

# The CoReACTER Manual, version 0.1

The Community of Researchers Assessing Chemical Transformations and Exploring Reactivity (CoReACTER) is a research organization that operates within the Department of Chemical Engineering (ChemE) at Carnegie Mellon University (CMU). This Manual documents our practice of science, from how we hire researchers to join the CoReACTER and manage data to how we apologize when we make mistakes. We even have guidance for how we may edit this document!

*Land Acknowledgment:* We acknowledge that CMU occupies the unceded lands of the Seneca (Haudenosaunee), Shawnee, and Lenape peoples, in what is now known as Pittsburgh, Pennsylvania in the United States of America. We recognize that we benefit from this unjust occupation, the colonization of this region, and the concomitant displacement and elimination of Native peoples. We strive to enter into respectful relationships with the peoples of these lands and with the lands themselves as we search for avenues for healing, justice, and reconciliation.

January 14, 2025

*The CoReACTER Manual* was heavily inspired by the *CLEAR Lab Book*, written by Prof. Max Liboiron (they/them) and the Civic Laboratory for Environmental Action Research. EWCSS expresses gratitude to Tony Abel (he/him), Jeremy Adams (he/him), Helen Bergstrom (she/they), and Adrian Davey (he/him) for designing the “Racial Capitalism, Colonialism, and an Anti-Oppressive Chemical Engineering” course at the University of California – Berkeley, which gave them a space to reflect on revolutionary praxis in university settings.

Copyright 2025, The Community of Researchers Assessing Chemical Transformations and Exploring Reactivity (CoReACTER), Evan Walter Clark Spotte-Smith

The text associated with this work is licensed under Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International<sup>o</sup>.

The software associated with this work is licensed under a **Modified 3-Clause BSD Non-ML/AI License**, which states:

*Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:*

- 1. Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.*
- 2. Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.*
- 3. Neither the name of the copyright holder nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.*
- 4. The source code and the binary form, and any modifications made to them may not be used for the purpose of training or improving machine learning (ML) algorithms, including but not limited to artificial intelligence (AI), natural language processing, or data mining. This condition applies to any derivatives, modifications, or updates based on the Software code. Any usage of the source code or the binary form in an ML/AI-training dataset is considered a breach of this License.*

*THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT HOLDER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.*

# Contents

1. Contributors .....	7
2. How to Use This Manual .....	8
3. Principles .....	9
4. Rules .....	11
5. Guidelines .....	12
6. ReACTERS .....	14
6.1. Who Is a ReACTER? .....	14
6.2. Hiring .....	14
6.2.1. General Procedure .....	14
6.2.2. Hiring Candidates With External Funding .....	16
6.2.3. Undergraduate Students .....	16
6.2.4. Master's Students .....	17
6.2.5. Doctoral Students .....	17
6.2.6. Postdoctoral Researchers .....	18
6.3. Joining from Outside the CoReACTER .....	18
6.4. Onboarding .....	19
6.4.1. Introductions .....	19
6.4.2. Access .....	19
6.4.3. First Week Checklist .....	19
6.4.4. Goals and Research Plan .....	20
6.4.5. Essential Readings .....	21
6.5. Expectations of ReACTERS .....	21
6.6. Exiting .....	21
6.6.1. Leaving Voluntarily .....	21
6.6.2. Being Asked to Leave .....	22
6.6.3. Before You Go .....	22
7. Lab Management .....	24
7.1. Projects .....	24
7.1.1. Starting a Project .....	24
7.1.2. Project Teams .....	24
7.1.3. Suspending and Ending Projects .....	25
7.2. Lab meetings .....	25
7.2.1. Group Organizational Meetings .....	25
7.2.1.1. Attendees .....	25
7.2.1.2. Location .....	25
7.2.1.3. Scheduling .....	26
7.2.1.4. Agenda-Setting .....	26
7.2.1.5. Meeting Procedure .....	26
7.2.2. Group Research Meetings .....	26
7.2.2.1. Location .....	26
7.2.2.2. Attendees .....	27
7.2.2.3. Scheduling .....	27
7.2.2.4. Agenda-Setting .....	27

7.2.2.5. Meeting Procedure .....	27
7.2.3. Project Team Meetings .....	27
7.2.3.1. Attendees .....	28
7.2.3.2. Scheduling .....	28
7.2.3.3. Meeting Procedure .....	28
7.2.4. <i>Ad Hoc</i> Meetings .....	28
7.2.5. Running Meetings .....	29
7.3. Decision-making .....	29
7.4. Training and Mentoring .....	30
7.4.1. Mentoring Relationships .....	30
7.4.2. Skill Training .....	30
7.5. Communication and Information Sharing .....	31
7.5.1. Within the CoReACTER .....	31
7.5.2. Outside of the CoReACTER .....	31
7.6. Funding .....	32
7.6.1. Pursuing Research Funding .....	32
7.6.2. Disclosing Funding .....	32
7.6.3. Managing Funds and Making Purchases .....	33
8. Accountability .....	34
8.1. Restorative Justice .....	34
8.2. Apologizing .....	35
8.3. Call In, Call Out .....	35
8.4. Community Involvement .....	36
8.5. Auditing .....	36
9. Open Science .....	37
9.1. Open Source Software .....	37
9.2. Open (and FAIR) Data .....	37
9.3. Open Access .....	37
10. Software .....	39
10.1. CoReACTER Stack .....	39
10.1.1. Programming languages .....	39
10.1.2. Software Engineering .....	40
10.1.3. Organizational Software .....	40
10.1.4. Chemical Simulations .....	41
10.1.5. Experiments .....	41
10.1.6. Other Chemical Software .....	41
10.2. Developing Software .....	41
10.2.1. Our Code, Their Code .....	41
10.2.2. Code Style .....	42
10.2.3. Software Maintenance .....	42
10.2.4. Licensing .....	43
11. Computing .....	44
11.1. Available Resources .....	44
11.2. Getting Started on an HPC System .....	44
11.3. Budgeting .....	45

11.4. High-Performance Computing Tips .....	47
12. Managing and Sharing Data .....	48
12.1. “Soft” Data and Institutional Memory .....	48
12.2. Metadata .....	48
12.3. Data Organization .....	49
12.4. Online Repositories .....	49
12.5. Cloud Storage .....	49
13. Writing, Presenting, and Communicating Science .....	50
13.1. Authorship .....	50
13.2. How We Write .....	52
13.2.1. Software .....	52
13.2.2. Process .....	52
13.2.3. Style .....	55
13.2.4. Policy on Generative Machine Learning .....	56
13.3. Where We Publish .....	56
13.3.1. Preprint Servers .....	56
13.3.2. Journals .....	57
13.4. Meetings and Conferences .....	58
13.4.1. When to Attend .....	58
13.4.2. Where to Present .....	59
13.4.3. Preparing a Presentation .....	59
13.4.4. Presentation Style .....	61
13.4.5. Travel Policy .....	63
14. Resources for ReACTERs .....	64
14.1. Health and Wellness .....	64
14.1.1. Disability and Accessibility .....	64
14.1.2. Physical Health .....	64
14.1.3. Mental Health .....	64
14.2. Funding .....	65
14.2.1. Scholarships and Fellowships .....	65
14.2.2. Research Grants .....	65
14.2.3. High-Performance Computing Allocations .....	65
14.3. Research Tools .....	65
14.4. Internships and Jobs .....	65
14.5. Other Resources .....	66
14.5.1. Resources for Students .....	66
14.5.2. Conferences and Workshops .....	66
14.5.3. Legal Resources .....	66
Bibliography .....	67

## 1. Contributors

The original draft of *The CoReACTER Manual* was written by **Evan Walter Clark Spotte-Smith (they/them)**. Evan (referred to in this text as EWCSS) is an Assistant Professor of ChemE at CMU. They are the Founder of the CoReACTER and, as a “Principal Investigator” (or PI), they are nominally responsible for the research organization.

Ideas for the original draft were improved and clarified through discussions with **Julia Isabelle McKeown (they/them)** and **Dr. Samuel M. Blau (he/him)**. Julia is a poet and folklorist who works with the Othering and Belonging Institute (OBI) at the University of California — Berkeley (UC Berkeley). Dr. Blau is a career research scientist at Lawrence Berkeley National Laboratory (LBNL) and a longtime collaborator of EWCSS’.

Parts of this manual are adapted from resources developed by other faculty at CMU. In particular, I want to acknowledge important inputs from **Prof. Rachel Kurchin (she/her)** and **Prof. John Kitchin (he/him)**.

## 2. How to Use This Manual

This document is rather long. We don't expect any ReACTER (or interested individual) to read *The CoReACTER Manual* cover to cover. Rather, we encourage you to read what you need, as it becomes relevant to you. If you're interested in joining the CoReACTER, you should read Section 3 to learn about our core philosophy and politics and Section 6 to see how to join (and leave) the organization. Our ethical ideas and guidelines are sprinkled throughout this manual, but you might want to read Section 8 to learn how we hold ourselves, each other, and our community members accountable or Section 9 to see our perspective on open science. For current ReACTERs, read Section 6.4 on onboarding, and then read whatever sections you need to do your work. For instance, if you're working on writing a paper or preparing a conference presentation, you probably want to read Section 13.

What if you have a problem with this manual? A problem could be a small typo, something that you think is inaccurate or poorly explained, or a major idea that you disagree with or think is missing. Whatever you want to change, we have a procedure for editing *The CoReACTER Manual*. Please submit proposed changes to the CoReACTER through a Pull Request (PR) on GitHub<sup>o</sup>. In your PR, please summarize your changes and explain your rationale for making them. We will discuss any proposed edits in a future group organizational meeting (see Section 7.2.1). Because *The CoReACTER Manual* is a foundational document for the CoReACTER, laying out our principles and our ways of engaging in scientific research, it is important that we have agreement regarding the text that is unanimous wherever possible.



### 3. Principles

**All science is political.** From the questions that we ask and the approaches that we take to answer them to the relationships that we build and the ways that we (and others!) use science to make decisions, we are engaging in politics and driving political processes at various scales. Rather than try to ignore the role of politics in our scientific praxis, we are intentional and thoughtful about our political positionality.

The CoReACTER is an anti-oppressive research organization, aiming to subvert practices that are harmful and/or destructive to the well-being of ourselves, other humans, non-human life, and the planet. In particular, we align ourselves with ideas from:

- (Intersectional) feminism
- Disability justice
- Environmental justice and sustainability

and we actively oppose the forces of:

- White supremacy, anti-Blackness, and racism
- (Cis)sexism, homophobia, and transphobia
- Capitalism and classism
- Colonialism
- Ableism
- Nationalism(s) of all forms
- Fascism and totalitarianism

Academia is a deeply flawed system that is embedded in other flawed and harmful systems. Because of this, we cannot completely avoid engaging in Capitalism (modern research universities operate as capitalistic enterprises; we also need to finance our research), colonialism (CMU's campus is on stolen land, within a colonized country and colonizing nation-state), and nationalism (most research that takes place in the United States is publicly funded, and publicly funded research generally supports American national interests); perhaps this is also true of some of the other evils that we have outlined. Nonetheless, we openly challenge and subvert oppressive paradigms and structures where and how we can.

Our core principles, which guide us as we practice science, are:

- **Honesty:** We are honest with ourselves, with our fellow ReACTERs, with the scientific community, and with the world. Honesty means truth-telling — saying what we know or believe to be right, whether it is convenient or not. It also means transparency. We strive to communicate openly throughout the scientific process, to share early and often rather than hoarding or hiding information.
- **Humility:** Often, “humble” is used to mean that one has a low opinion of one's self. Rather, when we say that we aim to be humble, we mean that we aim to avoid both excessive pride and self-deprecation. Through careful reflection and honest dialogue, we try to maintain an accurate, moderate understanding of ourselves, our knowledge, our behavior, and our place in our communities. Critical to effective humility is knowing our own limits, knowing what we know and what we don't know!
- **Collectivism:** No scientist is an island. We are all part of many communities — groups of friends, colleagues, biological and found families, neighborhoods, folks who go to the same hair salon. Our approach to our work centers these communities, rather than our own interests. We are motivated to use scientific inquiry and our expertise to make our communities and the broader world better than we found them. Our work seeks to uplift others, particularly those in need. We will never harm or put down others for our personal benefit.

- **Justice:** Historically, Western science has been practiced primarily by the elite members of society. This elite practice has created or supported (and continues, in many ways, to create and support) regimes of oppression, repression, subjugation, and exclusion. While the CoRe-ACTER operates within traditional academic structures, we aim to push against oppressive practices and create more just policies and outcomes, within the Academy and outside of it.
- **Kindness:** In a world of competition, we choose collaboration. In a world of selfishness, we choose to give freely. In a world that is hard and unforgiving, we choose softness. Where we find ourselves given the choice, we choose to be kind: to ourselves, to each other, and to those around us. We aim to be empathetic and understand the perspectives of others, and where we cannot empathize, we sympathize and listen.
- **Curiosity:** Science as a philosophy offers tools for problem-solving, and we employ these tools to address meaningful problems, from fundamental theoretical questions to global challenges like anthropogenic climate change and plastic pollution. However, we refuse to treat science only as a means to a practical end. Rather, we center the joy of inquiry and celebrate discovery as an end in and of itself. As we do our work, we seek to learn and to facilitate the learning of others.

## 4. Rules

When rules are just, well-designed, and logical, they should never be broken. We have identified a small number of such universal rules that ReACTERs should always adhere to. Non-universal policies — which may have exceptions or may only apply in certain situations — are described in later chapters.

1. **Health and well-being come before research.** If you are physically ill, **do not** come into the office or lab,<sup>1</sup> and **do not** work from home. If you are physically or emotionally exhausted, **do not** continue working; rest until you are no longer exhausted, and then you may return. If you are suffering from sorrow, heartbreak, grief, or other conditions of the mind, **step away from work** and take space to care for yourself. And if any of these afflictions — physical, mental, or emotional — are chronic and persistent, please try and discuss them with your colleagues (at least EWCSS) so that reasonable and helpful accommodations can be made.
2. **Keep lab and office space clean and safe.** The CoReACTER performs both experimental and computational research. “Safety” and “cleanliness” mean different things in these different contexts.<sup>2</sup> In all cases, though, high safety quality prevents health and environmental disaster, while cleanliness shows respect for one’s self, one’s colleagues, and one’s surroundings.
3. **Document everything.** We are an open science lab, and many of our papers include datasets that we publish and share with the community. We also write, contribute to, and maintain open source scientific software. Documentation is the difference between a useful and a useless dataset. It’s the difference between code being arcane and maintainable. If you’ve invented something new, it’s the difference between being able to patent that invention and not. Even just for your own purposes, it is the difference between a key breakthrough and completely forgetting an idea. If you’re in doubt, **document it**.

---

<sup>1</sup>This not only helps to take care of you, but it also protects others, particularly those who may be immunocompromised or at risk of disease.

<sup>2</sup>In the future, details will be discussed in later chapters.

## 5. Guidelines

1. **Communicate.** The saying goes that “silence is golden”, but in our experience, open communication is a lot nicer and helps a research organization run well. If you’re having a problem (technical, personal, or otherwise), ask for help! If you see someone else struggling, ask them what’s wrong or offer your support. Chat with your fellow ReACTERs to coordinate equipment usage and compute time, to brainstorm new ideas, and to work through conflicts. If you’re in doubt as to whether communication is needed, err on the side of over-communication.
2. **Set and respect boundaries.** If you let it, the worlds of Academia and scientific research will take everything from you — your time, your peace of mind, even your health. Boundaries are a way that we protect ourselves from these harms. Early and often, consider what you need to be happy and healthy, and state those needs as boundaries. These boundaries could involve how you communicate and what you communicate about; when you do and do not engage with work; and much more. Hopefully, those around you will respect clearly communicated boundaries, but you may need to periodically re-assert them. You should feel free in doing so, as maintaining and protecting your boundaries is ultimately good for you and those around you. Likewise, if your colleagues tell you their boundaries, listen, make note of them, and (assuming these boundaries are reasonable) take care not to cross those boundaries. If you do violate them in some way, recognize that harm, apologize, try to repair, and work to improve so that you respect your colleagues better in the future.
3. **Limit work time/take breaks.** In the CoReACTER, productivity is not our ultimate goal. We want to improve the practice of science and use science to better the world, but not at the expense of our own health and potential for joy. We further believe that time spent on research does not necessarily correlate with productivity. Rather, we understand that you will be a better scientist if you take time for rest, for personal relationships with friends, families, and loves, for hobbies, for fun, for absolutely nothing at all. We encourage ReACTERs to design sustainable patterns of labor that incorporate regular rest and relaxation. We also try to incorporate breaks into our rhythm as a research organization, with shared meals, games, informal discussions, retreats, and more.
4. **Recognize and fix mistakes early.** We will make mistakes. That means you will make mistakes, too. Rather than waiting for problems to arise, double-check yourself early (or ask a colleague to help look for mistakes!), so that you can correct them before they blow up (literally or figuratively). This advice applies to research projects — selecting parameters for calculations, designing experiments, performing analysis — and also to interpersonal interactions. Taking time to reflect and thinking about how you could have handled a situation better will help you to avoid tension and minimize the harm you do to others!
5. **Be on time.** Especially because we limit the time that we spend on work (see above), our hours in the office and the lab are valuable. Being on time is a way to show care and respect for yourself and for those around you. If you are worried about running late, try to give yourself buffer time so that others don’t need to wait for you. If you know that you’re going to be late, communicate that as early as possible so that others can respond and plan accordingly. We understand that things happen, and we will always try to give grace (especially acknowledging that relationships around time are culturally dependent), but it’s easier if you do your part!
6. **Attend meetings.** We know, there are a lot of meetings, possibly too many. Where possible, we encourage you to structure your work to minimize the number of meetings that are on

your plate. However, we strongly advise that you attend and be fully present for meetings that are essential for the CoReACTER (group organizational and research meetings) or for your specific project (project team meetings, meetings with collaborators and funders). These meetings will help you to stay up to date on what your colleagues are working on, help you to get feedback on your work, and (if well designed) build relationships and community.

## 6. ReACTERS

### 6.1. Who Is a ReACTER?

Throughout this document, we mainly refer to ReACTERS, rather than “group members”, “students”, or other terms. In addition to being cute, this serves a practical purpose for identity development and community building.

ReACTER is short for “Researcher Assessing Chemical Transformations and Exploring Reactivity”. With this meaning, pretty much any chemical scientist (e.g., chemist, chemical engineer, or materials scientist) could reasonably call themselves a ReACTER. The bar to entry is low. Are you interested in chemistry? In reactions? In transformative change (in a flask or in the Academy)? Great, you are welcome here!

When we refer to ReACTERS in this text and in our common usage, we are referring to folks who not only practice chemical science but are committed to the principles and practices of the CoReACTER that are laid out in this document. Note that that does not mean that you agree with everything that’s written here, or everything that someone in the CoReACTER says. This is a living document (see Section 2), and we expect that we will need to amend and expand our manual as time goes on, in response to brilliant ideas from folks like you! But if you read Section 3 and feel uncomfortable, frustrated, or like you disagree on major points, this may not be a community that you want to devote your energy and attention to.

For now, ReACTERS include EWCSS and the students that they advise. Over time, we expect our ranks to expand to include our core collaborators (see Section 6.3 for more discussion on how to join the CoReACTER and “officially” become a ReACTER).

### 6.2. Hiring

Though the CoReACTER differs in many respects from conventional research groups, our funding and personnel structures are much the same. EWCSS, as an Assistant Professor, is paid 10 months of the year to teach, and they pay their own salary for the remaining 2 months through the CoReACTER’s startup funds and grant money. Doctoral students, postdoctoral researchers, and staff researchers are paid by a combination of department funds, CoReACTER grants, as well as external scholarships and fellowships.<sup>3</sup> Undergraduates and Master’s students are typically compensated through research course credits, though financial compensation is sometimes available; for summer students, we also pay through grants.

