



Centre for
AI and Data Governance

*Early Career
Researchers' Workshop
31 March 2022*

INTRODUCTION

This report reflects upon the inaugural workshop for early career researchers (ECRs) facilitated by the Centre for AI & Data Governance (CAIDG) on 31 March 2022. A summary of the entire event can be found on our [blogpost](#).

The aim of the workshop, organised by the centre’s research associates, was to share and discuss the challenges faced by new or otherwise junior researchers, such as coping with the pressure of producing publications and tackling new research areas or methods. It was also an opportunity for ECRs from various institutions and disciplines to connect with one another.

This report focuses on observations from the following segments of the programme:

- ECR Internal Sharing on “Research Methods and Challenges”
- “Getting Published” with *Big Data & Society* editor Professor Hallam Stevens

Both sessions were an hour-long each. There were 20 ECR participants in attendance from various research institutes across Singapore. All work, or have an interest, in the research fields of AI and data governance and smart cities. Their disciplinary backgrounds include urban planning, social science, law, geography and computer science. They were also at different stages of their professional development, with a handful having obtained or being in the process of obtaining doctorates.



“Unfamiliar research methods and interdisciplinary work are ECRs’ top research concerns”

ECR Internal Sharing on “Research Methods and Challenges”

The sharing session took on a semi-casual roundtable format, facilitated by two CAIDG research associates. ECR participants were free to share as they wished, although the limited time meant that generally only two to three persons responded to each prompt. Prior to the event, many ECRs had cited unfamiliar research methods and interdisciplinary work as their top research concerns.

The discussion and sharing centred around three main topics:

- (i) Empirical data research methods
- (ii) Interdisciplinary research and teams
- (iii) Questions for computer science ECRs

(i) Empirical research methods

[Prompts: How many of you have worked in empirical studies? Out of you who have, would anyone like to share a challenge they faced?]

Approximately half of the ECRs have had experience doing either qualitative or quantitative empirical research. The ECRs who had first-hand experience collecting empirical data had plenty of advice to share with the group. One ECR reflected that thinking systematically about data cleaning was important, for their regression analysis was initially illogical and rectifying this was very time-consuming. On a related point, another ECR shared that recording and organising all research data was difficult, and so learning different tools and software to structure data proved to be helpful for their thesis. Referencing tools, such as Zotero, can be of immense help as well.

While interacting with human participants in fieldwork, it was recognised that inadvertently there would be an element of **positionality**, or worse, the risk of personal bias of the researcher creeping into the research. Namely, the researcher's identity, background, personality and knowledge influenced the dynamic they had with participants, and consequently their responses. For instance, research subjects may hide or play up certain things to project a better positive image of themselves or their community to the researcher.

On this front, “**capturing silence**” was seen as a challenge in fieldwork as silence could be interpreted in many ways, and what was not said could be as important as what was said. For example, in the youth privacy research field, silence from the teenagers could demonstrate their lack of understanding of privacy rights. A researcher shared their supervisor's comment that it is precisely these grey areas that research seeks to uncover and understand. A researcher's self-reflexivity is therefore important, as well as an awareness that their research may not be perfect.

(ii) Interdisciplinary research and teams

[Prompts: Let's talk about interdisciplinary research. What were your top challenges working in a team or when doing solo work?]

Approximately half of the ECRs worked mainly on team-based research projects, while the other half did mainly independent research work (with one or more supervisors).

An ECR who worked both in teams and alone reflected that working alone was more effective for time-sensitive work such as writing geopolitical analyses. However, challenges of working solo include the inability to evaluate an issue holistically. For teams-based research outputs like reports and co-authored work, there could be a productive clash of perspectives arising from different theoretical backgrounds, where team members can address each other's blind spots. Another researcher, a social scientist who recently joined a team of scientists, highlighted the challenge of communicating in a multi-disciplinary team, specifically “**translating**” across different disciplines. Having to familiarise himself/herself with scientific jargon such as CT values, primers and thermos imaging, the researcher sees his/her current role as an excellent opportunity to learn to speak across disciplines in a diverse team.

