of the project, including the project kick off meeting - up to three based on project needs
 - **Format:** Zoom
 - **Length:** 90 mins
 - **Timing:** to be agreed based on team availability
 - **Agenda:** to be agreed in advance

Data Protection

- What types of data will we be capturing?

- Interview data (transcripts)
- Focus group data (transcripts and notes)
- Survey data (if necessary)
- Collaborative notes documents from leadership workshops
- Collaborative notes documents from community co-creation phase
- Observations from community forums (e.g., email lists, Discourse, Slack)
 - When these conversations are in open forums, may collect verbatim conversations as transcripts
 - When these conversations are in closed forums, only take notes that document conversation topics and characteristics of the conversation (e.g., tone, relational information) without identifying the participants
- Any existing data that have been collected by Turing Way/JupyterHub
- Will any of it be sensitive?
 - Interview questions may probe on socially-sensitive topics (e.g., positive and negative experiences)
- Where will it be stored? How will we keep it safe? How long will we store it for?
 - We will be storing recordings off cloud in a storage system which encrypts the files at rest on local disks.
 - Encrypted archives / snapshots of this data will be stored in backblaze B2 for disaster recovery purposes.
 - Recordings will be transcribed using a local GPU using WhisperX and transcripts themselves will be worked on using a private Google Drive with access only available to the OM team.
- What will it be used for?
 - All qualitative data will be used to generate insights for the four initiatives and related research questions
 - Excerpts from interviews may be shared in reports with participant consent and thorough de-identification
- Who should participants contact if they want their data removed or corrected?
 - To be included in informed consent form
- Informed consent form - See below