Given this structure, the CoReACTER hiring process is somewhat different depending on whether the hire is for an undergraduate student, Master’s student, doctoral student, or postdoc. Here, we lay out our general procedure and then discuss specifics for these different positions.

#### 6.2.1. General Procedure

The first step in the hiring process is deciding whether or not to hire. Generally, this decision will be made in a group organizational meeting (see Section 7.2.1) once funding has been obtained. In the short term, while the CoReACTER is mainly funded by EWCSS’ startup funds, hiring decisions are less tied to what was written in a grant and can be based mainly on what is right for the organization. In making a decision to hire, the following factors should be considered:

- Are there important projects that cannot be completed (e.g., because of a lack of skills among ReACTERS) or will not be completed without additional assistance (e.g., because of a lack of available time)?

---

<sup>3</sup>This is a slight simplification; doctoral students can also be partially paid to serve as teaching assistants.

- Is there enough capacity in the CoReACTER at present to train and mentor a new ReACTER? If the new hire cannot be adequately supported, it would be irresponsible to bring them into the organization.
- What type of role (undergrad, Master’s student, doctoral student, or postdoc) is most appropriate? For grant-based hiring, this will normally be decided during the writing and budgeting process. But if not, consider:
  - How much time (e.g., hours per week) will be required to do the work considered?
  - How quickly are results expected? In general, we try hard not to rush, but time-sensitive demands are hard to avoid in grant-based scientific research. The earlier a researcher is in their career, the less quickly they can be reasonably expected to adapt to a new project, new techniques, etc.
  - How long will this funding last? If the funding would support a postdoc for only one year but a doctoral student for two years, it might be best to hire a doctoral student to ensure stability.

If the CoReACTER decides to move forward and hire for a position, the next step is to design a job posting and solicit applications.<sup>4</sup>

For each hire, the CoReACTER will assemble a search committee, typically involving at least 3 people. The exact committee size will vary, depending on the number of ReACTERS interested in the outcome of the hire as well as the number of applicants. Prior to any evaluations, the search committee will meet to discuss the goals for the hire, define criteria for assessment, and discuss bias mitigation techniques to improve fairness in the hiring process. Assessment criteria will vary from position to position, depending on the project being hired for, the nature of the position, and the needs of the CoReACTER at the time of the hire.

For hires requiring written applications (hiring undergrads and postdocs), all members of the search committee are expected to read and evaluate all written applications, including cover letters, *curricula vitae* (CVs), and other materials.<sup>5</sup> In cases where the number of applicants is very large, search committee members will each be allocated a subset of applications, but at least two search committee members will evaluate each applicant.<sup>6</sup>

Following the evaluation of written materials (where relevant), the search committee will convene to discuss the candidates, assessing which candidates should be invited to interview. Each candidate will be reviewed in a non-comparative manner, assessed on their strengths and weaknesses without reference to other candidates. Search committee findings will be presented at a group organizational meeting (Section 7.2.1), where there will be a discussion on the committee’s findings. The CoReACTER will always vote to determine who to bring on for interviews, but we expect to agree with the recommendations of the search committee.

In general, all interviews will be held virtually so as not to give an unfair advantage to candidates who are local or are able to travel to CMU. Candidates will be given a set of standard questions ahead of time, and they will be told that we may ask follow-up questions that are related to the provided questions. For positions where prior research experience is expected, candidates will be asked to prepare short (30-minute) introductions to their research. In these cases, questions

---

<sup>4</sup>TODO: We need a better sense of what hiring procedures at CMU look like.

<sup>5</sup>TODO: Is it important to anonymize applications? Will that even be possible, given the small-world nature of academia and the fact that folks have their publications on their CVs? Look for best practices in academic hiring.

<sup>6</sup>We don’t have an exact rule for determining what is a “very large” number of applications, but for now, let us say that any more than 30 applications would be considered a large amount.

can also address the research that was discussed. All search committee members are expected to attend all interviews.

After all interviews have been conducted, the search committee will again meet, discuss their evaluations, and report back to the CoReACTER in a group organizational meeting. Final hiring decisions will be made collectively.

### **6.2.2. Hiring Candidates With External Funding**

In the general procedure described above, we assume that the CoReACTER is soliciting external applications. However, it often happens that an individual looking for a research position will reach out to a ReACTER about a position (a so-called “cold call” or “cold e-mail”). In most cases, we will not make a hiring decision based on an unsolicited application, as this would be unfair. If we decide to hire a candidate without advertising a position and waiting for applications, then that effectively creates a back-channel, privileging those who communicate with ReACTERS before we post job listings and taking away opportunities from candidates who (reasonably) did not believe that the CoReACTER was in a position to hire.

We make an exception for candidates seeking to join the CoReACTER who already possess external funding. Typically, this includes incoming and current doctoral students who have obtained fellowships,<sup>7</sup> as well as soon-to-graduate doctoral students and current postdoctoral researchers with postdoctoral fellowships.<sup>8</sup> In this case, many of the criteria that we describe in Section 6.2.1 do not apply. For example, a doctoral student or postdoc with external funding is not expected to work on existing projects but should pursue their own proposed research. We will therefore consider individuals with external funding sources on a case-by-case basis, using the following criteria:

- Is the candidate’s proposed work in alignment with work that is ongoing in the CoReACTER, or work that we plan to pursue in the near- to mid-term? That is, would the candidate find collaborators within the CoReACTER?
- Does the candidate demonstrate alignment with CoReACTER principles, as well as a willingness to learn and grow?
- Is there enough capacity in the CoReACTER at present to train and mentor a new ReACTER?<sup>9</sup>

Note that the exception described here does not apply for individuals *wishing to obtain* external funding. That is, if an individual reaches out asking a ReACTER to sponsor their application for a scholarship or fellowship, we will not do so without first advertising the CoReACTER’s *general* willingness to serve as a sponsor, inviting proposals from the community, and equitably deciding among the proposals as a collective.

### **6.2.3. Undergraduate Students**

Hiring for undergraduates follows two different procedures depending on if the student’s work will take place during the academic year or during the summer. In both cases, we currently only plan to hire undergraduates from within CMU. However, we hope that this changes in the near future as we establish external connections.

For undergraduates working during the academic year, the CoReACTER will solicit applications by communicating to particularly relevant departments (e.g., Chemistry, ChemE, and Materials Science and Engineering or MSE). This will mainly involve sending the job description through

---

<sup>7</sup>Examples include the National Science Foundation Graduate Research Fellowship Program, the Hertz Fellowship, and the Ford Foundation Predoctoral Fellowships.

<sup>8</sup>Examples include the Schmidt Science Fellowships and the Arnold O. Beckman Postdoctoral Fellowship.

<sup>9</sup>This criterion is unchanged from Section 6.2.1.



department listservs including undergraduate students. EWCSS may also advertise in ChemE courses that they teach to undergraduates. Typically, we expect for CMU undergraduates working during the academic year to be compensated through research credits. However, if the financial circumstances of a selected qualified candidate necessitate direct financial compensation, we will compensate that individual through pay if at all possible.

For undergraduates who will work during the summer, we plan to sponsor at least one application for the Chemical Engineering Summer Scholars (ChESS)<sup>o</sup> program each year. ChESS is related to the CMU Summer Undergraduate Research Fellowship (SURF)<sup>o</sup> program and offers \$4,500 stipends for students performing summer research for 8-10 weeks. We will advertise our interest in working with students through ChESS before or shortly following the annual ChESS announcement through the CMU ChemE listservs. We may make additional summer hires external to ChESS/SURF; in these cases, we will advertise to all relevant departments at CMU, as described above for academic-year appointments. Even for students hired outside of ChESS/SURF, we intend to provide financial support for all summer researchers.

#### **6.2.4. Master's Students**

The CMU ChemE department has several Master's programs<sup>o</sup>. Among these, the Master's of Science in Chemical Engineering<sup>o</sup> (MS-ChemE) and Master's in Artificial Intelligence Engineering-Chemical Engineering<sup>o</sup> (MS-AI-ChemE) both require students to perform research-based projects, with the MS-ChemE degree requiring that students participate in a three-semester research project and the MS-AI-ChemE program requiring 12 project- or research-based credits.

In ChemE, Master's students are matched with faculty advisors. If the CoReACTER decides to hire Master's students, we will submit project descriptions to the department early in the Fall semester; these project descriptions will be developed by the hiring committee and voted upon by the CoReACTER. Students will be allowed to select projects that they are interested in working on, and then the ChemE Department Head and Graduate Advisor will use student preferences to assign students to projects.

#### **6.2.5. Doctoral Students**

First, note that we do not require written applications for PhD students, and we do not expect students to have performed research prior to beginning their doctoral studies.

EWCSS is an Assistant Professor of ChemE, our home department. In ChemE, PhD students do not enter the department in their first semester with an advisor. Rather, during their first semester, students hear about the research that different professors are performing, meet with professors, and then the department matches students with faculty, like with Master's students (see Section 6.2.4). This matching process takes into account the availability of faculty funding (*i.e.*, how many students each faculty member needs or wants to hire), student preferences, and factors such as seniority.<sup>10</sup>

Because pool-based hiring, as it is operated in ChemE, is based largely on student preferences, the CoReACTER will not need to make internal decisions regarding hiring, though we will typically discuss potential PhD students and our thoughts about them in a group organizational meeting.

---

<sup>10</sup>In CMU ChemE, junior faculty members are given preference for PhD student hiring, to help them build successful groups.

### 6.2.6. Postdoctoral Researchers

Postdoc hiring will mainly follow the general procedure outlined in Section 6.2.1. We will advertise all positions online (e.g., via BlueSky<sup>o</sup>, appropriate e-mail listservs, and the CoReACTER website) and through word of mouth (e.g., at academic meetings). Maintaining diversity in the applicant pool is critical for improving diversity in our organization and for improving equity in the hiring process. Therefore, we may target organizations serving minoritized researchers (e.g., the National Society of Black Engineers or oSTEM) for additional outreach and advertisement.

### 6.3. Joining from Outside the CoReACTER

If you're reading this, you're probably someone who is interested in the work conducted within the CoReACTER or our approach to scientific research, but maybe you're not in a position to join formally under the hiring mechanisms described in Section 6.2. Maybe you're a faculty member looking for collaborators who share a perspective or an ethic. Maybe you're a graduate student in another group at CMU or at another university. Or maybe you're not in the Academy at all, but you want to get involved in research. Regardless of your current position, you're asking: is there a way for me to be a part of this organization?

There certainly is!

The first step is to reach out to any ReACTER, inquiring about initiating a collaboration. In this communication, you should outline:

- Why you are interested in working with the CoReACTER
- How you would like to work with the CoReACTER<sup>11</sup>
- How long do you expect an initial collaboration to last?

When we receive an inquiry, we will discuss it at a CoReACTER group organizational meeting. If we agree that the collaboration would be fruitful — with the prospective new ReACTERs contributing positively to us as scientists and as a community, and with us in some way assisting the new ReACTERs — then we will work with the inquiring individual(s) to draft a Memorandum of Understanding (MOU), laying out rough plans and expectations.<sup>12</sup>

Once the MOU has been agreed upon by all relevant parties, the external collaborators are considered to be “Provisional ReACTERs”. Individuals in this provisional status have full rights to access CoReACTER resources and to participate in group decision-making. However, the expectation is that the relationship between the provisional ReACTERs and the CoReACTER will terminate completely at the end of the initial collaboration period (if the relationship does not end earlier by mutual agreement).

At the end of the initial collaboration period described in the MOU, we will discuss ongoing collaboration with the provisional ReACTERs. If all parties believe that the relationship has been positive and mutually beneficial, and that the relationship would continue to be beneficial, then the CoReACTER can agree to accept the external collaborator as a ReACTER. They will then be fully included in CoReACTER business with all of the associated rights and privileges until they decide to leave the CoReACTER or the relationship is otherwise terminated (see Section 6.6).

---

<sup>11</sup>This should be a short (1-2 page) project proposal, with sufficient technical detail that we could reasonably evaluate the project by reading the proposal and the cited literature.

<sup>12</sup>This is a form of gatekeeping, without a doubt. How do we balance the desire for an open community and the need to protect ourselves from harm? An area for discussion.

## 6.4. Onboarding

### 6.4.1. Introductions

Once you've officially begun your work in the CoReACTER, you should introduce yourself to your colleagues through an e-mail. This is important because you may not see everyone around the office, and because you should be able to plan and present yourself the way you want.

It hopefully goes without saying, but every ReACTER should introduce themselves. That means undergrads, graduate students, postdocs, visiting scholars, and even external collaborators.

Send your introduction e-mail to the CoReACTER listserv [coreacter-main.andrew.cmu.edu](mailto:coreacter-main.andrew.cmu.edu) with the following information:

- A picture of you. This doesn't need to be professional. You can show us you enjoying a hobby, spending time with loved ones, etc.
- Your preferred form of address, including your preferred name (which doesn't need to be a name you were given at birth!) and pronouns (e.g. "he/him/his", "she/her/hers", "they/them/their", "xe/xim/xir", etc.)
- What you'll be doing in the CoReACTER. For instance, "I'm a graduate student, and I'll be working on methods for chemical reaction networks". If you don't know what project(s) you're going to be working on, you can say that!
- Your background. This can include places you've lived before, things you've done (where you went to school, what jobs you've held, etc.), and any training you have
- Your goals. What are you trying to accomplish while you're in the CoReACTER?
- Some fun facts about you. What's your pet peeve? Have you ever had a brush with the supernatural? What do you think is the correct way to pronounce "gif"? If you can't think of any fun facts to provide, ask EWCSS!

### 6.4.2. Access

As a ReACTER, you'll need access to some physical and digital resources to participate in the organization and complete your work. Every ReACTER will need access to the following tools, and you should be invited shortly after joining the organization. If you do not receive invitations within one week of joining, send an e-mail to EWCSS.

- **CoReACTER GitHub organization** (<https://github.com/coreacter/>)
- **Zotero reference manager** (<https://www.zotero.org/groups/5748695/coreacter>): If you already have a Zotero account, send EWCSS your username. Make sure that your account is associated with your CMU e-mail address if possible, as this will give you free unlimited storage.
- **Zulip** (<https://coreacter.zulipchat.com>)
- **Anytype**

If there's something that you need that you don't see here, ask around!

### 6.4.3. First Week Checklist

Your first week will probably be hectic. While we don't expect you to hit the ground running, there are a couple of things that are important to get out of the way:

- **Introduction:** As noted above (Section 6.4.1), we want to get to know you! Try to send at least a quick introduction on your first day. If you can't manage that, send it out by the end of your first week. Note that this will require you to have access to the CoReACTER listserv!
- **Benefits:** If you are a postdoc who is eligible for benefits (insurance, retirement savings, etc.), sign up as soon as possible. Your benefits options will be explained during your HR

onboarding. You can also get information on the HR website<sup>o</sup>. When you're ready, make your benefits elections on Workday.

- **Trainings:** Most likely, you'll have some necessary trainings to get out of the way for your new role. For instance, you'll need to take a cybersecurity training. For most people, these aren't particularly fun, some of them are quite important. You should get them out of the way so that no one has to hound you down later.
- **Equipment:** You'll need some tools to do your work. Specifically, everyone will need a laptop and related peripherals (keyboard, mouse, monitor, etc.). Folks working in the laboratory will need personal protective equipment (PPE). Check with EWCSS to see if any appropriate equipment is available within the CoReACTER. If not, work with them to make the purchase. Generally, laptops will be purchased through the CMU book store<sup>o</sup>, though other vendors can be used.
- **Background reading:** Most ReACTERs will enter the CoReACTER with a project. You're not expected to make any progress on that project in your first week, but you should take some time to read up on your project and the related literature so that you can start asking questions and engaging in meetings. If you're funded by a grant (as opposed to EWCSS' startup funding or an external fellowship), read the grant proposal.
- **Get on the website:** Find a picture of you that you like and write a short bio describing your background and interests. Send these to EWCSS with the subject line "Onboarding: CoReACTER website <name>", where <name> is your name.

#### 6.4.4. Goals and Research Plan

After you've settled down and managed some of the initial logistics, it's time to start thinking about what you're going to do while working in the CoReACTER. If you are paid on a grant, there may already be a well-defined project that you're expected to work on. Alternatively, if you walk into the CoReACTER with a fellowship or other open-ended funding (including EWCSS' start-up funding; see Section 14.2), you might have considerable freedom to research what excites you. In any case, we want the time that you spend in the lab, at your computer, and/or in your office to be as well adapted to you as possible.

In one of your first one-on-one meetings with EWCSS (see Section 7.4), you should discuss your needs, desires, and goals. To guide this discussion, try to answer the following questions (adapted from a guide developed by Prof. John Kitchin of CMU):

1. What training did you receive prior to joining the CoReACTER? This could include prior schooling, previous jobs within or outside of academia, etc.
2. What did you like about your previous experiences? What didn't you like?
3. What motivates you to do research? How and why did you come here?
4. What are your core values? What guides you in your life, personally and professionally?
5. What are your strengths?
6. What are areas that you want to or need to work on, personally or professionally?
7. What skills do you hope to develop through your work with the CoReACTER?
8. What do you want to achieve in your time at CMU?
9. What are your career goals? Where do you want to end up in 5 years? In 10 years?
10. What do you need to succeed? What accommodations do you need for your work?
11. If you could study anything at all, what would it be? What about that project or subject draws you in?
12. What else should I know about you and your goals?

### 6.4.5. Essential Readings

Each ReACTER may work on different projects, use different tools, take different classes, and read different papers. In spite of this, it is important for ReACTERs to have a shared base of knowledge and understanding, both to aid technical conversations with a shared vocabulary and to facilitate radical communal action. Here is a (living, growing) list of texts that all ReACTERs should strive to read during their first six months in the CoReACTER. We try to limit this list to short, quick reads, so this should not be a significant burden.

1. *Mutual Aid* by Dean Spade: While the CoReACTER is not a mutual aid organization, we are fighting against many of the same forces that Spade describes, including systems of oppression, burnout, and a pervasive culture of perfectionism. This accessible text is an invaluable resource for developing and managing sustainable, anti-oppressive, mission-focused organizations like the CoReACTER.

### 6.5. Expectations of ReACTERs

The CoReACTER principles (Section 3), rules (Section 4), and guidelines (Section 5) serve as a minimal set of expectations for ReACTERs.

Beyond these “ground rules”, we have some additional expectations that we hope will lead to better outcomes for each ReACTER individually and for the CoReACTER as an organization, including:

- Taking **at least two (2) weeks of vacation** per year, in addition to university holidays.
- Actively seeking out new sources of information (e.g., papers, journals, and conferences) and resources (e.g., grants, fellowships, and computing resources).
- (For students) maintaining good academic standing and keeping track of your progress towards graduation.
- Pursuing activities that nourish you and that serve your communities. Your self-serving and other-serving activities don’t need to be the same, but they certainly can be!

Just as important, if not even more important, are the things that you are *not* expected to do. These include:

- Being present in the laboratory or office during fixed hours (other than regular meeting times) or beyond normative working hours (e.g., in evenings or on weekends, more than the hours expected based on your course credits or work contract).
- Communicating with any other ReACTER(s) when you are not working. This includes all technical and non-technical professional communications!
- Doing more than is sustainable. There is always more work that can be done, and this is perhaps especially true in the Academy. More projects, more papers, more service, more outreach, more more more! While you are encouraged to challenge yourself, your wellbeing is paramount (see Section 4, Rule #1).

