Gavin Abercrombie

Employment

2023-present	School of Mathematical and Computer Sciences, Heriot-Watt University, Scotland **Assistant Professor**

2020-2023   School of Mathematical and Computer Sciences, Heriot-Watt University, Scotland
Research Associate
2022-2025: EPSRC project: *Equally Safe Online*
2020-2022: EPSRC project: *Designing Conversational Assistants toReduce Gender Bias*
Principal Investigators: Verena Rieser, Ioannis Konstas

2017-20   School of Computer Science, University of Manchester, England
Research Assistant
Principal Investigator: Riza-Batista Navarro

2014-16   Centre for Language Technology, University of Copenhagen, Denmark
Research Assistant
Principal Investigator: Anders Søgaard

2006-13   The British Council
Teacher and Teacher Trainer

Education

2017-20   University of Manchester
PhD Natural Language Processing
Thesis: *Topic-centric sentiment analysis of UK Parliamentary debates.*
Supervisor: Riza Batista-Navarro; co-supervisor: Goran Nenadic

2013-16   University of Copenhagen
MSc IT & Cognition
Masters thesis: *Automatic sarcasm detection in context: The effects of class imbalance and manual labelling on supervised machine classification of Twitter conversations.*
Supervisor: Dirk Hovy

1995-99   Edinburgh College of Art
BA Design & Communication

Research visits

2021-22   Milan NLP group, Boconni University, Milan, Italy.
2018   Data & Web Science Group, University of Mannheim,, Germany.

Teaching

Heriot-Watt University

2023	Data Mining & Machine Learning
2023, 2024	F21CA Conversational Agents

2020-present	F21RP Research Methods and Requirements Engineering (course leader)
	F21/20CA Conversational Agents

University of Manchester

2018-20	Teaching Assistant:
	MSc COMP61332 Text Mining
	BSc COMP16321 Introduction to Programming 1

Funding and awards

2023	Google Research workshop sponsorship	USD $1,000
2022	EPSRC Grant *Equally Safe Online* (as Co-I)	£1,000,000
2022	Google Research workshop sponsorship	USD $2,000
2021	Scottish Informatics and Computer Science Alliance (SICSA) PECE award	£2,750
2018	University of Manchester Postgraduate Summer Research Showcase: Best Poster	

Publications

Frenda, S., **Abercrombie, G**., Basile, V., Pedrani, A., Panizzon, R., Cignarella, A.T., Marco, C. and Bernardi, D. 2024 Perspectivist approaches to natural language processing: a survey. *Language Resources & Evaluation*. Springer.

Poudhar, A., Konstas, I. and **Abercrombie, G**. 2024. A Strategy Labelled Dataset of Counterspeech. In *Proceedings of the 8th Workshop on Online Abuse and Harms (WOAH 2024)*. ACL.

Cercas Curry, A., **Abercrombie, G**., and Talat, Z. 2024. Subjective Isms? On the Danger of Conflating Hate and Offence in Abusive Language Detection. In *Proceedings of the 8th Workshop on Online Abuse and Harms (WOAH 2024)*. ACL.

Dinkar, T., **Abercrombie, G**. and Rieser, V. 2024. ReproHum #0927-03: DExpert Evaluation? Reproducing Human Judgements of the Fluency of Generated Text. In Proceedings of the Fourth Workshop on Human Evaluation of NLP Systems (HumEval) @ LREC-COLING 2024. ELRA and ICCL.

Sasidharan Nair, S., Dinkar, T. and **Abercrombie, G**. 2024. Exploring Reproducibility of Human-Labelled Data for Code-Mixed Sentiment Analysis. In Proceedings of the Fourth Workshop on Human Evaluation of NLP Systems (HumEval) @ LREC-COLING 2024. ELRA and ICCL.

Abercrombie, G., Vitsakis, N., Jiang, A. and Konstas, I. 2024. Revisiting Annotation of Online Gender-Based Violence. In Proceedings of the 3rd Workshop on Perspectivist Approaches to NLP (NLPerspectives) @ LREC-COLING. ELRA and ICCL.

Bonaldi, H., Chung, Y-L., **Abercrombie, G**. and Guerini, M. 2024. NLP for Counterspeech against Hate: A Survey and How-To Guide. In Findings of the Association for Computational Linguistics: NAACL. ACL.

Plaza-del-Arco, F., Cercas Curry, A., Cercas Curry, A., **Abercrombie, G**. and Hovy, D. 2024. Angry Men, Sad Women: Large Language Models Reflect Gendered Stereotypes in Emotion Attribution. In Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics. ACL.

Abercrombie, G., Cercas Curry, A., Dinkar, T., Rieser, R. and Talat, Z.. 2023. Mirages. On Anthropomorphism in Dialogue Systems. In Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing. ACL.

