

GRETA WARREN

Postdoctoral Researcher, Department of Computer Science, University of Copenhagen

grwa@di.ku.dk ♦ gretawarren.github.io ♦ [ORCID: 0000-0002-3804-2287](https://orcid.org/0000-0002-3804-2287)

Scientific interests: Human-Centred AI, Explainable AI, Natural Language Processing, Fact-Checking, Misinformation

EMPLOYMENT

Postdoctoral Researcher, Department of Computer Science, University of Copenhagen, Denmark **Mar 2024 - Present**

- Member of Copenhagen Natural Language Understanding research group and ExplainYourself project team (funded by European Research Council) researching human-centred explanations for automated fact-checking.
- Leading interdisciplinary human-computer interaction and computational research using quantitative and qualitative methods, collaborating with experts in NLP, HCI, and fact-checking stakeholders to address diverse needs for explainable AI.

Research Assistant, Trinity College Institute of Neuroscience, Trinity College Dublin, Ireland **Sep 2018 - Aug 2019**

- Recruited participants and collected electroencephalography (EEG), pupillometric and neuropsychological data, conducted data entry, processing and analysis, compiled hospital research ethics applications, delivered talk on brain injury research to patients and staff in National Rehabilitation Hospital, Dublin

EDUCATION

Ph.D., School of Computer Science, University College Dublin, Ireland **Sep 2019 - Jan 2024**

- Thesis: *User-centred counterfactual explanations for explainable AI*
- Supervisors: Prof. Mark Keane and Prof. Ruth Byrne.
- Examiners: Prof. Tim Miller (University of Queensland) and Prof. David Coyle (University College Dublin)
- Awarded multiple travel scholarships (€6000+; ACM SIGCHI, ACM FAccT, ICCBR)
- Collaborated with Accenture Labs, The Dock, Dublin on applied explainable AI projects (2022 - 2023)

B.A. (Hons), Psychology, Trinity College, University of Dublin, Ireland **Sep 2014 - May 2018**

- Thesis: *Effects of additive and subtractive counterfactual thinking on choice of political candidate*
- Supervisor: Prof. Ruth Byrne

TEACHING EXPERIENCE

Thesis Advisor, Department of Computer Science, University of Copenhagen **Jan 2025 - Sep 2025**
Copenhagen, Denmark

- Stefan Gottlieb Kramer, MSc. in IT and Cognition. Primary supervisor for Thesis: *Using Graph Neural Networks for Usable, Explainable Source Reliability Determination in Fact-checking*. Jan 2025-Sep 2025. **Grade: 10/12**

Internal Censor/Examiner, Department of Computer Science, University of Copenhagen **June 2025 - Present**
Copenhagen, Denmark

- Assessed and graded MSc. student thesis projects in Department of Computer Science.

Teaching Assistant, School of Computer Science, University College Dublin, Dublin, Ireland **Sep 2020 - Dec 2022**

- 2020-2022 Text Analytics, MSc course, coordinated demonstrators and administration, prepared lab materials, facilitated lab sessions and graded assignments
- 2021-2022 Data Science in Python, MSc course, prepared lab materials and facilitated lab sessions
- 2020-2022 Introduction to Programming, BSc course, prepared lab materials, facilitated lab sessions and graded assignments
- 2020-2021 Programming I & Object-Oriented Programming, MSc course, prepared lab materials, facilitated lab sessions and graded assignments

SKILLS

Quantitative Research Methods: Designing and conducting large-scale psychological & behavioural experiments (Prolific Academic, Alchemer, Survey Monkey, Qualtrics). Computational analysis of large scale datasets (e.g., social media posts) using NLP techniques.

Qualitative Research Methods: Gathering rich insights through semi-structured interviews, think-aloud studies, focus groups using methods from grounded theory & thematic analysis.

Software: Python, pandas, scikit-learn, nltk, R \LaTeX , HTML, SPSS, NVivo.

Languages: English (Native), Irish (Native), French (B1), Danish (A2)

HONOURS & AWARDS

DAAD AInet Fellowship , German Federal Ministry of Research	Nov 2025
Best Paper Award , Pre-ACL Workshop, Danish Data Science Academy	July 2025
Gary Marsden Travel Award , ACM SIGCHI	€4,000 Mar 2023
Doctoral Consortium Scholarship , ICCBR 2022	€500 Sep 2022
Travel Award , ACM FAccT 2022	€1,500 Jun 2022
1st Prize, Arthur Cox Alternative Perspectives Essay Competition , Trinity College Law Review	May 2018
First Class Award , Trinity College Dublin	Sep 2015
Entrance Exhibition , Trinity College Dublin	€150 Sep 2014
All Ireland Scholarship , All Ireland Scholarship Trust	€27,000 Sep 2014

PROFESSIONAL SERVICE

Associate Chair

ACM CHI Conference on Human Factors in Computing Systems	2026
<i>Understanding People: Statistical & Quantitative Methods Subcommittee</i>	