### 6.6. Exiting

Eventually, all ReACTERs will leave the CoReACTER (that, or the CoReACTER will cease to exist). In this section, we discuss procedures for leaving gracefully and responsibly.

#### 6.6.1. Leaving Voluntarily

ReACTERs may want to leave the CoReACTER for a variety of reasons. It may be that they no longer feel aligned with the CoReACTER’s mission, policies, or actions. Alternatively, the ReACTER may still feel committed to the CoReACTER but feel that they don’t have adequate time and/or energy to participate in CoReACTER activities.

Regardless, ReACTERs are free to go at any time, for any reason. However, we ask that you inform the CoReACTER, ideally at a group organizational meeting, at least two weeks before leaving the organization. This helps us to process and to make necessary plans (see Section 6.6.3 below).

We hope that the saying “Once a ReACTER, always a ReACTER” holds true. If, at some point after a researcher leaves the CoReACTER, they want to return, they will be welcomed back, with no discussion needed. However, given funding limitations, the CoReACTER may not be able to compensate the newly returned member, at least not immediately.

### **6.6.2. Being Asked to Leave**

We hope that this section of *The CoReACTER Manual* collects dust, that it rarely needs to be read or edited. That situation would imply that relations in the CoReACTER are healthy, that ReACTERs come and go with the natural flows of academic work. In the spirit of caution and preparedness, however, we must have a procedure in place for the unfortunate situation in which we must ask a ReACTER to leave, potentially involuntarily.

There are few circumstances that would merit a hard break in the relationship within the CoReACTER. Because of the loose, liberal definition of membership in our organization, typical pressures need not result in ruptured relationships and separation. For instance, even in the unfortunate and unlikely situation that the CoReACTER runs out of research funding and cannot financially support a ReACTER (especially a graduate student or postdoctoral researcher), the ReACTER in question need not leave the organization. Their role may need to change, and their involvement may become more limited as they pursue other avenues of research and/or work, but we will not ask a ReACTER in such a situation to leave.

Even a violation of our community principles, rules, or other norms need not lead to a researcher leaving the CoReACTER. We embrace the practice of restorative justice [1], in which we work with those who have caused harm to take responsibility for their actions, learn from their mistakes, and repair damage where possible (see Section 8 for more detail). A ReACTER who is committed to justice and personal responsibility will in general be allowed to remain in the CoReACTER, though their role in the organization may change as they rebuild relationships and undergo a process of growth.

On the other hand, a researcher who causes harm but cannot or will not take responsibility and put in the work necessary for repair does not have a place in the CoReACTER. Likewise, an individual who repeatedly causes harm without changing their behavior may not yet be capable of acting as a positive member of the community. In these cases, the CoReACTER may collectively decide to revoke an individual’s membership. This action should not be taken lightly, and care should be taken to ensure that all members of the CoReACTER have time to process and discuss the harms that have been caused thoroughly before asking a ReACTER to leave.

A ReACTER who has been asked to leave can still return at some later point. To do this, they must petition the CoReACTER and demonstrate, to the extent possible, how they have grown in ways necessary to create and maintain positive relationships and serve as a positive member of the CoReACTER community. The CoReACTER will discuss the petition and will vote on whether or not to restore the membership of the ex-ReACTER.

### **6.6.3. Before You Go**

This is a kind of checklist, somewhat similar to the one described in Section 6.4.3. If you are planning to leave the CoReACTER soon, or if you have been asked to leave, please work with your colleagues to ensure that you address the following tasks before you go.

1. **Software maintenance:** If, as part of your work in the CoReACTER, you have been involved in developing or maintaining software, care should be taken to ensure that the tools that you have worked on will be maintained. Even if you intend to continue maintaining your code after you leave the CoReACTER, it is generally a good idea to train at least one other ReACTER who is willing to step in as necessary for development and maintenance.
2. **Data:** *All* data that you produced as a part of your research should be backed up on physical hard drives accessible to the CoReACTER and on a remote cluster. If the data is unlikely to be used later, it should be placed in long-term tape storage; otherwise, storage on-disk is fine. Make sure your data are well documented and that others in the CoReACTER know what is where.
3. **Feedback:** While you should be providing feedback to your fellow ReACTERs (including EWCSS) throughout your tenure in the CoReACTER, we nonetheless recognize that departures are a good time for reflection. If possible, set up meetings with the folks you've worked closely with to provide mutual feedback. What worked? What didn't? What recommendations do you have for the CoReACTER or the people in it to continue to improve?
4. **Returning equipment:** If you have any materials owned by the CoReACTER (e.g., keys, a computer, office equipment, a hard drive, etc.), return them to EWCSS. Any computer should be wiped after it's been fully backed up.

## 7. Lab Management

### 7.1. Projects

Research in the CoReACTER is organized into projects. A project is an intellectual container, a research unit, consisting of some related ideas, goals, efforts, and personnel.

If that definition seems vague, that's because it is! Projects are a flexible concept and structure, allowing ReACTERs to organize as they need to effectively engage in team-based research. Projects are often related to funding sources<sup>13</sup>, but not necessarily. A project could be a short-term effort involving two or three folks that results in one paper (or less!), or it could be a multi-year engagement involving the entire CoReACTER.

#### 7.1.1. Starting a Project

Generally, projects are established by the CoReACTER as a collective. One or more ReACTERs proposes the new project, defining the scope and the (initial) team. If we agree that the project is a good idea and that we have the time and resources to devote to it, then the project is created!

We ask that ReACTERs only initiate projects if they have the time and resources (including physical, emotional, and mental energy) to devote to them *without harming themselves or others or significantly taking away from existing projects*. Ultimately, the assessment of time and resources is subjective and must be considered on a case-by-case basis. This is all the more reason to communicate about every new project as a team!

The CoReACTER does not own the ideas of ReACTERs, and ReACTERs are welcome to initiate projects outside of the CoReACTER. However, even in this case, we ask that ReACTERs propose new projects within the CoReACTER. This allows us to discuss how the new project may interact with and affect work within the CoReACTER. It also gives us a chance to take stock and (if necessary) redistribute work to allow for this new project to begin.

#### 7.1.2. Project Teams

A project needs a team who will execute on the ideas involved. Almost always, that team will consist of two or more people, as we encourage collaborative “team science”. Nonetheless, a one-person project team is valid.

When a project is created, the CoReACTER will discuss who will be involved and the roles that they will play. This initial team is just that: initial. We should expect that project teams will change over time, just as we expect that the scope of a project will change over time. When it comes to teams, it is important to remember that we are not in the business of exclusion. If someone wants to be involved in a project, if they have a legitimate reason to be involved (e.g., they want to learn a skill that is important for the project, or they think they will contribute a useful perspective to the project), if they have the time and resources to devote to the project, and if they act in good faith and as good community members, then they will always be welcome. This is true whether that individual is formally a ReACTER or not!

Project Teams may have “leaders” (*i.e.*, folks who play major roles in a project, who provide vision or guidance, etc.), but they do not have “owners”. Once a project begins in the CoReACTER, the project and its associated ideas are collective unless (collectively) decided otherwise.

---

<sup>13</sup>That is, a project could be defined by the ideas laid out in a funding proposal, the objectives laid out in the proposal, the research tasks undertaken to meet those objectives, and the folks paid by the particular funding source.



### **7.1.3. Suspending and Ending Projects**

Inevitably, all things must end. Just as it is with life, so too is it with projects. A project within the CoReACTER can be closed in two ways: suspension and termination.

If a project is suspended, that means that we still think the ideas involved are good, we'd still like to work on it, but we don't currently have the time or resources. If we agree to suspend a project, then the project team members should document the current status of the project as thoroughly as they can and should establish a time in the future to re-assess the project and discuss re-opening it.

If, on the other hand, a project is terminated, that means that we no longer think that the ideas are worth pursuing, or we do not think we will ever have the time or resources to devote to the project. In that case, all data related to the project should be backed up and archived, just in case it might be useful for another project down the line. Someone on the team should also be designated as the historian for the project, able and willing to answer questions about it if they come up in the future.

## **7.2. Lab meetings**

Within the CoReACTER, we hold three types of regular meetings: group organizational meetings, group research meetings, and project team meetings. In addition, we have a procedure for calling *ad hoc* ("when needed") meetings.

In this section, we detail how each of these types of meetings operate. Decision-making within meetings is discussed in the following section.

### **7.2.1. Group Organizational Meetings**

In group organizational meetings, we discuss all matters related to group management and well-being, from grants and staff hiring to meeting scheduling and social events. These organizational meetings are the CoReACTER's main structure for decision-making, which makes attendance and active participation important for all ReACTERS.

#### **7.2.1.1. Attendees**

Group organizational meetings are generally closed, including only ReACTERS. If the participation of outside parties is required for a particular item of business (for instance, to discuss a new collaboration with the potential collaborator in question), they may be invited if and only if the CoReACTER collectively agrees to invite them. Invitations are granted on a per-meeting basis. Outside parties, once invited to a meeting, have full rights to participate, including voting on matters of substance.

#### **7.2.1.2. Location**

Group organizational meetings will be held in person in EWCSS' office (Doherty Hall A207E). In-person attendance is encouraged, but to accommodate travel, family needs, health needs, etc. we will always offer virtual attendance. The video conferencing platform of choice for CMU is Zoom. We prefer Jitsi<sup>o</sup>; in addition to relying on open source code, Jitsi is somewhat more secure than Zoom (all video and audio is end-to-end encrypted), and Jitsi has better data privacy policies.

### 7.2.1.3. Scheduling

We will regularly (**once every 3-6 months**) determine a mutually agreeable time and place to hold group organizational meetings. As we do not typically expect such organizational matters to be urgent, these meetings will be scheduled to occur **once every two weeks**. Meetings are expected to last for **one hour**.

If organizational discussions are required urgently outside of the scheduled one-hour blocks, ReACTERS are encouraged to schedule an *ad hoc* meeting.

### 7.2.1.4. Agenda-Setting

**At least 24 hours** prior to a scheduled group organizational meeting, the Meeting Lead will send a preliminary agenda to all ReACTERS<sup>14</sup>. Typically, this preliminary schedule will include items that were scheduled for previous meetings but had not been discussed, as well as new topics that have been brought up since the previous meeting. Each item will be given a brief description as well as a suggested time allotted for discussion. The suggested time is not meant to create a rigid timetable, but rather to help ReACTERS keep the limited time budget in mind when planning and running the meeting.

ReACTERS will then have time to respond to the preliminary agenda, proposing additional items or suggesting to table items already on the agenda. At some point shortly before the meeting (preferably at least **one hour** before the meeting starts), the Meeting Lead will close discussion on the agenda and will send out an amended, final agenda for the meeting.

### 7.2.1.5. Meeting Procedure

Group organizational meetings will typically follow a standard pattern:

1. Greetings, introductions, and brief socialization (~5 minutes)
2. Review of the agenda. (typically less than 5 minutes) At this point, urgent items can be added or amended. We ask that ReACTERS try not to suggest large, last-minute changes and rather follow the agenda review process mentioned above.
3. The agenda is followed, with discussion led by the Meeting Lead (~45 minutes).
4. In the last ~5 minutes of the meeting, the Meeting Lead goes over the section of the agenda that has not been covered. If any items are urgent, an *ad hoc* meeting may be scheduled at this time. Otherwise, they will form the basis for a preliminary agenda for the following group organizational meeting. In addition, the next Meeting Lead will be selected.

## 7.2.2. Group Research Meetings

Group research meetings are regular meetings where we share our recent research findings. It is expected that, in each meeting, each ReACTER shares a small research update, which could include discussion of a paper or papers that they've read, a new software package that they're working with or writing, a new method or form of analysis that they're trying out, new data that they've collected or analyzed, etc. In addition to these brief round-robin updates, research meetings will hold space for longer-form discussion of new (or new-to-us!) ideas.

### 7.2.2.1. Location

Like group organizational meetings, group research meetings will be held in EWCSS' office and on Jitsi (or Zoom if needed).

---

<sup>14</sup>If the Meeting Lead is busy and unable to manage the meeting agenda, they should talk to EWCSS, who will try to help with meeting logistics.

### 7.2.2.2. Attendees

Group research meetings are open. All ReACTERs are expected to attend. In addition, meetings will be advertised to collaborators<sup>15</sup>, and any non-ReACTER who wants to attend need only reach out to EWCSS or the Meeting Lead to be given information for in-person and remote attendance.

### 7.2.2.3. Scheduling

As with group organizational meetings, we will regularly (**once every 3-6 months**) determine a mutually agreeable time and place to hold group research meetings. These meetings will be scheduled to occur **once every week**. Meetings are expected to last for **ninety minutes**.

For additional research discussions, ReACTERs are encouraged to set up informal discussions or carry discussions into project team meetings. *Ad hoc* meetings should only be called to discuss research matters if they are urgent and affect all or nearly all of the CoReACTER.

### 7.2.2.4. Agenda-Setting

Roughly **three days and no less than two days** prior to a group research meeting, the Meeting Lead will communicate with all scheduled presenters, who will communicate what they are interested in presenting to the Meeting Lead, along with any related materials<sup>16</sup>. If, for some reason, a presenter cannot attend a group research meeting or cannot present, at this time, the Meeting Lead should make efforts to solicit alternative presentations.

At some point prior to the start of the meeting (ideally **one day** before the meeting, but at least **one hour**), the Meeting Lead will finalize the agenda and communicate the finalized agenda to the CoReACTER and all other attendees.

### 7.2.2.5. Meeting Procedure

Group research meetings will typically follow a standard pattern:

1. Greetings, introductions, and brief socialization (~5 minutes)
2. Review of the agenda (typically less than 5 minutes). ReACTERs who have longer updates can at this point propose to present if time remains at the end of the meeting.
3. Brief updates (roughly 5 minutes per ReACTER, for a total of roughly 15 minutes). Each ReACTER describes what they've been working. The goal of these updates is to keep everyone relatively informed regarding each others' research and make sure that everyone is regularly getting feedback on their research. These brief updates may spawn more thorough, external conversations.
3. The agenda is followed, with discussion led by the Meeting Lead (45 minutes).
4. If time allows, additional presentations can be shared (time to be determined).
5. In the last ~5 minutes of the meeting, the next Meeting Lead and presenters are selected, with a preliminary discussion of what might be presented.

### 7.2.3. Project Team Meetings

Project team meetings are regular meetings that aim to drive forward a particular project (see Section 7.1). These meetings could involve sharing research updates, as in group research meetings, but could also involve future planning or concentrated work. The scope of project team meetings are based on the needs and desires of the particular project team.

---

<sup>15</sup>Question for later decision: should these meetings be publicly advertised?

<sup>16</sup>For instance, if the presenter is presenting a recent paper, they might share a copy of the manuscript so that the CoReACTER can read it.

### 7.2.3.1. Attendees

In general, project team meetings are open to the CoReACTER. If the project involves collaborators outside of the CoReACTER, those external collaborators should be invited to all project team meetings.

### 7.2.3.2. Scheduling

Project teams should decide as a group how often they want to meet. As a starting point, we advise that teams for active projects meet **once every two weeks**. Meeting duration is also up to the discretion of the project team, but **one hour** is a good starting point.

### 7.2.3.3. Meeting Procedure

We do not dictate how project team meetings are run. However, we provide some general guidelines to help meetings remain positive, healthy, and helpful.

1. Set intentions for meetings. Whether you set a formal agenda weeks ahead of time or decide on a plan the day of the meeting, do not meet unless you know what you're hoping to get out of the meeting.
2. Be consistent. While free-form, spontaneous meetings have their place (see Section 7.2.4), project team meetings should have a rhythm. They should be part of your normal research routine, like reading papers and checking on calculations.
3. Share the burden. No one ReACTER owns a project, nor does any one individual hold all of the responsibility for it. Distribute organizational tasks like scheduling, note-taking, discussion leading, etc. equitably, perhaps rotating from meeting to meeting.
4. Accommodate your fellow team members. Try to schedule so that as many people who want to attend can. Provide multiple options for attendance and participation. Ask what folks need to engage in the meeting!

### 7.2.4. Ad Hoc Meetings

Not everything can happen on a nice, pre-arranged schedule. Sometimes, something urgent comes up and demands discussion as soon as possible. Less severely, maybe you just ran out of time in a group organizational or research meeting, and there are a few lingering points that you want to get through before the next meeting.

Regardless of the motivation, an *ad hoc* meeting of the CoReACTER could be the answer. To propose an *ad hoc meeting*, you need to communicate:

- The rationale for the meeting (Why does this need to be an *ad hoc* meeting? Why couldn't this wait until the next scheduled meeting?)
- A plan for the meeting, including an expected duration and agenda
- The expected outcomes of the meeting, e.g. a decision on an important topic, or finalizing a paper before submission

In general, *ad hoc* proposals should be communicated to the CoReACTER as a whole. Once proposed, we hold a quick and informal vote to determine if this meeting needs to be held. If there is significant interest<sup>17</sup>, then we will schedule the meeting, trying to find a mutually agreeable time and date. Most *ad hoc* meetings are meant to be held for urgent or relatively time-sensitive matters and so should not be planned more than one week in advance.

---

<sup>17</sup>For now, this means that at least three people want this meeting to occur.

The proposer of an *ad hoc* meeting will serve as the Meeting Lead. If multiple ReACTERs propose an *ad hoc* meeting together, then they should choose among themselves who should serve as Meeting Lead.

### 7.2.5. Running Meetings

Each meeting will be run by a Meeting Lead. With the exception of *ad hoc* meetings, where the person/people to call the meeting serve as Lead(s), the Meeting Lead will be determined at the close of the previous meeting. The Meeting Lead has the following responsibilities:

- Setting the meeting agenda
- Leading discussion during the meeting
- Setting the pace of the meeting, *i.e.* directing the flow of conversation to ensure that as much of the agenda as possible is covered without stifling fruitful debate and discussion
- Designating note-takers and other organizational roles within the meeting
- Communicating a meeting summary (consulting any notes taken in-meeting) to the CoReACTER and (where appropriate) our collaborators following the meeting, including action items

Meeting Leads will rotate from meeting to meeting to ensure that meetings are not dominated by any one perspective. Likewise, Meeting Leads should take care to select note-takers and other assistants such that the distribution of organizational labor is distributed equitably.

### 7.3. Decision-making

As an anti-oppressive research organization, the CoReACTER is dedicated to community-centered, collective, and equitable decision-making. Here, we outline procedures that help us to live up to our ideals and aspirations while making the best decisions possible.

1. *Decisions are made democratically.* No one ReACTER has a greater ability to direct CoReACTER business than any other; in general, each individual is given an equitable opportunity to voice their opinions and to vote to decide on appropriate courses of action.<sup>18</sup>
2. *Involve as many stakeholders as possible.* When making a decision, we consider who will be affected. If possible, we try to call those stakeholders and affected parties into our decision-making process; this could involve consulting with external parties ahead of a meeting or inviting external parties to a CoReACTER group organizational meeting, where they will have full rights of participation. If a decision will disproportionately affect some ReACTERs or stakeholders, we try to privilege those affected perspectives in our discussions.
3. *Consensus is the goal.* Although we operate as a democracy, we are not aiming for simple majority rule. Ideally, we want to reach agreements that satisfy as many ReACTERs and other stakeholders as possible. This means that, in our arguments, we should try to win over *everyone*, and we should constantly be on the look-out for opportunities to compromise on our ideas (though not on our ideals!).
4. *Until it comes time to decide, discuss and listen without judgment.* The beginning of the decision-making process is the time to freely share ideas. Even if you come into the discussion with strong opinions, do not put down other ideas or push your own ideas. Rather, listen and carefully consider each person's perspective. Once everyone has said their piece and all of the ideas are in the air, then we can start to discriminate between ideas and find what might be the best course of action.

