

Research Interests

I have recently passed my PhD with integrated study at the University of Edinburgh, supervised by Björn Ross, Vaishak Belle and Zachary Horne. In my thesis I address how to measure bias in natural language processing (NLP) technologies using a human-centric approach, as it is truly human behaviour that determines the real world impact of these tools. My research spans from how the public responds to biased AI technologies, to how to measure bias in a psychologically grounded fashion, to analysing the ethical risks of generative models. My linguistics and psychology background affords me a unique perspective on the topics of fairness and ethics in NLP, in that I use social science theory and rigorous experimental design to gain a nuanced understanding of NLP harms. I have experience conducting large-scale automatic evaluations of safety issues in large language models, as well as human evaluation.

Publications

1. Eddie L. Ungless, Sunipa Dev, Cynthia L. Bennett, Rebecca Gulotta, Jasmijn Bastings and Remi Denton. 2025. *Amplifying Trans and Nonbinary Voices: A Community-Centred Harm Taxonomy for LLMs* In Proceedings of the Association for Computational Linguistics: ACL 2025. CORE: A*
2. Alice Ross, Cliodhna Hughes, Eddie L. Ungless and Catherine Lai. 2025. *Vocal cues to perceived gender: insights from trans men* Interspeech 2025
3. Eddie L. Ungless, Zachary Horne and Björn Ross. 2025. *"Till I can get my satisfaction": Open Questions in the Public Desire to Punish AI*. Extended Abstract to appear at CHI 2025. CORE: A*
4. Eddie L. Ungless, Nikolas Vitsakis, Zeerak Talat, James Garforth, Björn Ross, Arno Onken, Atoosa Kasirzadeh and Alexandra Birch. 2024. *The Only Way is Ethics: A Guide to Ethical Research with Large Language Models*. COLING 2025
5. Eddie L. Ungless, Nina Markl and Björn Ross. 2024. *Experiences of Censorship on TikTok Across Marginalised Identities*. ICWSM 2025
6. Atli T. Sigurgeirsson and Eddie L. Ungless[†]. 2024. *Just Because We Camp, Doesn't Mean We Should: The Ethics of Modelling Queer Voices*. Interspeech, 2024.
7. Charlotte Bird and Eddie L. Ungless[†], and Atoosa Kasirzadeh. 2023. *Ethical and Social Risks of Generative Text-to-Image Models*. In Proceedings of the 2023 AAI/ACM Conference on AI, Ethics, and Society.
8. Seraphina Goldfarb-Tarrant and Eddie L. Ungless[†], Esma Balkir and Su Lin Blodgett. 2023. *This prompt is measuring <mask>: evaluating bias evaluation in language models*. In Findings of the Association for Computational Linguistics: ACL 2023. CORE: A*
9. Eddie L. Ungless, Anne Lauscher and Björn Ross. 2023. *Stereotypes and Smut: The (Mis)representation of Non-cisgender Identities by Text-to-Image Models*. In Findings of the Association for Computational Linguistics: ACL 2023. CORE: A*
10. Eddie L. Ungless, Björn Ross, and Vaishak Belle. 2023. *Potential Pitfalls With Automatic Sentiment Analysis: The Example of Queerphobic Bias*. Social Science Computer Review. Impact Score: 4.1
11. Eddie L. Ungless, Amy Rafferty, Hrichika Nag, and Björn Ross. 2022. *A Robust Bias Mitigation Procedure Based on the Stereotype Content Model*. In Proceedings of the Fifth Workshop on Natural Language Processing and Computational Social Science (NLP+CSS)

[†] joint first author

Invited Talks

May 2025. Sex and Gender Research Meeting, University of Edinburgh: Invited to speak on gender bias in AI

April 2025. Loch Lomond Trust Futures Group: Invited to speak on ethical risks of adopting AI technologies

March 2024. Scottish AI Summit, Edinburgh: Invited to be part of a panel on queer data and generative AI.

Research Experience

October 2024 - February 2025. Student Researcher, Google Research: with Remi Denton and Sunipa Dev. Created granular taxonomy of harms to trans people from language models, based on a literature review, community survey and focus groups. Proposed pragmatic evaluation heuristics.

February - July 2024; March to May 2025. Research Assistant, University of Edinburgh, Informatics School: with Alexandra Birch and Atoosa Kasirzadeh. Produced a whitepaper on how to conduct ethical research with Large Language Models, including a literature review of useful resources. Later developed into public wiki.

January - March 2023. Research Assistant, University of Edinburgh, PPLS School: with Atoosa Kasirzadeh. Helped prepare a policy document for the Department for Culture, Media & Sport on the topic of ethical risks of text-to-image models. Conducted a literature review on the risks of discrimination, exclusion and harm.

2021-23. Developer, Biascan: Developed a prototype for a system which detects implicit bias in jobs listings, and suggests alternative wording.

Teaching Experience

July 2025. Postgraduate course, University of Hamburg. Guest Lecturer on Ethics for 'Ethics and Modern AI'.

January 2023, 2024. Postgraduate course, School of Informatics, University of Edinburgh. Guest Lecturer on Ethics for 'Doing Research in NLP'.

January - April 2023, 2024. Postgraduate course, School of Informatics, University of Edinburgh. TA for 'Case Studies in AI Ethics'.

October - December 2022, 2023. Undergraduate course, School of Informatics, University of Edinburgh. Tutor & Marker for 'Professional Issues' (Ethics Course).

October - December 2021, 2022, 2023. Postgraduate course, School of Informatics, University of Edinburgh. TA, Tutor & Guest Lecturer for 'Accelerated NLP'.

Professional Experience

February - August 2019. HYD, digital production agency (London, England). Deputy Head of Strategy and Innovation (Full-time).

- Managerial responsibilities: Responsible for day-to-day management of team of 4, including training new staff members, in addition to responsibilities of strategist role detailed below.
- Pitching products and services: Led pitches from conceptualisation to delivery.

August 2016 - February 2019. HYD, digital production agency (London, England). Strategist then Senior Strategist (Full-time).

- Developing product concepts: Created digital product concepts with a research led approach.
- Pitching products and services: Pitched to clients internationally, including Renault, Canon, Pepsi Co, BBC. This entailed liaising with internal production teams to ensure we presented attractive and competitive solutions.
- UX consultant: Worked closely with the design team to establish optimum UX.

Other Relevant Experience

- Reviewer for ACL, FAccT, AIES, WiNLP (ACL), TOCHI
- Given talks on public engagement and public speaking to CDT peers

Technical Skills

- Extensive experience with Python including finetuning, and intrinsic and extrinsic evaluation LLMs, confident using key NLP libraries such as pytorch, NLTK etc
- Static site design with Jekyll and Hugo; limited HTML/CCS
- Statistical analysis in SPSS, Smart-PLS, R
- Thematic coding in NVivo, Atlas

Education

2020 - 2025. PhD with integrated study in the School of Informatics. University of Edinburgh, UKRI CDT in NLP Edinburgh, Scotland.

Courses: Included applied machine learning, NLP, research methods, AI ethics.

2019 - 2020. MSc - Psychological Sciences (Distinction). UCL, London, England.

Courses: Included introduction to Python, experimental design, statistical analysis for psychology, qualitative analysis techniques, social psychology.

2013 - 2016. BA Hons - Linguistics (First Class). University of Cambridge, Cambridge, England.

Courses: Included computational linguistics, generative syntax, semantics, formal logic.