Another ECR who researched into politics revealed the challenge of teamwork on polarising research topics, for everyone has different views. Working in a group of like-minded people can result in an echo chamber of similar ideas, but working with people who may fundamentally disagree with you poses a real challenge: different personal standpoints may seep into the research and lower its quality, and the work may not present the best middle-ground from the team.

For ECRs who code, teamwork is preferred. According to one researcher, writing code also meant inadvertently writing in bugs, and therefore the practice of **cross-validation** (reading and checking of each other's code) is important. However, the reality in which computing engineers tend to work among themselves (and not so much in interdisciplinary teams) presents its own set of challenges. Domain experts were needed to unpack AI ethics, laws and regulations; it was difficult for engineers to understand and unpick the underlying philosophies and then write code that are in compliance.

A single discipline however may be so broad and diverse that it requires **negotiating a balance between theory and practice** within the team. An ECR observed that the urban planning field involves researchers from different backgrounds – some come from theoretical backgrounds (imagining cities) while others have science backgrounds, and they come together to discuss urban planning in smart cities. This can lead to potential disagreements on what to include in the research output: theories that some find beneficial may not be practical enough for others.

“Writing code also inadvertently means writing in bugs, and therefore the practice of cross-validation is important.”

Another researcher said that interdisciplinary research is only considered to be valuable when it produces policy relevant research. For instance, while philosophy has a lot to contribute to AI ethics, philosophical analysis often has to give way to concrete applications, i.e. what it means for law and policymakers.

Notably, in a team, the quality of the overall writing could either increase or decrease. The challenge of different writing styles within a team resonated with the ECRs present.

The challenges of solo work

For those who worked alone, the relationship they foster with their supervisors becomes understandably significant. One ECR shared their difficulties of working with two supervisors who had differing opinions on drafts of his/her work but did not otherwise communicate their disagreements with each other. Another ECR empathised and said that the **researcher-supervisor power dynamics** could be a challenge: to what extent can one disagree with their supervisor's ideas for a paper? Could an ECR disagree with their supervisor when they have a lot more experience (and authority)? One tactic employed was simply to follow their supervisor's instructions or to include their perspective and "hope they don't notice".



(iii) Questions for computer science ECRs

As the ECRs present were predominantly from social science and humanities backgrounds, the workshop presented an opportunity for them to air specific questions to their peers in computing about the work they did. After all, it is a perceived concern that social scientists (in the fields of AI and data governance and smart cities) do not sufficiently know about tech, although a few of the ECRs revealed to have some knowledge of coding or experience with predictive modelling.

Two particular questions were raised to the computing engineers:

1. Can you explain how you use code to answer your research question(s)?
2. Do you consider bias in your dataset before or during coding?

1. Can you explain how you use code to answer your research question(s)?

Response:

My research work involves building tools to uncover the fairness issues in the AI models. As many models can be accessed simply with the application programming interface (API), which is a line of code, we design our tool to treat the AI systems under testing as black boxes. Our tools have identified many fairness issues in state-of-the-art sentiment analysis systems. For instance, we have found that by changing male names to female names in the input, the model predicted sentiment changes from positive ones to negative ones. My job then is to analyse how and why this happens, by looking at the data and the model. Then, I will design some methods to fix the bias accordingly.

2. Do you consider bias in your dataset before or during coding? While the data can be noisy/biased/unclean, the code itself may be unfair/discriminatory. For example, when banks use postal code to determine approval for bank loans, this can discriminate people based on their socio-economic backgrounds.

Response: Well as a model user or a developer, especially for a model that will be deployed in the real world, we should pay attention to the data itself. But as a tester, we care less about bias when implementing the model, because our tool is meant to uncover the bias. Rather, we say that it is better if the model has bias so that we can detect and understand it. And thereafter, we try to uncover the bias as much as possible.

Thus, the short exchanges above demonstrate how researchers from different disciplines approach their work. In particular, the concept of bias in the field of AI and data governance might be understood differently in the computer science and social science disciplines.

“Getting published in a good journal is important for those interested in a career in academia.”