---

<sup>18</sup>EWCSS, as PI, is professionally and in some respects legally responsible for the work that takes place in the CoReACTER. Because of this imbalance in power, risk, and responsibility, there may be circumstances where EWCSS must make a decision unilaterally, without reaching consensus with the rest of the CoReACTER. Even in these cases, they commit to communicating regarding these decisions as much as possible with the CoReACTER, and to holding space for disagreement and dissent.

5. *Opinions are not only spoken.* Not everyone will always speak their mind openly. Some folks may be naturally less open about their feelings and opinions, and others may specifically be uncomfortable voicing an opinion on certain topics. Given these possibilities, it is imperative that we not treat a lack of dissent as consent. When making decisions, consider body language and facial expression. Consider who is and who is not talking. If everyone appears open and enthusiastic, that's a good sign that the discussion is on the right track. If not, there may be concerns or problems with the proposed plan that have not been discussed yet.
6. *When possible, do not make decisions in haste.* Making good decisions takes time and thought! If a decision can be reached easily without significant disagreement, then it's appropriate to seize the opportunity and finalize immediately. On the other hand, if folks are butting heads or if consensus feels far away, be willing to give the decision more time, or to step away and reconvene after everyone has had time to reflect and devise more creative solutions. Some decisions will truly be urgent and will not have the luxury of second (or third, or fourth) discussions, but don't create urgency if it isn't there!
7. *Be willing to iterate.* The CoReACTER might not get the decision right the first time. It can be tempting to commit to our decisions, to pick a path and not stray from it. We must resist this urge and, when necessary, reassess and revise our previous decisions. In doing so, we should consider who has been affected by our initial decision and how we can be accountable to them while potentially changing our own course of action.

## 7.4. Training and Mentoring

### 7.4.1. Mentoring Relationships

Most ReACTERS will be in training positions, whether they are students or postdoctoral fellows. The top priority for trainees is to learn and gain skills. Therefore, upon entering the CoReACTER, students and postdocs should work to identify mentors within and outside of the organization. Note that we say “mentors”, not “mentor”. It is unlikely that a single person can meet all of a trainee's personal, professional, and academic needs. Moreover, if for some reason one's mentor is unavailable (for instance, because of illness), then one would be bereft of necessary resources. We encourage ReACTERS to seek out multiple mentors that complement each other's skills and perspectives and to discuss with these mentors what you hope to gain from them (perhaps using your answers to the questions discussed in Section 6.4.4 as a guide). Note that the mentor-finding can be an ongoing process; you don't need to find all of your mentors at once!

Although training and mentoring are especially important for students and postdocs, we emphasize that teaching is not a one-way activity, with knowledge being passed from the “expert” to their pupils. Such a model of education and training would be antithetical to the CoReACTER, which aims to deconstruct academic hierarchy. Moreover, assuming that mentorship and teaching flow in one direction essentially discards many opportunities for more “senior” members of the CoReACTER (e.g., faculty) to grow! If someone has asked you to be a mentor, you should still expect to learn from them, and you should feel comfortable coming to them for help. Stay humble (Section 3) and ask for what you need!

### 7.4.2. Skill Training

If, in the course of your research, you find that you need to use a new-to-you piece of equipment, **do NOT** simply begin using said equipment. This goes for computational resources (e.g., a high-performance computing cluster) as well as for laboratory equipment (e.g., a glove box or spectrometer).

ReACTERs must be trained on each piece of equipment before they use it in their research. To streamline this training process, each piece of equipment will have a designated “point person” who will be responsible for training. The point person should develop a standard curriculum that demonstrates the basic functions of the equipment, best practices, and (where appropriate) necessary maintenance to new users. Prior to beginning training, they should also meet with the prospective user to learn what they aim to use the piece of equipment for. If possible, the point person should also show the new user how to use the equipment in a way that is similar to their planned use case(s).

Only after the point person has completed training and has verified that the new user is capable of safely and responsibly operating the equipment can the new user use the equipment independently. Training activities should be recorded in the CoReACTER Anytype (Section 10.1.3) for institutional memory purposes.

## **7.5. Communication and Information Sharing**

### **7.5.1. Within the CoReACTER**

There are no secrets within the CoReACTER.<sup>19</sup> What one ReACTER knows, we all should know, or at least we should all be *able* to know.

The structure of CoReACTER meetings (Section 7.2) makes the internal exchange of information particularly facile. Research updates will be shared consistently in group research meetings (Section 7.2.2) and project team meetings (Section 7.2.3), which are open to the CoReACTER. Similarly, updates about funding, group decision-making, collaborations, and more will be discussed in regular group organizational meetings (Section 7.2.1). This structure means that one would have to go out of one’s way to hide information from other ReACTERs, which you absolutely should not do (follow our first guideline; see Section 5).

For information that would not normally be shared in a research, project, or organizational meeting, ReACTERs are expected to be proactive about sharing information within the CoReACTER. As an example, say that you hear about a new graduate fellowship or a postdoc opportunity. This may not be relevant to the group as a whole, but if you know other ReACTERs who might be interested in applying, you should identify those individuals and pass the information along.

### **7.5.2. Outside of the CoReACTER**

Information sharing outside of the CoReACTER involves more compromise between our principles and guidelines. While we are a research organization that works to promote open science (Section 9), we also operate under a model of collective decision-making (Section 7.3). The general guiding principle in this case is that information will *eventually* be shared openly; however, the question of when to share must be made collectively, and until that decision has been reached, information should be assumed to be privileged.

We take this course because information, once shared, cannot be un-shared; there is no closing the Pandora’s Box of an idea. In deciding whether or not to share information with individuals or groups outside of the CoReACTER, the burden of proof should be on those wishing to maintain privacy (*i.e.*, the default decision should be that information be openly shared), but the weight of the decision means that deliberation is still needed.

---

<sup>19</sup>There will inevitably be exceptions to this statement, cases where not everyone should know something. In particular, one can imagine that some personal information may be shared on a need-to-know basis, being otherwise secret so as to protect individual privacy. However, in general, all information should be considered open to all ReACTERs.

## 7.6. Funding

The CoReACTER needs to maintain sufficient funding to continue operation. This makes us directly tied into Capitalist modes of value and production, intrinsically creating ethical conflict. Obtaining and managing funding also burdens the CoReACTER with ethical responsibilities. Here, we briefly outline how we go about pursuing funding, publicly disclosing our funding, and managing funding once we've obtained it.

### 7.6.1. Pursuing Research Funding

The funding process is morally fraught, and there are very few options for sustainable research funding that do not require some kind of ethical compromise. Understanding this, we are generally open to any funding source, though we try to move towards funding sources that are as aligned with our principles (Section 3) as possible.

There are, however, some areas that we will not consider because they are radically opposed to our principles, the good of the scientific community, and/or the wellbeing of society. The CoReACTER will not pursue or accept funding or other resources:

- From military sources (e.g., the United States Department of Defense), organizations whose major product or service relates to military technologies (i.e., members of the military-industrial complex), or from projects with direct military applications
- For projects that will further the extraction and utilization of fossil fuels, or that otherwise are likely to significantly increase greenhouse gas emissions and/or accelerate anthropogenic climate change. Note that this includes so-called “AI” companies like OpenAI and Anthropic
- That impede us from sharing our finding freely and openly, or that limit our ability to effectively collaborate

We note that “other resources” includes computational resources. This means that we will not, for instance, pursue access to or access supercomputers owned and operated by the Department of Defense.

All ReACTERS are encouraged to pursue funding. CMU allows postdocs to serve in lead PI and co-PI roles on grants. If you are a postdoc who has a good idea for a new research direction, you should expect to be a PI! Students, including doctoral students, unfortunately cannot serve in PI roles. However, they are encouraged to pursue external funding in the form of fellowships, to suggest grants and projects for the CoReACTER to pursue under different PIs (i.e., EWCSS or a postdoc), and to participate in grant-writing.

Regardless of who generates the idea for a funding proposal or who will serve as lead PI, all grant-writing and funding submissions must be approved by the CoReACTER. Typically, funding-related decisions will be made during group organizational meetings (Section 7.2.1), though in exceptional circumstances, an *ad hoc* meeting (Section 7.2.4) may be called for this purpose.

### 7.6.2. Disclosing Funding

It is important to be transparent about our funding. Money creates real and potential conflicts of interest, and an organization's finances are a reflection of their values and priorities.

The CoReACTER will always publicly disclose our funding sources on our website (<https://coreacter.org/funding/><sup>o</sup>). In line with standard publication practices, we will also disclose our funding and all related conflicts of interest in our publications (including preprints and blog posts) and public presentations.



### **7.6.3. Managing Funds and Making Purchases**

Within CMU, a grant's lead PI is responsible for project finances. For day-to-day operations, we follow this pattern. A lead PI can make small financial decisions (generally anything under \$1,000 is considered small) without consulting the CoReACTER or engaging group decision-making, and they will be expected to make orders, manage purchases, and prepare and amend grant budgets, as necessary.

CMU provides regular (usually monthly) reports on project finances. Grant lead PIs should present these reports to the CoReACTER during group organizational meetings. Group organizational meetings should also be used to make decisions regarding large (>\$1,000) purchases.

## 8. Accountability

Inevitably, we're going to screw up as individuals and as an organization. Thoughtful researchers that we are, we own our mistakes and, once they occur, seek to minimize the harm to our victims and the broader communities. Here, we outline a collection of practices keep us accountable in healthy and (hopefully) helpful ways.

### 8.1. Restorative Justice

Restorative justice is a transformative legal concept that seeks to replace (or, in some cases, supplement) punitive responses to wrongdoing (e.g., prison time, or in an academic context, academic probation or expulsion) with reparative, dialogue-based responses that empower the victim and allow the perpetrator of harm to be re-integrated into their community. Restorative justice concepts are not new *per se*, as many cultures around the world have practiced and still practice reparative, community-centered methods of justice, but modern practice of restorative justice began in the 1970's as a response to the apparent failings of the Western criminal-legal system. For a detailed review of restorative justice practices, their utility, and possible shortcomings, see Ref. [1].

It is our goal to maintain and, over time, strengthen the CoReACTER community and to create strong bonds with external communities. We therefore embrace the practice of restorative justice. There may be forms of harm and wrongdoing that cannot be forgiven, and there may be cases where we must break a bond of community and ask individuals to leave the CoReACTER, we will strive to repair relationships wherever possible.

The core of our restorative justice process looks something like this:

1. An individual or a group comes forward and explaining what has happened and how they or others in their community have been harmed
2. The CoReACTER creates a space for *intentional, facilitated dialogue*.<sup>20</sup> This space will allow for the direct participation of any alleged wrongdoers and any victims, which can include direct victims, bystanders, and members of larger communities. In this dialogue:
  1. The facilitator describes what has occurred and explains how the wrongdoing has caused harm, broadly defined
  2. Victims come forward to explain their perspectives and further describe how they were affected
  3. The wrongdoer explains what they did, as well as their intentions and motivations for doing so
  4. The wrongdoer accepts fault and responsibility for their actions and, ideally, apologizes to those harmed (see Section 8.2)
  5. We engage in a discussion of the root causes of the wrong and its harm
  6. The victims have the opportunity to forgive the wrongdoer, though they have no obligation to do so
  7. We collectively consider appropriate steps forward to repair and heal from the wrong
3. The wrongdoer follows the steps suggested in the facilitated dialogue, righting wrongs (to the extent possible) so that they may ultimately be reintegrated into their community/communities

Some keys to productive restorative justice practice:

- The wrongdoer as a person is separated from their actions. The actions may be wrong, but the person is *not* inherently bad or wrong.

---

<sup>20</sup>This can be thought of as a particular type of *ad hoc* meeting; see Section 7.2.4

- Restorative justice doesn't work by force. Victims need to move past a desire for retribution, and wrongdoers must open themselves up to listening to how they've hurt others, genuinely apologizing, and putting in the work to repair relations.
- To the extent possible, we should be forward-looking. While we should not forget what has happened, we should emphasize what we can change or improve going forward.

## 8.2. Apologizing

In Section 8.1 above, we state that *apology* is a major part of our restorative justice framework. But how do we give an apology that is in line with our values?

Here, we draw from the *CLEAR Lab Book*<sup>o</sup> and The Four Parts of Accountability & How To Give A Genuine Apology<sup>o</sup>, a 2019 blog post by Mia Mingus.

To make an apology:

1. Say "I'm sorry."
2. Name the harm/hurt
3. Name the impact (not the intention)
4. Name the actions
5. Commit to not doing the harm again

— Mia Mingus

When you are apologizing, in the context of a restorative justice dialogue or outside of it, keep the following in mind:

- You are not being attacked, and any accusations against you or wrongs you've committed do not make you a bad person.
- Focus on listening. Do not interrupt others, especially your victims. Focus your attention on understanding how your actions were harmful.
- Try to be empathetic, compassionate, and humble.
- If the other person/people consent to continued discussion, ask what you could have done instead
- Your apology will mean much more if you change your behavior than if you do not. This will be hard, but you must try!

## 8.3. Call In, Call Out

To "call <something> out" is to publicly criticize or fault that thing. Calling out has its place — it asserts community norms and can quickly stop negative behaviors. However, calling someone out can rupture relationships and can alienate an individual from their community.

When possible, we prefer *calling in*. When we call in, we (usually privately) explain to someone how their behavior is affecting us or others, letting them *into* another perspective (hence the name). Calling in can be used for minor harms or simply when there are differences in understanding or when norms are unclear (e.g., because different ReACTERs grew up in different cultures).

Key to a call-in, as in our restorative justice practice, is that we are focusing on *behavior*, not people. We also assume best intentions — we call in to help others gain understanding, assuming that they do not want to cause any harm.

## **8.4. Community Involvement**

What happens if a ReACTER is harmed by someone outside of the organization? What happens if our actions negatively affect those outside of our immediate community?

In principle, restorative justice calls for the involvement of representatives of all those affected by the wrongdoing. In honor of this notion, we should take effort to identify all those who have been affected and to bring them into our process. Doing so, we should be sensitive to their different needs and practices. Restorative justice remains relatively uncommon; some folks might not be willing to enter into dialogue with those who hurt them or with those they hurt, and some might want retribution. Moreover, it's possible that some will be uncomfortable with the idea of public or semi-public discussion, depending on the nature of the offense.

When involving individuals outside the CoReACTER, we should offer our approach but decide on a method for healing collaboratively. With that said, if we believe that others' methods of reconciliation and justice are harmful, we do not need to involve ourselves.

## **8.5. Auditing**

Not all wrongs have obvious victims or victims that can/will speak up. For instance, if we are harming the environment or contributing to pollution, who should call us in? Who holds us accountable?

The answer is: we need to. We commit to regular (e.g., annual) self-auditing. In this audit, we should review our group practices and consider how they align or do not align with our principles, rules, and guidelines. We should also consider how our actions may have unintentionally harmed others. In addition to the environmental considerations mentioned above, consider our participation in societal systems of oppression and toxic academic patterns.

The point of these audits is not to make ourselves feel bad, but rather to constructively identify areas for improvement and growth on an organizational level.

We have not yet performed an audit. When we do so, we will need to consider what to do with our audit's findings. Do we publish them openly and explain how we want to do better in the future (making us more publicly accountable)? Do we want to set up a committee or project team to track progress towards the goals we've identified? These are questions for future group organizational discussion.

## 9. Open Science

The CoReACTER is dedicated to open science, but what does that actually mean? “Open science” is a broad and vague term which means different things to different folks (see [2]). For us, when we speak of “open science”, we mean that the practice of science should be transparent; that there should be minimal barriers to access the outputs of scholarly scientific work; and that science and research should be practiced within communities that extend beyond the boundaries of a research group, a university, or academia as a whole.

Here, we briefly describe some of our positions on open science and related practices. We emphasize that this page is meant only as an introduction to this discussion!

### 9.1. Open Source Software

Much of our research involves developing scientific software. In addition to building and contributing to large-scale software projects, we frequently write code for data generation, analysis, and visualization. As such, open software is a significant part of our open science mission!

We value open source software because it lowers the barrier to use and modify computational tools, potentially broadening the impact of our research efforts. Open source software can also lead to better, more correct code! Because our software is transparent, other researchers and programmers can inspect, test, and modify the code that we write, potentially finding bugs and optimizations that we missed or adding new features that are beyond our capacity.

We commit to making all software developed by the CoReACTER available open-source. Currently, we use GitHub<sup>°</sup> to host our software. We prefer restrictive licenses such as the GNU General Public License<sup>°</sup> that ensure that derivative software projects remain open-source, though we are willing to use more permissive licenses such as the MIT license<sup>°</sup> or The Unlicense<sup>°</sup>. We will only spend our time contributing to other software projects if they are open-source, and we try to use open-source codes for our research whenever possible (see Section 10.1).

### 9.2. Open (and FAIR) Data

In the age of data science, data is power: power to discover patterns, power to create models, power to accelerate research, and more. As we are committed to mass empowerment, it is only fitting that we in the CoReACTER want to share data widely. Open data further improves the reproducibility of scientific research.

When we generate data that we believe may be valuable for future research, we publish that data openly under permissive Creative Commons<sup>°</sup> licenses. But making data publicly available is only the first step. We understand that, to maximize the benefit that data can have on the scientific community, it should be **FAIR**[3] — **f**indable, **a**ccessible, **i**nteroperable, and **r**eusable. To this end, we develop and contribute to public data infrastructure, such as the Materials Project<sup>°</sup>, and we work with our collaborators around the world to make the data that we generate as useful and easy-to-use as possible.

### 9.3. Open Access

Many scientific publications are closed-access: institutions (e.g., libraries and universities) or individual readers must pay for the privilege of reading. This slows the spread of information, particularly to those who lack resources and privilege. The CoReACTER strongly believes that science should serve the public good. As a consequence of this belief, we hold that published

scientific outputs such as journal articles, book chapters, and whitepapers should be publicly available at no cost.

For more detailed information about our approach to open access publication, see Section 13.3.

## 10. Software

Currently, CoReACTER research is more computational than experimental. We therefore put a lot of time thinking about software. In this section, we describe some of our thoughts and practices related to software. In particular, we discuss our preferred software suite, our interactions with third-party software, and our policies regarding software development.

### 10.1. CoReACTER Stack

CoReACTER is a research collective that attempts to empower individuals while thinking and acting collectively. This creates tension with regard to choices regarding software. The individualistic lens says that each ReACTER should be able to choose a software suite that works for them, while the collective lens favors standardization to facilitate collaboration and interoperability.

In writing this section, we lean softly on the side of collectivism. The software described in this section is strongly preferred over equivalent or near-equivalent alternatives. If you don't have a strong reason to use something else, please adhere to the software choices that we list here. If you strongly believe that a particular project or application demands an alternative stack, please communicate this in a one-on-one meeting with EWCSS, a group organizational meeting, or a project team meeting, as appropriate.

#### 10.1.1. Programming languages

For most programming tasks, we use Python<sup>o</sup>. This is partly historic — many of the scientific software tools that EWCSS and collaborators have previously developed are written in Python — and partly an acknowledgment that Python is presently the most widely used language for high-level scientific software. For package management, we further encourage the use of either Anaconda/miniconda<sup>o</sup> or mamba<sup>o</sup>. EWCSS learned how to code many years ago from *Learn Python the Hard Way*<sup>o21</sup>, but more recently we've seen the *W3Schools* tutorial<sup>o</sup> being mentioned online.

