An Interview with Claire Benn

Claire Benn is an Assistant Professor at the University of Cambridge and Course Leader of the MPhil in Ethics of AI, Data, and Algorithms at the Leverhulme Centre for the Future of Intelligence. In this interview, Editor-in-Chief and current MPhil student Mahera Sarkar sits down with Dr. Benn to reflect on the programme's first year.

What motivated you to create the MPhil alongside the MSt?

The primary motivation was to create a more traditional Master's experience through an intensive, in-person programme. The MSt caters to a more professionally established audience, while the MPhil aims to open this opportunity to a broader group of people who may not have as much professional experience but are still passionate about AI Ethics.

Which seminar paper was your favourite this year and why?

I made a conscious decision to attend every seminar paper. With a background in both philosophy and experimental psychology, I found it fascinating to gain new conceptual understanding, especially from John Burden's course on Evaluation of AI Systems. John did a fantastic job bridging the technical and theoretical assumptions. His seminars highlighted overlaps between his own work and my module's perspectives, and made me realise we needed a stronger technical foundation in our curriculum. As a result, John will run a technical module at the beginning of next year to introduce students to the basics of algorithms, machine learning, and AI.

As Course Leader, what challenges did you face during the MPhil's first year?

There were a few challenges during the inaugural year. As the sole leader of the course, I've always worried that I might lean too much towards my areas of expertise. Finding a balance when teaching between covering material that was not too specialised for some but also not too obvious for others was tricky. Additionally, the administrative burden was significant. Building infrastructure and anticipating how to tailor our programme while incorporating insights from other departments were crucial but challenging tasks. One of the

I've been impressed and thrilled with how the cohort has bonded. Students spend time

most difficult aspects was finding a balance between being flexible and setting precedents for fairness. Ensuring individual student needs are met whilst establishing policies that are equitable is extremely important and something that I am constantly evaluating.

Was the programme's interdisciplinary approach beneficial or challenging?

Interdisciplinarity is essential. It's vital for quality research and understanding these complex issues. Exposure to different academic backgrounds only enriches discussion, especially in the course's Work in Progress seminars. It's valuable for those entering public corporations due to the breadth of perspectives. However, maintaining rigour without disciplinary bias and setting a standard for marking can be challenging. Those pursuing further study may also find it difficult to re-integrate into discipline-focused programmes. That's why we created a balance between the introductory module, which is purposefully broad, and discipline-specific electives to ensure students still get to specialise whilst gaining the skills to critique their own fields.

Did the cohort bond well?

together and share their work. Support groups have organically formed, and students show genuine interest in each other's presentations.

The discussions we have had as a class have been spirited and engaging and I am so pleased that it has been driven by the students themselves.

Why is the MPhil assessed entirely through research?

There was never a debate about exams, which require a core corpus of shared knowledge. Instead, I wanted students to develop a shared vocabulary but also have the confidence and skills to produce novel specialised research. The focus on coursework allows them to build skills like project management, self-reflection, and recursive improvement. I don't want to tell people what to write about. Instead, students tell me what they're working on - the range of incredible projects have blown my mind.

What role do you see graduates of the course playing in the future development of AI?

I have high hopes for graduates of this course shaping AI discourse across multiple avenues. This could occur in public policy, private companies, or academia. I also envision a "soft influence" through a cultural shift addressing AI concerns. I often find public discourse to be pitched too high or overly sensationalised so by having those workplace or dinner-table discussions, I see this cohort as custodians and gateways for meaningful public engagement.

Looking ahead, what are your priorities and goals for the MPhil's second year?

As I said earlier, John Burden will lead a technical introductory module, and we're considering a new elective on legal, regulatory, and policy aspects. We will definitely incorporate feedback from the current cohort to refine the curriculum. I also plan to create a college family structure to connect the new cohort with past students. Finally, I am passionate about finance not being a barrier to this opportunity so hopefully we can secure more funded places for future students.

Do you have any advice for prospective applicants?

When reviewing admissions, I simply want to find people who will enjoy the course and do well. It is my responsibility as Course Leader to ensure students succeed and thrive during the course. I therefore encourage applicants to show their suitability through prior experience or enthusiasm. Given that the course is solely assessed on the basis of academic writing, applicants should demonstrate strong writing skills and thrive in an intensive, independent research environment. The level of independence that you have in this course is both an incredible opportunity and a challenge. As a result, people should be flexible but be prepared to be decisive and settle on what they want to pursue.

Any book recommendations?

I'm a huge fan of sci-fi and fantasy, and I believe in reading whatever you're passionate about – even if it is not specifically about AI. In my lectures this year, I spoke about Ursuala Le Guin's 'The Left Hand of Darkness". It has nothing to do with technology and yet it makes you reflect on what it means to be human, which is just as important as any AI-specific literature.

For further details on the MPhil in Ethics of AI, Data, and Algorithms, visit:
https://www.lcfi.ac.uk/education/mphil