“Getting Published” with Big Data & Society editor Professor Hallam Stevens

Big Data & Society is one of the distinguished journals in the field, publishing interdisciplinary work about the implications of Big Data for societies.¹ For ECRs, getting published in a good journal is important for those who are interested in a career in academia. The session provided ECRs the opportunity to speak directly with a *BD&S* editor, offering ECRs insights into the sometimes seemingly opaque workings of a journal, as well as advice as to how they might further their careers.

Comments on interdisciplinarity

Professor Stevens first prefaced his session by commenting on interdisciplinary work (in the field of AI and data governance), stressing that *interdisciplinarity should still have a basis in each discipline*. Speaking as a professor of interdisciplinary studies, he said that while different perspectives are needed to address real-life problems, such perspectives need to be grounded from within a discipline – namely having basis in its research methods and literature. As such, Professor Stevens remarked that trying to foster interdisciplinary work at research institutes is not always easy nor successful.

As for interdisciplinary work in the tech field, Professor Stevens noted that a “fundamental inequality” or discrimination between disciplines would be an obstacle to interdisciplinarity in practice. In particular, he highlighted that the

¹ <https://journals.sagepub.com/home/bds>

humanities and the social sciences run the risk of becoming ‘handmaidens’, if engineers only see social scientists as communicators for their work, for example having the role of convincing the public of decisions that are already made, such as the introduction of autonomous vehicles.

“Interdisciplinarity should still have a basis in each discipline.”

Instead, true collaboration between disciplines would mean that social scientists are involved from the beginning, with their comments and input being taken seriously and factored into the project. For instance, the project should be grounded in literature (which is for the most part grounded in disciplines) and if it involves the collection of empirical data, the data should be collected according to some kind of disciplinary norm.



The same applies to paper outputs. Professor Stevens pointed out that good papers are ultimately still grounded in specific methods that come from specific disciplines, both methodologically and in the literature review. Producing papers that showcase interdisciplinary work does not mean that researchers are licensed to abandon the methodological aspects and uniqueness of each discipline and “throw everything together” into a paper.

Tips for getting published

Trite as it sounds, the first thing to do is to find journals that are suitable avenues for the ECR’s work. While some might try approaching the journal directly, Professor Stevens said that this is not necessarily effective. A more intuitive way would be to read the journal, understand its scope and aims, and ask around for advice (i.e. tapping into one’s research circles). Ultimately, it would be helpful to keep in mind the journals that correspond to one’s research interests.

Reading the journals of interest, as well as reading widely in general, is likely to pay off: this helps to determine the topics of research that are being published and are therefore publishable. Further, reading the journals would help in appreciating the scope of research coverage that would be deemed appropriate. Scope is always a big issue, noted Professor Stevens.

Having determined the scope that a piece of work aims to cover, Professor Stevens said that it must be made clear to the reviewers and editors what its contribution is and how it may be advancing the field or adding to the literature. This could mean presenting a new dataset that no one else has, or applying an idea to a new geographical context. Whatever it is, the work's contribution must be made clear (if the editors are not convinced about this, it is unlikely for them to accept the article).

One ECR asked about which journals they should look out for if they are theorists “in a sea of engineers” within their team. Professor Stevens responded that are interdisciplinary journals that publish on AI and big data. It is about finding the right journals, and to do this it is helpful to talk to mentors. One should also not preclude the journals in their own discipline (such as philosophy or law journals) if their work is better suited there, said Professor Stevens (*Editors' note: this drives home the message that interdisciplinarity has its basis in disciplines!*).

“For an early career researcher, simply getting started in the publishing game is paramount.”

Having laid out these general rules of thumb, Professor Stevens pointed out that for an early career researcher, simply getting started in the publishing game is paramount. While journal articles are the gold standard, the peer review process can be awful or even dysfunctional, and as such they are viewed as a “necessary evil”. There are alternatives to journals that may be easier, such as special collections in books, which have a different peer review process. In fact, better work might get published in these avenues, said Professor Stevens. Such opportunities may arise from collaborations where an institute gets invited to write a chapter, and also via networking and participating in conferences. It is important to realise that journal publishing is only one medium of publishing, that is often too linked to old-fashioned university values, and perhaps it is more crucial to have one's ideas read and commented upon through different pathways, such as conference submissions, commentaries and blog posts, as these are also productive ways to get started before writing gets refined for eventual release.