Recently, Julia<sup>o</sup> has emerged as an alternative high-level language to Python. Julia is in general much faster than Python and has some language features that are highly beneficial for scientific programming, including strong dynamic typing with multiple dispatch for parametric polymorphism, native, high-level functionality for working with matrices, excellent support for parallelization, multithreading, and automatic differentiation, and the ability to quasi-natively run C code. If a lot of that doesn't mean anything to you, just know: Julia is easy to use, like Python, but is tuned for high-performance computing and scientific programming tasks. Julia is our second-favorite language, and we encourage ReACTERS to write code in Julia, particularly if it needs to be fast and if it doesn't depend heavily on existing Python libraries. There are plenty of Julia resources here<sup>o</sup> if you want to learn!

When we *really* need speed, we prefer to write in C. C is an extremely well-supported language with decades of compiler optimization. It's harder to learn and harder to use properly than Python or Julia, but your end result will likely be much more efficient, on average. Note that we do not encourage the use of either C++ or Fortran. C++ offers some benefits over C, such as templates for parametric polymorphism, but the increased complexity often outweighs the benefits. Fortran is an amazing language that (for tight numerical tasks) is often even faster than C, but no one in the CoReACTER at present has expertise in Fortran, so if you run into problems, you'd largely be on your own. And we don't want anyone to have to struggle alone!

---

<sup>21</sup>Back then, it was a free online resource!

We do not currently have a recommended resource for learning C; if you know of one, please suggest it and consider adding it to this guide.

For small command-line utilities and shell scripting, we encourage Zsh<sup>◦</sup> but also accept Bash<sup>◦</sup>, which remains more widely available on high-performance computing platforms. If you use Zsh, please consider using Oh My Zsh!<sup>◦</sup> to manage your configuration. There’s a nice Bash Guide<sup>◦</sup> since Bash and Zsh are so similar, this will help you regardless of which you use.

### 10.1.2. Software Engineering

Git<sup>◦</sup> is the CoReACTER version control system of choice. Git is lightweight, fast, and supports many features that are necessary and helpful for collaborative programming. If you’re unfamiliar with Git, you might want to check out some of the following resources:

- This list of resources from GitHub<sup>◦</sup>
- W3Schools Git Tutorial<sup>◦</sup>

Git beginners may also want to use a graphical user interface (GUI) rather than use Git directly through the command-line. We do not recommend any one particular Git GUI, but in line with our preference for free and open source software, you might want to consider:

- GitHub Desktop<sup>◦</sup>
- ungit<sup>◦</sup>
- lazygit<sup>◦</sup>
- Magit (for Emacs users)<sup>◦</sup>

For writing text and code, we recommend using VSCodium<sup>◦</sup>. VSCodium is a free/libre and open source distribution of Microsoft’s VSCode, with (almost) all of the nice features and none of the tracking or telemetry. It’s fast, it’s extensible, it can easily handle any (common) language you throw at it. There’s little not to love.

### 10.1.3. Organizational Software

We use Anytype<sup>◦</sup> as our knowledge base. Anytype is open source<sup>◦</sup>. It’s extremely flexible; while an Anytype “Space” will always use a graph-based representation under the hood, users can define arbitrary types and relations, making it easy to organize your ideas however makes the most sense for you. As the user, you have the power to host all of your data locally, but you can also host it end-to-end encrypted on the Anytype servers. For personal notes that are unrelated to the CoReACTER, you’re welcome to use whatever knowledge management system you want. But for anything CoReACTER-related (even your own personal notes), please use Anytype!

We use Zotero<sup>◦</sup> as our reference manager of choice. We choose Zotero because it is open source<sup>◦</sup> and because it is easy to connect with common web browsers<sup>◦</sup> for single-click storage of an interesting document. Through CMU, we get unlimited cloud storage of papers for free!

For messaging within the CoReACTER, we use Zulip. Our Zulip organization is <https://coreacter.zulipchat.com><sup>◦</sup>. Because we are an academic research organization, the CoReACTER has a free “Standard” account, which gives us access to nice features like unlimited search<sup>22</sup> (the normal “Free” tier for Zulip stores only the previous 10,000 messages). However, direct messaging platforms like Zulip are not appropriate for storing and maintaining institutional ideas. If you’re talking about something important on Zulip — if you make a big breakthrough on a project, if you find a useful resource, etc. — please put it on the Anytype!

---

<sup>22</sup>TODO: need to request sponsorship



The CoReACTER maintains a (for now modest) social media presence. Currently, we engage only on BlueSky<sup>°</sup>, an open source<sup>°</sup>, federated platform based on Twitter. If you're on BlueSky, please consider following and sharing our posts to spread awareness of what's going on in the CoReACTER!

#### 10.1.4. Chemical Simulations

For molecular density functional theory (DFT) calculations, we prefer ORCA<sup>°</sup>. ORCA is not open source, but it is free for academics, making it one of the most accessible quantum chemistry codes available. It's also actively developed, with some state-of-the-art features. We also use Q-Chem<sup>°</sup>, which requires a paid license. Where possible, we use QuantumEspresso<sup>°</sup> for plane-wave DFT calculations. It's free, it's open source, and it has many desirable features built in!

When we run molecular dynamics (MD) simulations, we primarily use OpenMM<sup>°</sup>. OpenMM is open source, and it's written to take great advantage of GPU acceleration. For this reason, it's also a natural choice for running simulations using machine learning interatomic potentials (MLIPs)<sup>°</sup>.

The CoReACTER uses and contributes to the Materials Project<sup>°</sup> (MP) software suite. This includes pymatgen<sup>°</sup>, a general-purpose analysis toolkit for chemistry and materials science; custodian<sup>°</sup>, a just-in-time error handling system; and emmet<sup>°</sup>, a framework for building chemical ontologies, databases, and APIs.

For high-throughput computational workflows, we currently prefer QuAcc<sup>°</sup>, developed by the Rosen Group<sup>°</sup> at Princeton University. QuAcc relies on the Atomic Simulation Environment<sup>°</sup>.<sup>23</sup>

#### 10.1.5. Experiments

TODO: THIS

#### 10.1.6. Other Chemical Software

CMU offers free licenses for ChemDraw<sup>°</sup>, which we use for creating chemical diagrams. This is paid, non-free software, but unfortunately current open source/free alternatives are severely lacking.

## 10.2. Developing Software

### 10.2.1. Our Code, Their Code

The world of software is rife with packages that reinvent the wheel.<sup>24</sup> Before you make a new GitHub repository or write any code, consider if a new package is actually needed. Search online, on sites like GitHub and in the scientific literature, and see what (free and open source) packages are out there.<sup>25</sup> Try them out. Can you install the code? Does it work? Does it do what you need it to?

Often, it'll be the case that there isn't a piece of software that does *exactly* what you want to do, but there are codes that come close. Many developers and maintainers welcome feature requests and/or pull requests! Consider reaching out by e-mail or through an issue (for repositories on GitHub). In your communication, detail what you're trying to do, what a solution would

---

<sup>23</sup>Two things here: 1) we should have sections for setting up these codes. 2) We should play around with different back-end systems for QuAcc. Do we want to stick with jobflow?

<sup>24</sup>For instance, check out this list of workflow engines<sup>°</sup>.

<sup>25</sup>If there are paid packages in the space that you're working on but not open source alternatives, that may be a sign that the code that you're planning to write is very much needed.

look like, and indicate if this is something that you could do on your own or if you would need help with your implementation.

If you're sure that your needs are not met by existing codes and that it would not be possible to include your desired features into any existing codes (e.g., it's too technically difficult, the developers don't have space to help, etc.), then making a new package is the correct choice.

The next step is to figure out your toolkit. What language are you going to write in (see Section 10.1.1 for the CoReACTER's stance on programming languages)? Are there standard interfaces, datatypes, or libraries for the types of things your code will need to do? For instance, if you're working with reaction networks and need a graph representation, can you use existing graph libraries, or do you really need to develop your own data structure(s)? Here you're again trying to stand on the shoulders of giants as much as possible, and you're also trying to prevent standard proliferation<sup>◦</sup>.

### 10.2.2. Code Style

We want our code to be readable by others. You shouldn't need to be a ReACTER to contribute to a CoReACTER code base! For this reason, we mostly adhere to existing style guides and community standards.

For Python, we mainly follow the Google style guide<sup>◦</sup>. We use black<sup>◦</sup> for auto-formatting code,<sup>26</sup> One area where we don't adhere completely to the Google standard is in *type annotations*. Python is a dynamically typed language built around duck typing<sup>◦</sup>. It was never really designed for a strong, modern type system, and the solutions that have been developed so far (including type annotations and mypy<sup>◦</sup>) leave much to be desired. While we encourage type annotations in function signatures as a form of documentation to indicate function behavior, we *do not* require the use of type annotations in general.

For Julia, we generally try to adhere to the official style guide<sup>◦</sup>, and for C, we mainly stick to the Linux kernel guidelines<sup>◦</sup>.

### 10.2.3. Software Maintenance

As a general rule of thumb, ***if you write it, it's your responsibility***. Think of your code like a child. You created this wonderful thing, and now you need to take care of it until and unless it can take care of itself (unlikely for a software package) or it meets its end.

That rule of thumb applies broadly. If you make a new package, you're responsible for the package. If you add a new feature to an existing package, you're responsible for it. Where possible, each developer should indicate what files, features, etc. they worked on so that users and other developers/maintainers know who to come to for help!

The exception to this rule is when someone else has agreed to be responsible for your code. Often, for small code bases, it's more efficient for one person (or a small team) to have knowledge of all of the code written and be responsible for maintaining it.

Regardless of the maintenance structure that you take for your (our) code bases, keep the following in mind:

1. The `main` branch is sacred. Code should only be added to `main` if it meets our standards, including our coding style (see Section 10.2.2 above) and the following guidelines. Less polished code can be made public in non-`main` Git branches.

---

<sup>26</sup>TODO: set up basic configs for e.g., black and pylint.

2. *Everything* should be documented. Classes/structs/data types should be described in terms of what they're meant to do and how they're meant to be constructed. Each publicly exported (*i.e.*, non-helper) function should have a docstring describing its inputs, outputs, and behavior.<sup>27</sup>
3. *Everything* should be tested. 100% test coverage, meaning that every line of code is evaluated at least once in a test, is the goal. When writing tests, try to consider “normal” cases, rare/edge cases, and errors. In general, we prefer unit testing<sup>◦</sup>, but end-to-end testing<sup>◦</sup> might be more appropriate in some cases.
4. Triage. If the code that you maintain is public-facing, you'll probably have users submitting issues, questions, bugs, and feature requests at some point. This is a good thing — it means that people are using or want to use your code! But it's easy to get overwhelmed by the labor of software maintenance. Try to stay on top of these requests, but triage to make sure that you're only responding to them when you need to. A bug that's preventing your code from being installed or from operating as expected is really bad, and you should put fixing such a bug high on your priorities list. On the other hand, a feature request might be considered a low priority; if/when you have time, you can work on it, but you don't need to jump on it right now.
5. Set milestones for releases. On a similar note to point 4 above, you want to make sure that your work in software development and maintenance has well defined stopping points. Set regular goals for yourself and your collaborators. Try to work towards those goals regularly. When they're met, make a new release, think about what should come next, and then step away from your code for a while to give yourself a well-earned break.
6. Have an exit plan. No one will stay in the CoReACTER forever. Eventually, every developer and maintainer will need to move on to other projects. Try to plan ahead and train others in using and developing your code significantly before you're planning to leave. This is particularly important if you've been the sole or the main developer of a code.

#### 10.2.4. Licensing

As discussed in Section 9.1, we use open source licenses for all software developed by our group, generally preferring licenses that require derivative code to remain open source.

It is important to select an appropriate license early. Once a licence is chosen, it binds the software, and it can be legally tricky to change a license once any code has been written under that license.<sup>28</sup>

To avoid headaches, we recommend that you meet with your project team as soon as it is clear that you will be developing a new software tool and ideally before any code has been made public. Consult tools<sup>◦</sup> for open source license selection<sup>◦</sup>, and consider:

- Who you want to use your code, and under what circumstances
- How your code might inspire derivative works, including add-ons and plugins
- What other code you might want to use in your project (*i.e.*, dependencies) and what licenses those codes use

Be sure to consider license compatibility<sup>◦</sup>. Some different open source licenses play nicely with each other, but others don't!

---

<sup>27</sup>Examples are a nice addition to docstrings but aren't necessary in all cases.

<sup>28</sup>If there's only one developer, it's not as much of a hassle in most cases, though some licenses are irrevocable, like the “Apache” 2.0 license<sup>◦</sup>.

# 11. Computing

## 11.1. Available Resources

Currently, the CoReACTER has access to the following resources:

- The Tartan Research Advanced Computing Environment (TRACE)<sup>◦</sup> cluster: TRACE is CMU Engineering's high-performance computing (HPC) cluster, housed at the Pittsburgh Supercomputing Center (PSC). We have access to some private nodes, which we can run on at high priority, and we can also run low-priority jobs when there are idle nodes. Currently, TRACE has two partitions: **i**) a "hybrid" partition with dual AMD EPYC 7713 (64-core/128-thread @ 2.0 GHz) CPUs, 2 TB RAM, and one NVIDIA Ampere A40 GPU with 48 GB on-GPU RAM; and **ii**) a CPU-only partition with dual AMD 9474F (48 Core/96-thread @ 3.6 GHz) CPUs and 768 GB RAM. A more GPU-heavy partition is in the works. We can use TRACE for any research purpose, though one of our nodes was purchased through a project with Toyota Research Institute (TRI).
- The PSC Bridges-2<sup>◦</sup> cluster. Bridges-2 has three partitions: regular memory, for CPU-only jobs with moderate memory needs (<256 GB); extreme memory, for CPU-only jobs requiring up to 4 TB of RAM; and GPU, for jobs requiring acceleration with up to 8 GPUs per node. We have access to Bridges-2 through two projects, currently: EWCSS' Carnegie Bosch Institute fellowship project on metal-mediated electrochemical ammonia synthesis and the TRI project on solid-state synthesis.
- NCSE Delta<sup>◦</sup>, an NSF-sponsored cluster (like Bridges-2) with CPU-only nodes and a variety of nodes with between 4 and 8 GPUs. We have access to Delta only through the TRI project on solid-state synthesis.

All ReACTERs working on compute-heavy tasks should get access to TRACE. Accordingly, in the examples included below, we'll be focusing on how one would do things on TRACE. If your project has additional compute resources outside of TRACE, you should talk to EWCSS or someone else on that project to learn how to gain access to and use those resources.

## 11.2. Getting Started on an HPC System

First, you need an account. Unless you are a PI on a project with an HPC allocation, you will typically need an administrator or PI to grant you access. For TRACE, this will involve EWCSS adding you to the CoReACTER group.

Once you have access, you will typically log into an HPC system through the Secure Shell (SSH)<sup>◦</sup> protocol. The most straightforward way to do this is *via* a terminal command-line, but there are also ways to access remote servers through e.g., VSCode<sup>◦</sup>. To SSH into TRACE:

```
$ ssh <USERNAME>@trace.cmu.edu
```

You'll be prompted for your password, which is the same as your AndrewID password.

If you want to use `zsh` instead of the default `bash` shell, instead execute:

```
$ ssh <USERNAME>@trace.cmu.edu -t "zsh --login"
```

You should be able to log into TRACE as soon as EWCSS adds you. However, it might take a day or two to get access to the group file system (located at `/trace/groups/coreacter`). You'll want to wait until you have file system access, because most of your files should be located in the group storage. If it takes more than two days, reach out to `trace-admin@cmu.edu`<sup>◦</sup>.

Once you've logged on and have file system access, you'll want to get your programming environment set up. We recommend that you create a personal installation of `mamba` (for Python environment sandboxing and dependency management) and `juliaup` (for managing versions of Julia).

To download and install `mamba` (through Miniforge):

```
$ cd /trace/groups/coreacter/<your_username> # move to the install directory
$ mkdir software && cd software
$ wget https://github.com/conda-forge/miniforge/releases/download/24.11.2-1/
Miniforge3-Linux-x86_64.sh # download install script
$ chmod +x Miniforge3-Linux-x86_64.sh && ./Miniforge3-Linux-x86_64.sh
```

Then, follow the prompts, being sure to install into your sub-directory of the CoReACTER file system. Once `mamba` is installed, you'll need to make sure it's active and incorporated in your shell startup script (`~/.bashrc` for `bash`, `~/.zshrc` for `zsh`).

To install `juliaup`:

```
$ curl -fsSL https://install.julialang.org | sh
```

Then, follow the installation prompts. By default, `juliaup` will only install the most recent Julia version,<sup>29</sup> and this is the version that will be launched when you run `$ julia` in the terminal. To install and use other versions, see the documentation<sup>o</sup>.

Once your programming environment is set up (making sure that both `mamba` and `juliaup` are in your `$PATH` variable), what you do will depend on what your research looks like. But if you haven't used an HPC system before, you might want to start out by playing around with the queueing system for running calculations.

To manage calculations, TRACE (and most HPC clusters) uses the SLURM<sup>o</sup> scheduler. SLURM decides what calculations go to what compute nodes, which jobs get how much memory, how many threads, and how many GPUs, and more. For a basic guide to running calculations on TRACE using SLURM, see the TRACE docs<sup>o</sup>.

Once you're comfortable with SLURM and running basic calculations, you'll probably want to try out some realistic chemical simulations. Thankfully, some simulation software is already pre-installed. You can see the available software "modules" by running:

```
$ module avail
```

Let's say you were interested in running a density functional theory calculation using QuantumEspresso. You see from `module avail` that there are multiple QE modules available, so you choose one (in this case, the module with GPU support):

```
$ module load quantumespresso/7.1-nvhpc22.7-cuda11.7
```

Now, you'll have access to QE executables and some necessary environment variables!<sup>30</sup>

### 11.3. Budgeting

As a relatively young, relatively small research group, we do not have the luxury of a large computational budget. We have (or will soon have) a few nodes on TRACE, some small NSF

---

<sup>29</sup>At the time of writing, version 1.11.

<sup>30</sup>Two notes here: first, while you can load a module from a login node, you should NEVER run any heavy computations there. All serious calculations should use the SLURM queue. Second, to use QE, you'll still need to do some setup, for instance by downloading and installing your preferred pseudopotentials.

allocations through ACCESS and NAIRR, and that's it. It is very likely that we will not always have access to all of the computing resources that we need to run that calculations that we want. We need to budget.

Computational budgeting is critically important. If not done carefully, we can prematurely run out of our allocations on clusters like Bridges-2. While we can't run out of compute on TRACE, all ReACTERs need to share the same nodes. Having an idea of how long your calculations will take allows us to share and use our resources more effectively.