Programme Committee Member / Reviewer

Association for Computational Linguistics (ACL) Rolling Review*	2025-
ACM CHI Conference on Human Factors in Computing Systems*	2025-
Annual Meeting of the Cognitive Science Society (CogSci)	2024-
ACM Conference on Intelligent User Interfaces (IUI)	2023-
AAAI/ACM Conference on Artificial Intelligence, Ethics, and Society (AIES)	2022-
Workshop on Explainable AI at International Joint Conference on Artificial Intelligence (XAI @ IJCAI)	2023-2024
Workshop on Cognitive Aspects of Knowledge Representation at Knowledge Representation (CAKR @ KR)	2023
Irish Conference on Artificial Intelligence and Cognitive Science (AICS)	2021

*Special recognition for outstanding review (2025)

Journal Reviewing

International Journal of Human-Computer Interaction	2025
Philosophy & Technology	2024
Decision Support Systems	2023
Expert Systems	2022

DEPARTMENTAL SERVICE

PhD Interview Panel, Department of Computer Science & Statistics, May 2025
Trinity College, University of Dublin

- Assessed and interviewed candidates for Irish Central Bank PhD Programme in AI & Data Science

Postdoctoral Network Committee, Department of Computer Science, Nov 2024–
University of Copenhagen

- Organised career-support and networking events for postdoctoral researchers

PUBLICATIONS

* denotes joint first authorship

Conference Proceedings

- Nadav Borenstein*, **Greta Warren***, Desmond Elliott, and Isabelle Augenstein. Can community notes replace professional fact-checkers? In Wanxiang Che, Joyce Nabende, Ekaterina Shutova, and Mohammad Taher Pilehvar, editors, *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 535–552, Vienna, Austria, July 2025. Association for Computational Linguistics. URL: <https://aclanthology.org/2025.acl-short.42/>, doi:10.18653/v1/2025.acl-short.42 [Core A*]
- **Greta Warren**, Irina Shklovski, and Isabelle Augenstein. Show me the work: Fact-checkers' requirements for explainable automated fact-checking. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, CHI '25, New York, NY, USA, 2025. Association for Computing Machinery. doi:10.1145/3706598.3713277 [Core A*]
- **Greta Warren**, Eoin Delaney, Christophe Guéret, and Mark T. Keane. Explaining multiple instances counterfactually: User tests of group-counterfactuals for XAI. In Juan A. Recio-Garcia, Mauricio G. Orozco-del Castillo, and Derek Bridge, editors, *Case-Based Reasoning Research and Development*, pages 206–222, Cham, 2024. Springer Nature Switzerland. doi:10.1007/978-3-031-63646-2_14
- **Greta Warren**, Ruth M. J. Byrne, and Mark T. Keane. Categorical and continuous features in counterfactual explanations of AI systems. In *Proceedings of the 28th International Conference on Intelligent User Interfaces*, IUI '23, pages 171–187, New York, NY, USA, 2023. Association for Computing Machinery. doi:10.1145/3581641.3584090 [Core A]
- **Greta Warren**, Barry Smyth, and Mark T. Keane. Better counterfactuals, ones people can understand: Psychologically-plausible case-based counterfactuals using categorical features for explainable AI (XAI). In Mark T. Keane and Nirmalie Wiratunga, editors, *Case-Based Reasoning Research and Development: 30th International Conference, ICCBR 2022, Nancy, France, September 12–15, 2022, Proceedings*, pages 63–78, Berlin, Heidelberg, 2022. Springer-Verlag. doi:10.1007/978-3-031-14923-8_5

Journal Articles

- **Greta Warren**, Ruth M. J. Byrne, and Mark T. Keane. Categorical and continuous features in counterfactual explanations of AI systems. *ACM Trans. Interact. Intell. Syst.*, 14(4), December 2024. doi:10.1145/3673907 [Q2 rank]

- Catherine N. Moran, David P. McGovern, **Greta Warren**, Rónán Ó Grálaigh, Joanne P. M. Kenney, Alan Smeaton, and Paul M. Dockree. Young and restless, old and focused: Age-differences in mind-wandering frequency and phenomenology. *Psychology and aging*, 36(2):252, 2021. doi: [10.1037/pag0000526](https://doi.org/10.1037/pag0000526) [Q1 rank]

Peer-reviewed Workshops and Symposia

- Nadav Borenstein*, **Greta Warren***, Desmond Elliott, and Isabelle Augenstein. Can community notes replace professional fact-checkers? In *CHI'25 Workshop on News Futures*, 2025. URL: <https://arxiv.org/abs/2502.14132>, arXiv:2502.14132
- **Greta Warren**, Mark T. Keane, Christophe Gueret, and Eoin Delaney. Explaining groups of instances counterfactually for XAI: A use case, algorithm and user study for group-counterfactuals. *IJCAI-23 Workshop on Explainable Artificial Intelligence (XAI)*, 2023. doi:10.48550/arXiv.2303.09297
- **Greta Warren**, Mark T. Keane, and Ruth M. J. Byrne. Features of explainability: How users understand counterfactual and causal explanations for categorical and continuous features in XAI. In *IJCAI-ECAI'22 Workshop: Cognitive Aspects of Knowledge Representation*, 2022. doi:10.48550/arXiv.2204.10152
- Jörg Cassens, Lorenz Habenicht, Julian Blohm, Rebekah Wegener, Joanna Korman, Sangeet Khemlani, Giorgio Gronchi, Ruth M. J. Byrne, **Greta Warren**, Molly S. Quinn, and Mark T. Keane. Explanation in human thinking. In *Proceedings of the 43rd Annual Meeting of the Cognitive Science Society*, 2021. URL: <https://escholarship.org/content/qt9k6291nk/qt9k6291nk.pdf>