Professor Stevens said to the audience of ECRs that one of the pitfalls of young researchers is getting distracted from their publishing goals. His piece of advice: Be strategic, identify the best journals for you, don't get distracted and “go all out”. To

hit their publishing goals, it would be best to have multiple works in the pipeline in different stages, so that if one does not get accepted, there are others to submit.

Handling the peer review process

It was acknowledged by everyone that the peer review process places the author in an unequal power dynamic in relation to the reviewer. An ECR raised that while the comments they receive can be constructive, sometimes they are not, which makes handling such comments challenging or even frustrating. Professor Stevens agreed that this was precisely why peer review can be dysfunctional, as it does not always encourage polite feedback. Speaking from personal experience, he said “I read the reviews and then put them in a drawer!”, thereby highlighting the importance of taking a healthy distance from the feedback. This way, one might be in a better frame of mind to review the feedback and address the concerns of the reviewers. While the criticism may seem negative, a detailed review demonstrates effort on the reviewers’ part to deeply engage with the work.



Therefore, Professor Stevens advised that it is always good to adopt a positive attitude and demonstrate a willingness to accommodate and accept critical reviews. With that said however, Professor Stevens also highlighted that abiding by publication ethics is important, and that authors can refer to the COPE principles on how editors and reviewers should behave,² and report behaviour that may stray from these principles.

“I read the reviews and then put them in a drawer!”

Another ECR asked how one could navigate the challenge of their reviewers not having background knowledge in their research topic. In such a scenario, Professor Stevens suggested that the author could respond to both the journal’s reviewers and

² <https://publicationethics.org/core-practices>

the editors; the editor, in particular, usually has a lot of discretion in the final publication decision. He recommended reading the conditions in the journal about how they derive a decision. Authors can also seek clarification regarding this. For instance, if Reviewer 1 suggests minor revisions while Reviewer 2 rejects the work from the get-go, the author could check with the editor whether it is a necessary condition to satisfy Reviewer 2 in order to get published. Authors are well-entitled to express their standpoints to the editors. If it would be unrealistic to get a reviewer to change their position, an author might even choose to withdraw, said Dr Stevens.

Finally, in the closing of his session, Professor Stevens encouraged ECRs particularly from social science and humanities backgrounds to make their arguments to their science and tech peers confidently, and to highlight and discuss within interdisciplinary teams the socio-ethical problems of technology. It is within this crucible that interdisciplinary work can flourish.

“It is important for ECRs to highlight and discuss within interdisciplinary teams the socio-ethical problems of technology.”

CONCLUDING THOUGHTS

The conversations in these sessions have laid bare the challenges of research in socio-technical fields, namely the issues of positionality and capturing silence when engaging with the research subject in empirical work, as well as translating/communicating across disciplines and grounding the research in each discipline in interdisciplinary work, which need to be reflected upon for interdisciplinary work to have its envisioned impact. In addition, ECRs might be uninitiated to challenges such as dealing with the publications or with supervisors, and forums such as this workshop provide the opportunity to troubleshoot these concerns.

It is also clear from these meta-conversations that sharing across disciplines is a great way to promote cross-disciplinary understanding, such as appreciating the different approaches to AI

bias. All in all, an ECR should expect being exposed to unfamiliar research methods and challenges as part of his/her journey. In these emerging fields of AI, data governance and smart cities, continued conversations among researchers, and also others, will be very much required – an overall learning experience in order to achieve true interdisciplinarity.

Centre for AI and Data Governance
21 April 2022

(Report prepared by Ong Li Min and Sharanya Shanmugam with the assistance of Felicia Lee and Sivakami d/o Arunachalam. With thanks to Professor Hallam Stevens, Willow Wong and Professor Mark Findlay for reviewing the draft.)



Centre for
AI and Data Governance