Computational budgeting is mostly guesswork. While we don't expect you to predict how many CPU-hours or GPU-hours you need, we will ask that you make a back-of-the-envelope estimate supported by small tests. To effectively benchmark:

1. Make your wish list. If you had infinite compute, what would you like to do? How many systems would you simulate? How large would your systems be? How large would your ML models be, and how much hyperparameter tuning would you do?
2. Determine your total budget. If you have an allocation on e.g., NERSC or ACCESS, the upper bound will be the total allocation. Usually you won't want to use an entire allocation on a single calculation campaign, so consult with the other folks on the project(s) associated with those allocations. If you'll be running on TRACE, the total budget is harder to estimate. As a starting point, think of how many days  $d$  you're willing to wait for your results. Then, your estimated budget  $B = 24d\frac{N}{3}$  or  $B = 8dN$ , where 24 is the number of hours in a day,  $N$  is the total number of nodes available,<sup>31</sup> and  $B$  is given in node-hours. To convert into core-hours, simply multiply by the number of cores available per node.
3. Identify realistic problems of different sizes. This usually isn't terribly hard. For density functional theory, the cost scales roughly as  $n^3$ , where  $n$  is the number of basis functions. So, pick some molecules or structures with different numbers of elements and run a calculation similar to what you'd like to run (e.g., a self-consistent field calculation or a geometry optimization). For molecular dynamics, cost scales roughly linearly with the number of atoms, so likewise, try running near-identical simulations with more and more atoms. If you're training an ML model or testing out an algorithm, you might need to be slightly more creative to create realistic problems of different scales, but generally, problem sizes scale with the number of datapoints, the number of model parameters (for ML), the complexity of the algorithm, and (if you're running calculations in parallel, especially on multiple nodes) the communication between processes.
4. Run your standard(-ish) workflow on your different problem sizes and track the cost. If you're not sure the optimal way to run these jobs (e.g., in terms of number of processes, amount of memory, etc.), you might also run the same problem with different job parameters. What jobs finish quickly? What jobs take close to the wall-time limit?<sup>32</sup> Which jobs don't finish, or can't be run at all?
5. Go back to your wish list and calculate roughly how much it would cost, based on your trial runs. Compare this estimate to your total budget. If the wish list estimate is less than the total budget, you're golden. Go for it and run everything you want. If it's higher (as is usually the case), you need to start trimming calculations.
6. Remember that in step 2, you're figuring out your absolute maximum budget; you don't need to use all of it!

---

<sup>31</sup>here, we're assuming that you're using the low-priority queue and that, at any given point, you have at most  $\frac{1}{3}$  of the TRACE nodes at your disposal.

<sup>32</sup>For TRACE, currently 48 hours

## 11.4. High-Performance Computing Tips

1. *Always run test jobs before production jobs.* This is necessary for budgeting (see steps 3 and 4 in Section 11.3 above). Test jobs also make sure that your workflow behaves as expected. One of the worst feelings in the world looking back on your production runs and realizing that most or all of them failed in some way because of a simple mistake. Try not to let it happen to you!
2. *Stay organized.* Give your data names that mean something to you. Use metadata consistently, and use timestamps to keep track of what happened when. Somewhere (ideally in the organization AnyType), keep a log of what calculations you've run, how you set them up, and what the status is.<sup>33</sup> Much of this is expanded upon in Section 12 below.
3. *Keep track of errors and how you fix them.* Once you've been running on HPC systems for a while, you'll notice the same issues popping up again and again. Each system has its own quirks, and there are certain common pitfalls that come with using SLURM as well as certain codes. By keeping track of your errors, you'll save time not only for yourself, but also potentially for other ReACTERs in the future!
4. *Avoid premature optimization, but do optimize.* As we note above (Section 11.3), our computing resources are precious, and we want to be efficient. However, your first job is always to get something working. Once that's done, and *only* once that's done, should efficiency be among your worries.
5. *Make judicious use of checkpointing, batch launching, and low-priority queues.* These are all valuable ways to improve efficiency! By checkpointing, you ensure that, if your job fails part-way through, you can continue without restarting from the beginning. Batch launching means running multiple jobs within the same job; if your workflow consists of many small parts, it's a good way to make sure that you're always using all of your allocated computing resources. Finally, many HPC systems (but not all) have a low-priority queue. Using such a queue will allow you to run your jobs faster (and often at low or no cost), but your job can be pre-empted and terminated early. Coupling low-priority queue usage with checkpointing is an ideal method for getting more bang for your computing buck.
6. *Coordinate compute usage.* Communicate, communicate, communicate! Most of our research is computational, so most of us will have some need of our HPC resources most of the time. If you work with your fellow ReACTERs one-on-one or in a project team or group organizational meeting, we might be able to get everyone the resources that they need. If we don't coordinate, then it's likely that at least some ReACTERs won't get what they need and therefore won't be able to make necessary progress.

---

<sup>33</sup>Are they ready to run? Running? Did they fail? If so, do you know what you need to do to fix them?

## 12. Managing and Sharing Data

Data is everything. It's a resource that we share with the community and that we leverage for data science. Retaining data helps us to understand what we did and potentially allows others to reproduce our analyses. Given the prime importance of data in our research, we want to take care to generate, maintain, and distribute data intentionally and carefully. This chapter outlines how we plan to do that.

### 12.1. "Soft" Data and Institutional Memory

Most of this chapter will cover "hard" data, *i.e.*, the stuff that you get at the end of a simulation or experiment. While that is certainly important, it doesn't cover all of the data that we generate. Actually, one could argue that "soft" data, *i.e.*, data that is not directly related to a scientific observation but is necessary for research activity or organizational function, is even more vital! A simulation can always be re-run; an idea, once lost, might be gone forever.

You should maintain soft data for your future self. You will forget things that you want to remember, but taking careful notes and managing "soft" data effectively will make this less common and less disastrous when it does happen. Even if "soft" data management didn't help you at all, we would still ask you to do it for the benefit of the CoReACTER as a whole. The reality of the Academy is that turnover is high. Master's students are typically gone in one to two years, while undergrads and PhD students typically work in a lab for at most four or five years. Postdocs might stay for anywhere from six months to 3 years. When folks leave, they take with them institutional knowledge and memory. Frequently, they will also leave unfinished projects behind that their colleagues might want to take to the finish line. Keeping good records is a way to make things easier for those who come after you.

Our organization AnyType Space is the main place where you should store ideas, meeting notes, important records, decisions, tutorials, information related to computational simulations, and more. When you join the CoReACTER and gain access to the AnyType, you should create a ReACTER subspace<sup>34</sup> where you'll put all information that is personally relevant to you. You should feel free to organize this document however suits you best. For information related to a project that affects more than just you, you should store that information or link it in the appropriate Project document.

In addition to the AnyType, we use an ELN package to manage our laboratory inventory and experimental data (see Section 10.1.5). While information about simulation runs can be put in the AnyType, information about your experiments, laboratory equipment, and experimental protocols should be stored in your ELN.

### 12.2. Metadata

Data is rarely useful without context. To make sense of a datum, we want to know how it was collected, and when; what other data it's related to (*e.g.*, is this one point of a time series?); for what project or purpose the datum was collected; what simulation or experimental parameters were used during data collection; and potentially more!

We provide this information by including *metadata*, *i.e.*, data about our data. The inclusion of high-quality metadata is important for our own purposes, improving simulation and experimental reproducibility and preventing organizational memory loss. Moreover, metadata are a

---

<sup>34</sup>TODO: set this type and template up and create an AnyType tutorial



major factor in the FAIR data guidelines (Section 9.2). If we want to share our data and have it be useful for our academic communities, we need metadata!

While the exact metadata that you include with a dataset will vary based on the nature of the data, as a starting point consider including:

- A unique (and, ideally, meaningful) identifier
- Timestamps for when the data was generated, analyzed, modified, etc.
- Where applicable, any software packages used to generate the data, with version information
- Where applicable, information about the hardware or equipment used to generate the data
- Any non-default or non-standard parameters used to generate the data
- Any relevant labels or warnings

### 12.3. Data Organization

With a good set of metadata, you could conceivably name your files whatever you like. After all, who cares? Just look in the metadata and see what this molecule is named, or when this experiment was run.

For automated data retrieval, this approach is reasonable, but sometimes we as humans must go looking for files and directories. Therefore, how we organize our data matters. Our general advice, which we hope all ReACTERs follow:

- Other than root directories (e.g., “~/documents” or “~/data/postdoc”), all directories should be dated, using the YYYYMMDD format.
- If your data is labeled or tagged, it’s often helpful to include those labels in directory or file names.
- In general, *make file names mean something*, if you at all can. Whether it’s an ID or some basic parameters, having a human-readable filename is much nicer than the alternative.
- Nest your directories to limit clutter.

### 12.4. Online Repositories

There are nowadays many online data repositories. Figshare<sup>o</sup> is a particularly easy-to-use repository that we have used many times. For relatively small collections of data and data dumps, it’s perfectly appropriate.

We make persistent archives of our software on Zenodo<sup>o</sup>. In addition to providing snapshot archives that can be synced with GitHub, Zenodo provides digital object identifiers (DOIs) for your software, making it easier to cite. It’s pretty cool!

In the past, EWCSS has worked closely with the Materials Project<sup>o</sup>, which hosts a large core collection of material and molecular properties as well as user-contributed collections through MPContribs<sup>o</sup>. If your data uses the Materials Project or could interface easily with their collections, consider reaching out to their staff!

### 12.5. Cloud Storage

Through CMU, we have unlimited cloud storage through Google Drive. This isn’t the most secure, but for non-private data,<sup>35</sup> that’s fine. Use Google Drive as a backup solution!

Small files can also be stored in our AnyType. We currently pay for 256GB of cloud storage. That’s not enough for a bunch of experiments or simulations, but it’s plenty for small images, text files, and the like.

---

<sup>35</sup>which includes most of our data

## 13. Writing, Presenting, and Communicating Science

### 13.1. Authorship

Authorship is a currency in the Academy. To get jobs out of school, to gain promotion and tenure, one must be an author on many papers and on “high-impact” papers. More than that, in the chemical sciences, candidates for awards and jobs are often judged based on where they sit in an author list. In our field, first authors are typically viewed as those who performed the majority of the work and did the majority of the writing, while the last and/or corresponding authors are viewed as the intellectual leaders of the project.

Because of the value that authorship holds in academic research, the author’s list is a space where power dynamics play out. “Ghost authors”, often junior researchers, women, and folks holding other marginalized identities, perform research but are not given credit or acknowledgment for their labor. On the other end, “honorary authors” (sometimes called “guest” authors or “gift” authors) are (typically powerful, senior, and/or privileged) individuals that are added to an author list without performing any significant labor. [4], [5]

We have been victims of authorship inequity. We have witnessed inequities occurring, and we have heard of inequities in papers that we have not been a part of. To maximize fairness, we have devised a set of standard procedures for determining authorship. We will attempt to follow these guidelines for all papers that any ReACTER is a part of, and that the CoReACTER is participating in as an organization. However, we recognize that we will not hold significant power in all scientific teams that we are a part of and on all papers that we contribute to. Because many of the practices that we propose are foreign to chemical engineering and related disciplines, we expect push-back from collaborators, and we reserve the option to negotiate new procedures as needed.

These guidelines are inspired by the work of Liboiron et al. [6].

First, we note that **authorship is an ongoing discussion**. We begin conversations about authorship early — as soon as possible after we realize that a project is likely to become a paper. These early conversations set expectations regarding who will be involved in a project and, more specifically, who will play what role(s). In determining a preliminary author’s list, we also establish a preliminary idea of value — what aspects of the project are most important, and how will we recognize different forms of labor for a given project.

As a project evolves, these initial decisions may no longer be appropriate. The scope of the project may change, requiring us to reassess which aspects are most important. Someone may have done much more labor than they initially intended, or someone may have had to step away from a particular project because they did not have as much energy to devote to it as they expected. When anyone working on the project — including someone not initially listed as an author! — notices a significant divergence between the initial decisions and the reality of the project, they should initiate a discussion on authorship and labor. Even if no major disagreements have been identified, authorship must be revisited by all individuals involved in a project before any document (e.g., abstract, manuscript, or preprint) is finalized and submitted. In these conversations, we always strive for consensus (see Section 7.3), and we aim to reach decisions on authorship order that accurately reflect the work done and that are amenable to all involved.

How should we decide who should be an author, and who should merely be acknowledged? **What constitutes authorship?** Many ethical guidelines (including the Vancouver Protocol<sup>o</sup>, as

well as guidelines from societies like IEEE<sup>o</sup> and the Royal Society of Chemistry<sup>o</sup>) agree broadly on a few points. To be an author, one must:

1. Make a significant intellectual contribution to the work
2. Take ethical responsibility for the complete work
3. Participate in the writing and/or review of a final work (e.g., manuscript or book chapter)

But this raises another question: what constitutes a “significant intellectual contribution”? Ultimately, we recognize that there will never be a perfect set of criteria to unambiguously determine significance, but we ground our subjective determinations:

- **Capital is not an intellectual contribution.** In contrast with the Contributor Roles Taxonomy (CRedit)<sup>o</sup>, we do not believe that one should be considered an author if their main or sole contribution is the provision of funding or resources (laboratory equipment, compute hours, etc.). The only exception would be if acquiring the resources required a significant intellectual effort; for instance, if a new instrument had to be designed or created for a particular project.
- **Generic mentorship does not count towards authorship.** Mentorship is a vital aspect of academic training, and in CoReACTER, we value mentorship of all kinds, between members of all levels of experience and all positions (Section 7.4). Because mentorship is a process of intellectual guidance, mentoring a researcher through a project can constitute a contribution and earn someone authorship. However, the mentorship in question should be *specific and relevant* to the particular project. In a commonly encountered example: a PI should not be listed as an author on a paper for a project they did not contribute to only because they trained a lead author.
- Care work and organizational labor are intellectual contributions and can be significant. We have already mentioned mentorship. In addition, care work and organizational labor can include sharing information between team members; planning and running meetings; taking notes and keeping records; and managing equipment.
- To determine if a contribution was “significant”, ask: **could this project, as it exists, have been completed without that contribution?** If a piece of equipment broke during a project but someone fixed it to allow data collection to continue, that repair work is significant, and the person performing the repair should be considered for authorship. On the other hand, if someone performed regular equipment maintenance that was not directly related to the project, that is probably not significant; the project may well have been completed without that maintenance occurring. Note that it does not matter, to us, if a particular labor could have been done by someone else. All that matters is that it was done by a particular person/group of people and that it was important.

We hope that this goes without saying, but critically, authorship is determined by each author’s contributions. This means that a graduate student can be corresponding author as easily as a professor, if they are the major intellectual leader on a given project. Rank and status are not irrelevant in authorship decisions (see below), but where we consider them, we tend to favor those with less power.

In general, we try to take a big-tent approach. If someone has met all of our criteria for authorship, we add them gladly. If we’re in doubt if someone’s contributions merit authorship, we try to err on the side of generosity. We believe that sharing credit, including privileged positions (e.g., by having co-first or co-corresponding authors) hurts no one, so in cases where there are good-faith disagreements regarding who led a particular work, we try to offer to share credit.

Finally, working towards equity in authorship, we consider power and social context. Each team member holds intersectional identities (related to gender, race, ethnicity, ability, age, title or rank, class, nationality, etc.) and may be privileged or marginalized in scientific contexts as a

result. Papers offer social capital, but that social capital may have variable value because of disparities in privilege. For instance, a given paper holds less social value for a senior professor who has written 200 papers than it would for a junior researcher with 30 papers under their belt who is up for tenure. Notably, “considering” power imbalances does not mean that we will always elevate junior researchers or researchers from marginalized backgrounds, but rather that we will in all cases consider using authorship to increase the social capital of authors.

## 13.2. How We Write

### 13.2.1. Software

Our software of choice for writing papers is LaTeX, a typesetting language first released in 1984. LaTeX gives users (writers) a high degree of freedom, and it is a *de facto* standard in many areas of scientific publishing. ReACTERs who are not familiar with LaTeX should consult online references:

- [Overleaf](#)<sup>o</sup>
- [Learn LaTeX](#)<sup>o</sup>
- [LaTeX Wikibooks](#)<sup>o</sup>

Remember that your fellow ReACTERs are always a great resource!

In addition to being widely used, LaTeX is a typesetting language of choice because it can be used collaboratively. We use Overleaf<sup>o</sup> to develop LaTeX-based documents in teams. If you make an account using a CMU e-mail, you’ll have free access to Overleaf Pro features, which is important to access features such as unlimited collaborators, Zotero integration, and full document history.

In some cases, journals and other publishing venues will specify a non-LaTeX writing format. For instance, the *Journal of Open Source Software* requires that manuscripts be written using Markdown. In these cases, you should follow the journal’s guidelines. For alternative formats, you should still write publicly such that your collaborators can easily edit and add comments. For instance, if using Markdown, you can add the manuscript text to a GitHub repository, and Google Docs can be used for Word-like editing.

If, by some stroke of luck, a journal or other publishing venue has no formatting requirements or will accept a compiled document (e.g., a PDF), we encourage ReACTERs to consider Typst<sup>o</sup>. Typst is a relatively new typesetting language (publicly released in 2023) that offers some considerable benefits over LaTeX. Perhaps most critically, Typst offers incremental compiling, so changes can be reflected in your document near-immediately. In contrast, LaTeX documents generally need to be recompiled following each change. Depending on the size of the document, compiling could take several minutes, severely slowing down the process of writing and revising. Typst can be edited and compiled locally, and there is an Overleaf-like Web app for collaborative editing. Typst is still in development, but it’s already usable for full-scale projects. In fact, this manual is written in Typst!

### 13.2.2. Process

Writing is fundamentally a creative and generative exercise. Because of this, it is also a highly individual process, and each writer may have their own method that works best for them. We encourage ReACTERs to experiment with different approaches and find what works! In the end, the best paper is a finished one, not one that follows a particular formula.

With that said, here we provide an example process. For folks who are not comfortable with scientific and technical writing, or for folks writing an academic manuscript for the first time, we encourage you to at least try following this method start-to-finish once. If you hate it, great! You've learned something that doesn't work for you. But you might find that's it's a helpful way for you to process your thoughts and get your writing done with minimal headaches.

### **Suggested process:**

1. **Answer the big questions.** Before you put any words to paper, you need to figure out what, exactly, you're writing. You can try to answer these questions yourself, but it may be helpful to talk some of these through with your mentors or collaborators. "Big questions" include:
  - What question does this paper answer?
  - What will this paper add to the literature?
  - What, if anything, is this paper responding to?
  - What is the central claim or thesis of the paper? In other words, if someone were to read this paper, what core idea do you want them to take away?
  - What kind of paper is this? Is it a forward-looking perspective or a review of the literature? Is it a thought piece, putting forward a novel idea or theory? Does it describe a new method, a new piece of software, or a new dataset? Is it a proof of concept? Are you doing a deep-dive into one material or chemical system, or are you exploring chemical space?
  - Who is the intended audience? Is this paper aiming to inspire a general audience, or is this meant to support research in your sub-field?
  - How much do you have to say about your topic? Is there one big result, or many results that add up to one big conclusion?
2. **Pick a target.** At this point, you know the story you're trying to writing, why it's important, and who you want to read it. This gives you enough information to pick out a preprint server, journal, or other publication venue.

What makes a journal the "right fit" for a paper? This seemingly simple question is a challenging one. You can think of journal selection as a multi-objective optimization problem, where you need to consider an array of different factors and criteria, including but not limited to:

- a) *Readership:* Where will your audience be most likely to find and read this paper?
- b) *Reputation:* What journals do you respect, that are known for quality work in a field related to your paper?
- c) *Speed:* What journals will review your paper reasonably quickly?
- d) *Likelihood of acceptance:* What journals publish papers with similar styles and content as yours?
- e) *Policies:* What journals are friendly to authors – for instance, allowing authors to deposit pre-prints, charging low or no fees for open access publishing, etc.?

Ultimately, you're unlikely to find one perfect journal that meets all of your criteria, but you can hope to find a reasonable fit for your target. To help you come up with a list, consider our guide (see Section 13.3) when selecting journals. In particular, this may help you brainstorm journals, and it may also help you avoid predatory publishers. If you're new to scientific publishing, we also recommend that you think about what journals you like to read from, and that you consult your colleagues.

We note that peer review can be challenging, and your paper might be rejected for reasons legitimate and illegitimate alike. In the spirit of always being prepared for the worst, it's a good idea to select one or two "backup" journals for if your first-choice journal doesn't look favorably upon your manuscript.