Preprints

- Jingyi Sun*, **Greta Warren***, Irina Shklovski, and Isabelle Augenstein. Explaining sources of uncertainty in automated fact-checking, 2025. URL: <https://arxiv.org/abs/2505.17855>, arXiv:2505.17855
- Isabelle Augenstein, Michiel Bakker, Tanmoy Chakraborty, David Corney, Emilio Ferrara, Iryna Gurevych, Scott Hale, Eduard Hovy, Heng Ji, Irene Larraz, Filippo Menczer, Preslav Nakov, Paolo Papotti, Dhruv Sahnan, **Greta Warren**, and Giovanni Zagni. Community moderation and the new epistemology of fact checking on social media, 2025. URL: <https://arxiv.org/abs/2505.20067>, arXiv:2505.20067
- **Greta Warren**, Mark T. Keane, and Ruth M.J. Byrne. Preferences for simple explanations in diagnoses and predictions. 2024. doi:10.31234/osf.io/738ne

Invited Talks

- **Greta Warren** (December, 2025). “I’m not sure because...”: Communicating uncertainty in human-AI collaboration. *Metacognition in Generative AI Workshop, EurIPS Conference, Copenhagen, Denmark*
- **Greta Warren** (December, 2022). Simplicity and complexity in explanations of diagnoses and predictions. *Reasoning and Imagination Lab, Trinity College Dublin, Ireland*
- **Greta Warren** and Eoin Delaney (November, 2022). Group counterfactual explanations for AI predictions. *Accenture Labs at The Dock, Dublin, Ireland*

Panel Discussions

- **Greta Warren** (December, 2025). Topic: Metacognition in Humans and Generative AI. *Metacognition in Generative AI Workshop, EurIPS Conference, Copenhagen, Denmark*

Oral Presentations

- **Greta Warren** (April, 2025). Show Me the Work: Fact-Checkers’ Requirements for Explainable Automated Fact-Checking. *CHI Conference on Human Factors in Computing Systems, Yokohama, Japan*
- **Greta Warren** (April, 2025). Can community notes replace professional fact-checkers? *News Futures Workshop @ CHI 2025, Yokohama, Japan*

- **Greta Warren** (March, 2025). Show Me the Work: Fact-Checkers' Requirements for Explainable Automated Fact-Checking. *Nordic Pre-CHI Conference, Copenhagen, Denmark*
- **Greta Warren** (March, 2023). Categorical and continuous features in counterfactual explanations of AI systems. *International Conference on Intelligent User Interfaces, Warrane/Sydney, Australia*
- **Greta Warren** (July, 2022). Features of explainability: How users understand counterfactual and causal explanations for categorical and continuous features in XAI. *Workshop on Cognitive Aspects of Knowledge Representation at International Joint Conference on Artificial Intelligence, Vienna, Austria*
- **Greta Warren** (September, 2022). Counterfactual Explanations for Explainable AI (XAI). *Doctoral Consortium at International Conference on Case-Based Reasoning, Nancy, France*
- **Greta Warren** (June, 2021). Counterfactual Explanations in Explainable AI. *Annual Meeting of the Cognitive Science Society, Online*

Poster Presentations

- **Greta Warren** and Nadav Borenstein (December, 2025). Can community notes replace professional fact-checkers? *ELLIS Unconference, Copenhagen, Denmark*
- **Greta Warren** and Nadav Borenstein (July, 2025). Can community notes replace professional fact-checkers? *Association for Computational Linguistics (ACL), Vienna, Austria*
- **Greta Warren** and Nadav Borenstein (July, 2025). Can community notes replace professional fact-checkers? *Pre-ACL Workshop, Copenhagen, Denmark*
- **Greta Warren** and Nadav Borenstein (June, 2025). Can community notes replace professional fact-checkers? *Copenhagen NLP Symposium, Copenhagen, Denmark*
- **Greta Warren** (July, 2024). Explaining multiple instances counterfactually: User tests of group-counterfactuals for XAI. *International Conference on Case-Based Reasoning, Mérida, Mexico*
- **Greta Warren** (March, 2023). Categorical and continuous features in counterfactual explanations of AI systems. *School of Psychology Research Symposium, Trinity College Dublin, Ireland*
- **Greta Warren** (September, 2022). Better counterfactuals, ones people can understand: Psychologically-plausible case-based counterfactuals using categorical features for explainable AI (XAI). *International Conference on Case-Based Reasoning, Nancy, France*