- 3. Outline.** Next, you need to figure out how you're going to tell your story. The two most important things that you want to get out of an outline are logic and structure. Logic ensures that you are effectively guiding your reader through all of the technical details. You want to ensure that each piece of your paper logically connects to what came before and what comes after, so that your reader is never confused and rarely if ever has to jump back and forth between different sections. Structure allows you to break down your text into digestible pieces, again supporting the reader.

To outline, begin by breaking down your paper into large sections. These sections will vary depending on the journal you've chosen but may include an Introduction, Methods section(s), Results, Discussion, and Conclusion. Consider what information you need to share with your reader to convince them of your central thesis (see point 1 above) and what information fits best in what section. Next, further sub-divide each section. Do you need multiple sub-sections? If so, how should the sub-sections be ordered such that the reader has the necessary information to understand each sub-section? Continue this process of breaking the paper down into smaller and smaller pieces until you've reached the level of describing the content of individual paragraphs.

After you've determined how the text of your paper will be organized, consider the figures. What are the most important ideas of your paper? What are the ideas that may confuse readers the most? And finally, what data do you need to convince the reader that your ideas are correct? Each of these points — the critical idea, the confusing concept, and the data — should be represented by figures. Once you've outlined what your figures need to do, you need to outline how they'll do it. For each figure, try drawing a crude sketch on a piece of paper (or on a drawing app, if you prefer). Then try showing it to a colleague and see if they understand, without a verbal explanation, what you're trying to convey. After they've given their explanation, have a discussion with your colleague where you try to understand what may make the idea in your figure more clear. Keep in mind that many readers will look at figures before they read the text of your paper — to the extent possible, the figures need to be clear on their own.

Finally, go through your outline and (attempt) to identify information that a) is not necessary at all to tell the story that you want to tell or b) is important to your story but may be too technical for an average reader. Information of the first kind should be removed from your outline, while information of the second kind — figures and text alike — should be moved to the paper's Supporting Information.

- 4. Write.** The time has finally come. You can (metaphorically) put pen to paper and give birth to your masterpiece!

Writing a manuscript can be daunting, especially for folks with limited experience in academic science communication. To avoid being overwhelmed, it's good to break down the task into smaller pieces of work. You might choose to begin with the Methods section, which generally requires little argumentation, just clear reporting of facts and parameters. Alternatively, you might want to begin with figures and then write the text around explaining those figures. Regardless of what you start with, it's generally advised that you write your abstract and conclusions last. These sections look at the manuscript as a whole, and it'll be hard for you to summarize your work before you've actually written it.

Just as it can help to break down the goal of writing a paper, it can help to break down the writing time. Try pacing yourself, committing to a couple of hours of writing and editing every day (taking time to rest, of course!). Even if you hate writing, hopefully you'll find the short time commitment manageable, and the paper will get done in no time!

### 13.2.3. Style

On style, we strive to strike a balance between the artistic and the practical, the individual and the standard. As a creative practice, scientific writing should never be reduced to a formula, and individual writers should be able to express their ideas in an individual voice. At the same time, the main goal of scientific writing is to share information, which requires clarity of ideas. Here, we outline some general writing principles that we hope will lead to clear, informative papers without significantly reducing ReACTERS' freedom.

1. **Make your writing accessible.** This could also be written as “**be kind to your readers**”. In the end, your manuscripts will be read by human beings. For your work to be as widely appreciated as possible, you should work to make your writing accessible. For figures, this means using color schemes and patterns that can be distinguished by folks with colorblindness, and you should provide descriptive figure captions such that viewing the figure is not necessary to understand the work. For text, avoid unnecessary jargon as well as culture-specific idioms and references (at least, without explanation). Always introduce a concept using its full name before providing an abbreviation, and if possible given space limitations, provide a reasonably thorough Introduction section so that non-expert readers are able to engage with the manuscript. When it comes to math, try not to skip steps, and always explain each new term in an equation!
2. **Avoid hype.** Too much of modern science is sales — trying to sell that your method is better than everything that has come before, selling a new model or a new material. In part, this is done so inflate the importance of a work, so that it may be accepted in shiny, “high-impact” journals like *Nature* (and all of the sub-journals, e.g., *Nature Chemistry*, *Nature Chemical Engineering*, *Nature Materials*, *Nature Energy*, *Nature Computational Science*, and *Nature Communications*), *Science*, *Energy and Environmental Science*, etc. As humility is one of our principles, we push back against hype. We should provide accurate, balanced accounts of our findings, emphasizing shortcomings as well as triumphs. We should recognize the successes of others and realize that most science is incremental, that true breakthroughs are few and far between. Practically speaking, this means avoiding superlatives (“the biggest”, “the best”, “the fastest”) and over-emphasizing novelty.
3. **Write with conviction.** Just as you should avoid exaggerating how excellent your ideas are, you should not couch your ideas in uncertainty. If you believe that a previously held idea is erroneous, and that you have adequate evidence to demonstrate that, say it. Rather than “The idea that [XXX] may not be accurate in this case”, say “Though previous reports suggested [XXX], our findings indicate that [YYY].” On a related note, while you should always be mindful of how your work may be critiqued, you should not give concessions to the reviewers in your head, but only the reviewers that you actually have to face. Do not soften your language because you think you’ll be challenged. Wait for the challenge, and then be willing to either amend your work in response to it or to forcefully rebut it. Remember: in any paper that you write, **you** are the expert!
4. **(Usually) avoid the passive voice.** In English, there are two “voices”. In the “active voice”, the subject of a sentence performs the action. For example, in the preceding sentence, the subject is “subject”, the verb is “performs”, and it is the subject that is doing the performing. On the other hand, in the “passive voice”, the action is received. Again looking at the previous sentence, “action” is the subject, but it is not doing anything. Rather, something else receives the action. Passive voice often feels awkward, at least for native English speakers. Sentences in the passive voice can often be written more directly using the active voice. This guideline uses the term “usually” because of one major exception: methods. Typically, we write our



methods sections in the passive voice, as we are describing what was done to obtain our results, rather than what we do.

5. **Maintain focus.** One paper tells one story. That story should be consistent, and it should be clear how each section of the paper serves that story and promotes your central thesis. Try to avoid side tangents and arguments that do not closely connect to your main take-away ideas. For short tangents that you believe to be important for your readers: put them in the main text only if you do not think they will be distractions. Otherwise, or if the tangent is longer than one or two sentences, put it in the Supporting Information. If you find yourself wanting to make a large number of asides, consider if you should be trying to write two separate papers, rather than one.

#### 13.2.4. Policy on Generative Machine Learning

Large-scale generative machine learning (ML) models, including large language models (LLMs, e.g., ChatGPT) and diffusion-based image generators (e.g., DALL-E), suffer from a host of ethical issues, including theft of intellectual property [7], [8] and devaluation of creative labor [9]; mis- and disinformation [10], [11], [12]; exploitation of low-wage workers [13], [14], [15]; exorbitant, unsustainable energy and water usage [16], [17], [18], [19]; promotion and reinforcement of systems of oppression and social bias [20], [21], e.g., racism [22], [23], sexism [24], ableism [25], homophobia, and transphobia [26]; and more. Though widely used for information retrieval, including on search engines like Google, generative ML operates on a purely statistical and stochastic basis, with the information generated having no necessary grounding in truth. In philosophical terms, LLMs are “bullshit” [27].

Considering the immense harm perpetrated by generative ML models, their widespread use is morally unjustifiable. Moreover, use of generative AI takes away from our potential to learn and grow as writers and graphic designers, to be creative and to generate art with our work. For these reasons, ***the CoReACTER does not condone the use of generative ML in our writing and scientific communication for ANY purpose and in ANY form.***

### 13.3. Where We Publish

#### 13.3.1. Preprint Servers

We are committed to making our research findings publicly available and open to the community as soon as reasonably possible. To that end, we are committed to posting all manuscripts in some form on preprint servers.

Preprints are different from journals (discussed below), in that they do not act as gate-keepers. In most traditional preprint servers, such as arXiv<sup>o</sup> and ChemRxiv<sup>o</sup>, manuscripts are near-immediately available for free after submission, because posted manuscripts do not undergo peer review. In such cases, readers must assess the value of a paper without relying on the assurances of quality provided by the traditional acceptance process of a journal.

There are some alternatives that act as compromises between the relative anarchy of arXiv and the gate-keeping of journals. Some preprint servers, such as ChemRxiv, allow readers to provide comments on posted scientific works. These comments could be used by readers to critically assess a paper, similar to the benefit of reviewer reports provided in transparent peer review. There are also efforts such as the “Peer Community In”<sup>o</sup> (PCI) preprint servers that provide “recommendations”. Preprints can still be posted without peer review, but following submission, “recommenders” can assess a work. In this system, a recommended work is considered to have the same degree of expert approval of a peer-reviewed manuscript in a journal.



We do not have strong preferences on preprint servers, as long as they are completely free and open-access and are community-accepted. Because our work is often designed to be read by chemists, chemical engineers, and materials scientists, ChemRxiv is usually an appropriate venue. However, we are open to other options! For instance, our work on chemical reaction networks may find a good home in the PCI in Network Science<sup>o</sup>.

### 13.3.2. Journals

There are easily hundreds of journals related to chemistry, chemical engineering, materials science, and computational science that we could publish in. We, as scientists and authors, are the reason that the journals have readers, clout, and prestige. We have the luxury of choice. We should use our power by submitting to journals that best fit our papers and our values.

And we must consider this choice carefully, as this is a political choice with significant potential ramifications. Many journals, particularly those that publish open access articles (see Section 9), charge the authors of an accepted manuscript a fee. By choosing a journal, therefore, we do not only support it indirectly by encouraging readership and by providing a valuable “product” (our ideas, words, and data); we also give the journal our direct financial support. Publishing an article is, to some extent, an endorsement of the editorial policies of a journal; we can therefore use our submissions (or lack thereof) as solidarity, as protest, as boycott.

So, what do our principles tell us about which journals to support? How to we make this ethical, political, and strategic decision?

Because we are fundamentally anti-Capitalist, we prefer journals that are run on a not-for-profit basis. Ideally, these would be community-run (such as the *Journal of Open Source Software*), but journals operated by research societies (e.g., the American Chemical Society (ACS), Royal Society of Chemistry (RSC), and the Institute of Physics (IOP)) are also aligned with this goal.

Because we believe in open science, we prioritize journals that allow for open access publishing. In particular, we’re big fans of “diamond” open access journals (e.g., *ACS Central Science*, *Chemical Science*), which charge neither readers nor authors. “Gold” open access journals charge the authors a one-time article publishing charge. While we are happy to publish in gold open access journals, we must be careful to avoid exploitative and excessive fees, such as those charged by *Nature*<sup>36</sup> and its related journals.

Finally, we specifically avoid predatory publishers that charge fees but do not engage in substantive peer review and editing. We currently avoid journals owned by Multidisciplinary Digital Publishing Institute (MDPI) and Frontiers; though individual journals may be more or less predatory, the publishers’ policies and strategies are harmful enough that we keep our distance in general.

Some specific journals that we’ve had positive experiences with include:

- *The Journal of Open Source Software* (Open Journals)
- *Digital Discovery* (RSC)
- *Chemical Science* (RSC)
- *ACS Energy Letters* (ACS)
- *The Journal of Physical Chemistry Letters* (ACS)
- *Journal of the American Chemical Society* (ACS)

---

<sup>36</sup>\$12,290.00 as of May 2024

## 13.4. Meetings and Conferences

Scientific meetings and conferences are complex, with significant benefits and drawbacks on individual and structural levels. Conference presentations are undoubtedly an important form of scientific communication and offer unique opportunities to engage in critical dialogue with experts in your field and adjacent fields, many of whom you may not regularly communicate with. Many argue that meetings are also important for developing and maintaining scientific networks [28], [29], though the power of conferences for networking depends on one's positionality, degree of extroversion, and social comfort. On the other hand, meetings are a financial burden on academics, particularly those from under-resourced regions and institutions [30]. They are in many ways environmentally unsustainable [31], [32], from the long-distance travel required to attend to the “swag” often given out by the conference organizers and partners. And just as conferences can help to create community and develop networks, they can also contribute to social isolation, alienation, and othering [33], [34], [35].

For now, the CoReACTER supports ReACTERS attending meetings and conferences, and we have a set of loose guidelines to make conferences as positive as possible for the organization and its members. To provide social and emotional support (and because conferences are more fun with friends), we generally try to send at least two ReACTERS to a given event. Because of the financial burden of meeting attendance and the environmental impact of travel, we advise ReACTERS to go to no more than four meetings in a year and to prioritize local meetings (especially in the American Northeast and Eastern Canada) over meetings in distant areas of North America or on different continents (for more information on travel, see Section 13.4.5). In general, we advise that ReACTERS only go to conferences if they have work to present, but exceptions will be made for students or postdocs in early stages of their positions as well as workshops and summer schools that are primarily educational in nature.

In the rest of this section, we cover when and where to present one's work and how to prepare a conference presentation. Much of the material covered here is also relevant to writing papers (discussed above), so we get to keep things brief.

### 13.4.1. When to Attend

If you are an early-stage graduate student or postdoc,<sup>37</sup> you should feel free to go to at least one conference whenever, just because you want to and think that you'll gain something valuable.

Once you're deeper in your research program and you've gotten at least one conference under your belt, you can determine if you should attend a particular conference by answering the following questions:

1. Is this conference essential (or very important) for my educational or professional development? You might answer “yes” to this if:
  - You're on the job market and the conference offers opportunities for job-seekers
  - The conference is a workshop or summer school on a technique or application that is vital to your research
2. Is my research at a good stage for a presentation? Generally, to answer “yes” to this question, you should:
  - Have published the work that you want to present; or
  - Be in the process of writing or editing that work; or
  - Be sufficiently far along in the project that you reasonably *could* write a manuscript describing it, even if you want to work on the project further.

---

<sup>37</sup>Whatever that means to you! We are not trying to put hard limits in terms of time, number of papers, etc.

In addition, you should have the agreement of all authors on the work that the work is ready to be presented and that you are the right person to present it (generally meaning that you were an intellectual leader on the project, *i.e.*, a (co-)first or (co-)corresponding author).

3. Are the other attendees of this meeting the right audience for my work? This question may be answered “yes” if:
  - You are aware of other attendees (including organizers) who:
    - you often cite
    - often cite your work
    - you collaborate with
    - you plan to/hope to collaborate with in the near future
  - The meeting or symposia descriptions specifically mention your area of research or the topic that you want to present on

When you find a conference that you think you should attend, bring it up at the next group organizational meeting. In general, if finances allow, we'll try to make reasonable meeting attendance possible.

### 13.4.2. Where to Present

We encourage ReACTERs to attend a mixture of large and small conferences, as they provide different benefits. Large conferences will allow you to meet more people, hear about more diverse research topics, and potentially (though not necessarily) reach a larger audience. On the other hand, smaller conferences are often better for learning more about specific, more narrow topics and developing deeper relationships.

Some large conferences that we've enjoyed include:

- The ACS National Meetings (Spring and Fall)
- The American Institute of Chemical Engineers (AIChE) Annual Meeting<sup>38</sup>
- The Electrochemical Society (ECS) Meetings

And for smaller and single-topic conferences, we recommend:

- The Gordon Research Conferences (GRCs) and Gordon Research Seminars (GRSs). These are held on a variety of topics, including:
  - Electrochemistry
  - Batteries
  - Computational Chemistry
  - Computational Materials Science and Engineering
  - Polymers
  - Catalysis
  - ...and more!
- The International Meeting on Lithium Batteries (IMLB)
- The American Conference on Theoretical Chemistry (ACTC)

### 13.4.3. Preparing a Presentation

**Software:** The CoReACTER does not have a strong preference for presentation software. You are free to use Keynote<sup>39</sup>, Microsoft PowerPoint<sup>40</sup>, LibreOffice Impress<sup>o</sup>, Google Slides, etc. to make slide decks. For the more programmatically inclined, you might want to look into using

---

<sup>38</sup>For job seekers: AIChE has the annual postdoc and faculty candidate poster sessions, which can be helpful for securing a position when you're ready to leave the group.

<sup>39</sup>If you use Mac OS X.

<sup>40</sup>We have free access through CMU.

LaTeX or Typst; both typesetting languages offer templates and utilities for presentation design (see Beamer<sup>o</sup> for LaTeX and typslides<sup>o</sup> or slydst<sup>o</sup> for Typst). We typically recommend making a poster in a vector graphics software like inkscape<sup>o</sup>, but tools like PowerPoint also work!

Regardless of what software you use, you should export a PDF version. You should also ensure that all images used in your slides are available separate from your slide deck file, so that even if someone in the group cannot access your presentation, they can recreate specific slides or visuals.

**Structure:** Most academic presentations fall into one of four broad categories. Each of these has their own rules (documented in some minor detail below), but typically, a presentation should be structured as:

1. *Introduction:* Introduce the general topic, but most importantly, *motivate* your work. Why did you embark on this project?
2. *Methods/tools:* Unless your presentation is theory- or method-focused, this probably won't be a major component of your presentation, but in order for your audience to understand *what* you did, they need to know *how* you did it
3. *Results:* Highlight what you learned *and why that matters* for your audience. Make sure you link back to your motivation — did you reach your goal? Move towards it?
4. *Conclusions/future directions:* If your audience was asleep for most of your presentation, what should they take away from your talk? And what are you aiming for next (or, what should the field aim for next)?

Unless you did all of the work for your presentation alone (which is, nowadays, very uncommon), you should acknowledge your collaborators, mentors, funding, etc. in some way in your presentation. Typically, these acknowledgments are put at the end of a presentation, but we prefer to put them at the beginning. Your collaborators and colleagues should not be an afterthought!

Specific structure tips for different categories:

1. **Posters:**

- A mainstay of conferences and symposia
- You'll have a static visual display — this should have the same major features outlined above, but severely abbreviated
- Two approaches to a poster: i) a stand-alone document that tells the whole story if you step away; or ii) a set of figures and other materials that you can reference in a discussion. Both are valid, but since you should be at your poster most of the time, the second is usually better!
- You should only prepare a short “elevator pitch” that is ~1-2 minutes in length. For the most part, you'll be discussing, not lecturing!

2. **Brief oral presentations:**

- The standard conference talk; typically no more than 15 minutes
- Involves a pre-planned monologue followed by question-and-answer (Q&A)
- With so little time, you need to keep your introduction short (no more than 5 minutes!)
- Similarly, methods should be kept brief unless your presentation is methods-focused
- A few well-explained pieces of data are better than many rushed, unexplained data

3. **Mid-length oral presentations:**

- Typically reserved for invited conference talks
- Between 15 and 30 minutes
- You should be able to tell one complete story (*i.e.*, tell the story of one full-length research paper)

- You can take more time to walk through the background and methods, though most of your talk should still be results and analysis

#### 4. Long oral presentations:

- Very rare for students and postdocs, but more common for faculty/staff researchers
- Include seminar talks, “job talks”, keynote presentations, etc.
- More than 30 minutes, but rarely longer than 1 hour
- Typically will tell the story of more than one paper. In this case, you might break your results down into multiple sections, but you will still typically want one section on background and motivation (motivating the broader area/narrative) and one conclusion

#### Timeline:

1. *Abstract submission* typically takes place somewhere between 6 and 9 months before major conferences.<sup>41</sup> For chemistry and other physical science conferences, you’ll only be expected to submit a short, one-paragraph description of your proposed work. It’s okay if you don’t know exactly where your research will be at the time of the conference; describe what you think you’ll be able to present. For computer science conferences, you’ll typically be expected to submit a full manuscript that has not been published previously. Before you submit, make sure that all of your co-authors have looked over and approved what you’ve written.
2. At least 3 months before the conference, you should *make logistical arrangements*. This includes booking hotels, flights/train tickets, and/or rental cars. You can make these purchases directly and be reimbursed, but we recommend that you work with purchasers at CMU, who can help make purchases for you.
3. Begin *preparing your presentation* one month before the conference. Make outlines, draft a poster or slides, and, importantly, show them around! Your colleagues and peers can help you understand where your message is clear and where you can improve your communication.
4. (For posters only) plan on *printing your poster* at least one week and ideally two weeks before your presentation. This will give you time to fix things if the first printing somehow doesn’t work out!
5. In the week before your presentation, *practice*. If you’re new to presenting at conferences, we recommend sharing at a group research meeting or creating an *ad hoc* meeting for the purpose of practice. While it’s important that you’re prepared to give your prepared presentation, remember that it’s just as important to be comfortable answering questions on the fly!

#### 13.4.4. Presentation Style

As in writing (see Section 13.2.3), we provide some general style guidelines that you should consider when preparing posters and slide decks. These guidelines are not meant to be overly prescriptive but are designed to help improve your presentations, both in terms of the clarity with which you present and the appreciation that your audience will have for your ideas.

1. **You are the focus.** Your poster or slides should supplement and support what you’re saying, not distract from it. If your audience is paying more attention to your visuals than to you, something is wrong! Common advice to reduce the weight of visuals and reduce the likelihood of distraction:
  - Minimize the use of text on slides
  - One slide, one message. Don’t let the audience linger on an old idea that you’re not talking about anymore, and don’t let them get ahead of you. Keep them on your pace.

---

<sup>41</sup>Smaller conferences sometimes operate on tighter timelines

- Minimize (or, if possible, eliminate) animations. There are rare cases where they're helpful (e.g., a visualization of a molecular dynamics trajectory that clearly illustrates a phenomenon that you're describing), but in general, animations are very likely to distract your audience.
2. **Hold the audience's hand through a journey.** You are an expert who has been working on your research topic for months or years. The audience might have a background in your general area of study, but they almost certainly do not know your topic as well as you do. Even if you pace yourself, you'll probably be giving the audience a lot of new information all at once, so part of your job is to make that information digestible. Introduce concepts at the beginning and as necessary throughout your talk. Try to minimize the amount of knowledge that you are assuming your audience holds. If you're worried that something might be unclear, try to devise a metaphor, visual representation, or concrete example.
  3. **Avoid math unless absolutely necessary.** If text on a slide is a possible source of distraction, math is a sure-fire way to cause confusion. Some folks will immediately stop paying attention to try to interpret your equations. Some will tune out and glaze over. Each time you find yourself adding an equation to a presentation think: is there any other way to convey the same information? It could be an image, a schematic, even some text! If you determine that an equation (or even a series of equations!) is necessary for your presentation, put in the extra effort to walk through those equations slowly and carefully (see point 2 above), explaining each term, what it means, and how each equation leads to the next.
  4. **Motivate throughout your presentation.** Again, your audience will be trying to process a large amount of information in a short amount of time. In the effort to understand your methods or results, to take notes, or to jot down questions, they may lose sight of what came before. Verbally link each idea with the previous one and to your overarching goals.
  5. **Resist the temptation to cram.** It should go without saying, but please, let your instinct be to cut back rather than to add more into a presentation. This is true across the board, whether you're preparing a 10-minute conference presentation or a one-hour seminar talk. If you really think that you won't have enough to say for the whole allotted time, then consider if there are any ideas that you haven't explained fully. From a pedagogical perspective, it's usually better to explain your content more thoroughly than to jam in more new content.
  6. **Anticipate questions.** Most presentations will involve a Q&A section, so there will always be time for your audience to ask questions and for you to clarify any points that caused confusion. However, you shouldn't rely on the Q&A to answer your audience's questions! When you're designing your presentation and practicing, consider what the audience might ask about on each slide<sup>42</sup> and think about if you can incorporate the answer in your presentation. Sometimes, the answer will be "no", for instance because the answer would distract from your main point or could cause more confusion. But even if you find that you can't or don't want to change your presentation, thinking about possible questions is still worthwhile, because you'll be less likely to be surprised when the Q&A comes around.
  7. **Balance explanation and entertainment.** This is a tricky one. You want to keep your audience engaged and energized so that they pay attention and can learn effectively. However, your goal in a presentation is ultimately to express and explain scientific ideas and results. Consider this balance when you're designing a presentation. If you find that an area of your presentation is dry, consider if you can incorporate humor, playfulness, or excitement. On the other hand, when you're practicing, ask your audience if they found anything distracting. Above all, make sure that you're not sacrificing the precision and clarity of your statements for entertainment value.

---

<sup>42</sup>This might be an easier exercise to do with a friend.

#### **13.4.5. Travel Policy**

Air travel is environmentally unsustainable and is nontrivially contributing to anthropogenic climate change [36]. Where possible, we prefer to drive (ideally carpooling) or rely on public transit like trains<sup>43</sup> to get to conference venues. If we do need to fly, we prefer non-stop flights over flights with multiple legs to reduce both travel time and fuel consumption.

As a general rule, we advise that ReACTERs attend at most one international conference per year. ReACTERs wishing to travel internationally more than once in a year should petition at a group organizational meeting, explaining the importance of their international travel for their research or career development.

---

<sup>43</sup>Pittsburgh' Union Station has direct routes to places like Washington, D.C., Philadelphia, and New York City

## 14. Resources for ReACTERs

This is a living, growing list of resources. We are always looking for more resources that might help members of our organization and our community!

### 14.1. Health and Wellness

#### 14.1.1. Disability and Accessibility

- Office of Disability Resources<sup>o</sup>: For information about student disability accommodations.
- Accommodations for Faculty and Staff Members<sup>o</sup>: For staff disability accommodations.
- SLICE<sup>o</sup>: Sign Language Incorporation in Chemistry Education.
- LabSigns<sup>o</sup>: An effort from Pacific Northwest National Laboratory to develop American Sign Language signs for STEM (currently focused on environmental science)

#### 14.1.2. Physical Health

In addition to CMU Health Services<sup>o</sup>:

- The Metro Community Health Center<sup>o</sup> offers medical, dental, and mental healthcare, regardless of your ability to pay.
- Similarly, services at Central Outreach Wellness Center<sup>o</sup> (including HIV/STI testing, PrEP clinics, gender-affirming care) are offered on a sliding-scale basis and will never be denied because someone cannot pay.

#### 14.1.3. Mental Health

- Counseling and Psychological Services (CaPS)<sup>o</sup>: For short-term mental health care on campus. Check out the FAQ here<sup>o</sup>.
- Mental Health Resources<sup>o</sup>: These include general provider searches, as well as crisis resources and specialized resources for disordered eating, substance issues, LGBTQIA+ mental health, and survivors of sexual and physical violence. If you need immediate help, call 412-268-2922 (CaPS), 1-888-796-8226 (the re:Resolve crisis line<sup>o</sup>) 988 (behavior and mental health crisis services), or 911 (medical emergency).

#### **Signs of a mental health crisis (taken from CMU<sup>o</sup> adapted from University of Alaska Anchorage<sup>o</sup>):**

1. Feeling hopelessness, worthlessness, depressed, angry or guilt
2. Withdrawal from friends, family and activities that used to be fun
3. Changes in eating or sleeping patterns
4. Feeling tired or exhausted all of the time
5. Trouble concentrating, thinking, remembering or making decisions
6. Restlessness, irritability, agitation or anxious movements or behaviors
7. Neglect of personal care
8. Reckless or impulsive behaviors (e.g., drinking or using drugs excessively or being unsafe in other ways)
9. Persistent physical symptoms (e.g., headaches, digestive problems or chronic pain) that do not respond to routine treatment
10. Thoughts about death or suicide

If you or someone around you is exhibiting these symptoms and patterns, please seek out help.



## 14.2. Funding

### 14.2.1. Scholarships and Fellowships

- CMU Fundwise<sup>o</sup> organizes a list<sup>o</sup> of funding resources.
- The Johns Hopkins Postdoctoral Funding Database<sup>o</sup>: What it says on the tin. This list is pretty biased towards medical research, so a lot of it might not be relevant to you, but it's always worth checking out!
- The National Science Foundation Graduate Research Fellowship Program (NSF GRFP)<sup>o</sup>
- Hertz Foundation fellowship<sup>o</sup>
- Ford Foundation Fellowships<sup>o</sup>
- Carnegie Bosch Institute Postdoctoral Fellowship<sup>o</sup>
- President's Postdoctoral Fellowship Program<sup>o</sup>
- Schmidt Science Fellowship<sup>o</sup>
- The Beckman Foundation<sup>o</sup> offers fellowships for undergraduates, postdocs, and early-career faculty

### 14.2.2. Research Grants

- The CMU Engineering Research Accelerator<sup>o</sup> maintains a list of funding opportunities<sup>o</sup>
- Grants.gov<sup>o</sup> is a one-stop shop for basically all US federal government grants, including those from NSF, DOE, and NASA.

### 14.2.3. High-Performance Computing Allocations

- The National Energy Research Supercomputing Center<sup>o</sup> offers HPC allocations, but only for those with Department of Energy Basic Energy Sciences<sup>o</sup> research funding.
- ACCESS<sup>o</sup> is the NSF's current HPC platform. You can request compute time even if you don't have *any* funding associated with your proposed project. Note that, in general, only postdocs and faculty can be PIs on ACCESS projects, but an exception is made for graduate students with NSF GRFP funding.
- The National AI Research Resource (NAIRR)<sup>o</sup> program is a new, multi-agency effort to promote so-called "AI" research in the US. It's currently in a pilot stage, so it's unclear what the program will look like in the future.

## 14.3. Research Tools

- OpenAlex<sup>o</sup> is a free catalogue of scientific publications with a powerful API. Think of it as an alternative to Google Scholar.
- Connected Papers<sup>o</sup> is a cool service. Provide information about one journal article and get recommendations for other papers to look at!
- arXivist<sup>o</sup> will send you personalized lists of arXiv<sup>o</sup> preprints based on your interests! A great way to keep up with the literature.

## 14.4. Internships and Jobs

- DOE Science Undergraduate Laboratory Internships (SULI)<sup>o</sup>: Paid internships for students to work at one of the US national laboratories. Summer and semester internships are available.
- NSF Research Experiences for Undergraduates (REU)<sup>o</sup>: Paid research program through NSF. There are REU sites all over the country, so look around and see if anything interests you!
- CMU Highway to Undergraduate Research in the Academic Year (HURAY)<sup>o</sup>: HURAY fellows are paid to do research for one year during the semester. It's a great way to get involved in research while getting compensation to support yourself!

- CMU Career & Professional Development Center (CPDC)<sup>o</sup>: The CPDC can help you practice for interviews and provide career counseling, among other services.
- Chemjobber<sup>o</sup>: Chemjobber is dedicated to helping chemists find jobs and to understanding the chemical job market. His chemistry faculty job spreadsheet is particularly useful if you're looking for jobs in academia!

## 14.5. Other Resources

### 14.5.1. Resources for Students

- CMU Chemical Engineering Graduate Student Association (ChEGSA)<sup>o</sup>: ChEGSA supports grad students in many ways, including hosting social and professional events.
- The ChemE Graduate Student Handbook<sup>o</sup>

### 14.5.2. Conferences and Workshops

- Nessa the Chemist's Conference Database<sup>o</sup>: A curated list, mostly focused on organic and automation chemistry.

### 14.5.3. Legal Resources

- Legal consultation<sup>o</sup>: Every CMU student gets one free 20-minute legal consultation. Consider reaching out here first!
- Affordable legal aid is also offered by:
  - Neighborhood Legal Services<sup>o</sup>
  - The Allegheny County Bar Foundation Pro Bono Center<sup>o</sup>
  - The Pennsylvania Legal Aid Network<sup>o</sup>

## Bibliography

- [1] C. Menkel-Meadow, “Restorative justice: What is it and does it work?,” *Annu. Rev. Law Soc. Sci.*, vol. 3, no. 1, pp. 161–187, 2007.
- [2] B. Fecher and S. Friesike, “Open science: one term, five schools of thought,” *Opening science: The evolving guide on how the internet is changing research, collaboration and scholarly publishing*. Springer International Publishing, pp. 17–47, 2014.
- [3] M. D. Wilkinson *et al.*, “The FAIR Guiding Principles for scientific data management and stewardship,” *Scientific data*, vol. 3, no. 1, pp. 1–9, 2016.
- [4] J. Schofferman, F. T. Wetzel, and C. Bono, “Ghost and guest authors: you can't always trust who you read,” *Pain Medicine*, vol. 16, no. 3, pp. 416–420, 2015.
- [5] J. A. Teixeira da Silva and J. Dobránszki, “Multiple authorship in scientific manuscripts: ethical challenges, ghost and guest/gift authorship, and the cultural/disciplinary perspective,” *Science and engineering ethics*, vol. 22, pp. 1457–1472, 2016.
- [6] M. Liboiron *et al.*, “Equity in author order: a feminist laboratory’s approach,” *Catalyst: Feminism, Theory, Technoscience*, vol. 3, no. 2, pp. 1–17, 2017.
- [7] A. Karamolegkou, J. Li, L. Zhou, and A. Søgaard, “Copyright violations and large language models,” *arXiv preprint arXiv:2310.13771*, 2023.
- [8] T. S. Goetze, “AI Art is Theft: Labour, Extraction, and Exploitation,” *FACCT '24: Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency*, pp. 186–196, 2024.
- [9] S. Chesterman, “Good models borrow, great models steal: intellectual property rights and generative AI,” *Policy and Society*, p. puae6, 2024.
- [10] M. R. Shoaib, Z. Wang, M. T. Ahvanooy, and J. Zhao, “Deepfakes, misinformation, and disinformation in the era of frontier AI, generative AI, and large AI models,” in *2023 International Conference on Computer and Applications (ICCA)*, 2023, pp. 1–7.
- [11] D. Xu, S. Fan, and M. Kankanhalli, “Combating misinformation in the era of generative AI models,” in *Proceedings of the 31st ACM International Conference on Multimedia*, 2023, pp. 9291–9298.
- [12] K. Jaidka *et al.*, “Misinformation, Disinformation, and Generative AI: Implications for Perception and Policy,” *Digital Government: Research and Practice*, 2024.
- [13] B. Perrigo, “Exclusive: OpenAI Used Kenyan Workers on Less Than \$2 Per Hour to Make ChatGPT Less Toxic,” *TIME*, Jan. 2023, Accessed: Jan. 04, 2025. [Online]. Available: <https://time.com/6247678/openai-chatgpt-kenya-workers/><sup>o</sup>
- [14] A. Williams, M. Miceli, and T. Gebru, “The Exploited Labor Behind Artificial Intelligence,” *Noema*, Oct. 2022, Accessed: Jan. 04, 2025. [Online]. Available: <https://www.noemamag.com/the-exploited-labor-behind-artificial-intelligence/><sup>o</sup>
- [15] J. Bartholomew, “Q&A: Uncovering the labor exploitation that powers AI,” *Columbia Journalism Review*, Aug. 2023, Accessed: Jan. 04, 2025. [Online]. Available: [https://www.cjr.org/tow\\_center/qa-uncovering-the-labor-exploitation-that-powers-ai.php](https://www.cjr.org/tow_center/qa-uncovering-the-labor-exploitation-that-powers-ai.php)<sup>o</sup>
- [16] D. Patterson *et al.*, “Carbon emissions and large neural network training,” *arXiv preprint arXiv:2104.10350*, 2021.
- [17] J. McDonald, B. Li, N. Frey, D. Tiwari, V. Gadepally, and S. Samsi, “Great power, great responsibility: Recommendations for reducing energy for training language models,” *arXiv preprint arXiv:2205.09646*, 2022.
- [18] S. Samsi *et al.*, “From words to watts: Benchmarking the energy costs of large language model inference,” in *2023 IEEE High Performance Extreme Computing Conference (HPEC)*, 2023, pp. 1–9.

- [19] P. Li, J. Yang, M. A. Islam, and S. Ren, "Making AI less" thirsty": Uncovering and addressing the secret water footprint of AI models," *arXiv preprint arXiv:2304.03271*, 2023.
- [20] N. Lee, Y. Bang, H. Lovenia, S. Cahyawijaya, W. Dai, and P. Fung, "Survey of social bias in vision-language models," *arXiv preprint arXiv:2309.14381*, 2023.
- [21] Y. Guo *et al.*, "Bias in Large Language Models: Origin, Evaluation, and Mitigation," *arXiv preprint arXiv:2411.10915*, 2024.
- [22] A. Field, S. L. Blodgett, Z. Waseem, and Y. Tsvetkov, "A survey of race, racism, and anti-racism in NLP," *arXiv preprint arXiv:2106.11410*, 2021.
- [23] V. Hofmann, P. R. Kalluri, D. Jurafsky, and S. King, "AI generates covertly racist decisions about people based on their dialect," *Nature*, vol. 633, no. 8028, pp. 147–154, 2024.
- [24] C. B. M. R. Costa-juss and N. Casas, "Evaluating the underlying gender bias in contextualized word embeddings," *GeBNLP 2019*, vol. 33, 2019.
- [25] B. Hutchinson, V. Prabhakaran, E. Denton, K. Webster, Y. Zhong, and S. Denuyl, "Social biases in NLP models as barriers for persons with disabilities," *arXiv preprint arXiv:2005.00813*, 2020.
- [26] V. K. Felkner, H.-C. H. Chang, E. Jang, and J. May, "Winoqueer: A community-in-the-loop benchmark for anti-LGBTQ+ bias in large language models," *arXiv preprint arXiv:2306.15087*, 2023.
- [27] M. T. Hicks, J. Humphries, and J. Slater, "ChatGPT is bullshit," *Ethics and Information Technology*, vol. 26, no. 2, p. 38, 2024.
- [28] S. Subramanian, "The importance of conferences and networking in PhD and postdoctoral training," *Current Protocols Essential Laboratory Techniques*, vol. 21, no. 1, p. e44, 2020.
- [29] D. J. McClements, J. McClements, and I. F. McClements, "Conferences and Networking: Getting to Know the Science and the People Behind It," *How to be a Successful Scientist: A Guide for Graduate Students, Postdocs, and Professors*. Springer, pp. 173–192, 2024.
- [30] N. Kiesler and D. Schiffner, "Conferences are Exclusive by Nature," in *Proceedings of the 2024 on RESPECT Annual Conference*, 2024, pp. 288–292.
- [31] A. Glover, Y. Strengers, and T. Lewis, "The unsustainability of academic aeromobility in Australian universities," *Sustainability: Science, Practice and Policy*, vol. 13, no. 1, pp. 1–12, 2017.
- [32] R. Raven *et al.*, "Transitioning to sustainable academic conferences needs more experimentation and reflection," *Global Sustainability*, vol. 6, p. e16, 2023.
- [33] C. Oliver and A. Morris, "(dis-) Belonging bodies: negotiating outsider-ness at academic conferences," *Gender, Place & Culture*, vol. 27, no. 6, pp. 765–787, 2020.
- [34] N. S. Sabharwal, E. F. Henderson, and R. S. Joseph, "Hidden social exclusion in Indian academia: gender, caste and conference participation," *Gender and Education*, vol. 32, no. 1, pp. 27–42, 2020.
- [35] C. Oliver and A. Morris, "Resisting the "academic circle jerk": precarity and friendship at academic conferences in UK higher education," *British Journal of Sociology of Education*, vol. 43, no. 4, pp. 603–622, 2022.
- [36] S. Gössling and A. Humpe, "The global scale, distribution and growth of aviation: Implications for climate change," *Global Environmental Change*, vol. 65, p. 102194, 2020.