Abercrombie, G., Jiang, A., Gerrard-Abbott, P., Konstas, I. and Rieser, R. 2023. Resources for Automated Identification of Online Gender-Based Violence: A Systematic Review. In The 7th Workshop on Online Abuse and Harms (WOAH). ACL.

Belz, A., Thomson, C., Reiter, E., **Abercrombie, G**., Alonso-Moral, J.M., Arvan, M., Braggaar, A., Cieliebak, M., Clark, E., van Deemter, K., Dinkar, T., Dušek, O., Eger, S., Fang, Q., Gao, M., Gatt, A., Gkatzia, D., González-Corbelle, J., Hovy, D. et al. 2023. Missing Information, Unresponsive Authors, Experimental Flaws: The Impossibility of Assessing the Reproducibility of Previous Human Evaluations in NLP. In Proceedings of the Fourth Workshop on Insights from Negative Results in NLP, pages 1–10. ACL.

Leonardelli, E., **Abercrombie, G**., Almanea, D., Basile, V., Fornaciari, T., Plank, B., Rieser, V., Uma, A., and Poesio, M. 2023. SemEval-2023 Task 11: Learning with Disagreements (LeWiDi). In Proceedings of the 17th International Workshop on Semantic Evaluation (SemEval-2023), pages 2304–2318. ACL.

Abercrombie, G and Rieser, V. 2022. Risk-graded Safety for Handling Medical Queries in Conversational AI. In *Proceedings of the 2nd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics*. Association for Computational Linguistics.

Bergman, A. S., **Abercrombie, G**., Spruit, S., Hovy, D., Boureau., Dinan, E., Y-L and Rieser, V. 2022. Guiding the Release of Safer E2E Conversational AI through Value Sensitive Design. In Proceedings of the 23rd Annual Meeting of the Special Interest Group on Discourse and Dialogue. ACL.

Dinan, E., **Abercrombie, G**., Bergman, A. S., Spruit, S., Hovy, D., Boureau, Y-L and Rieser, V. 2022. SafetyKit: First Aid for Measuring Safety in Open-domain Conversational Systems. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics* (ACL).

Abercrombie, G. and Batista-Navarro, R. 2022. Policy-focused Stance Detection in Parliamentary Debate Speeches. *Northern European Journal of Language Technology*.

Cercas Curry, A, **Abercrombie, G.** and Rieser, V. 2021. ConvAbuse: Data, Analysis, and Benchmarks for Nuanced Abuse Detection in Conversational AI. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*.

Abercrombie, G., Cercas Curry, A, Pandya, M. and Rieser, V. 2021. Alexa, Google, Siri: What are Your Pronouns? Gender and Anthropomorphism in the Design and Perception of Conversational Assistants. In *Proceedings of the 3rd Workshop on Gender Bias in Natural Language Processing*.

Abercrombie, G. and Batista-Navarro, R. 2020. ParlVote: a corpus for sentiment analysis of political debates." In *Proceedings of the 12th Language Resources and Evaluation Conference*.

Abercrombie, G. and Batista-Navarro, R. 2020. Sentiment and position-taking analysis of parliamentary debates: A systematic literature review. *Journal of Computational Social Science*.

Abercrombie, G., Nanni, F., Batista-Navarro, R. T., & Ponzetto, S. P. 2019. Policy preference detection in parliamentary debate motions. In *Proceedings of the 23rd Conference on Computational Natural Language Learning (CoNLL)*.

Abercrombie, G. and Batista-Navarro, R. 2019. Semantic Change in the Language of UK Parliamentary Debates. In *Proceedings of the 1st International Workshop on Computational Approaches to Historical Language Change*.

Abercrombie, G. and Batista-Navarro, R. 2018. Identifying opinion-topics and polarity of parliamentary debate motions. In *Proceedings of the 9th workshop on computational approaches to subjectivity, sentiment and social media analysis* (pp. 280-285).

Abercrombie, G. and Batista-Navarro, R. 2018. 'Aye' or 'No'? Speech-level sentiment analysis of Hansard UK parliamentary debate transcripts. In *Proceedings of the 11th Language Resources and Evaluation Conference*.

Abercrombie, G. and Hovy, D. 2016. *Putting Sarcasm Detection into Context: The Effects of Class Imbalance and Manual Labelling on Supervised Machine Classification of Twitter Conversations*. ACLSRW.

Abercrombie, G. 2016. A rule-based shallow-transfer machine translation system for Scots and English. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*.

Press and media

Mirages of Humanity: The Problem with Conversational AI. *InterMedia* Vol 1 Issue 4. December 2023.

Sweet, J. Before a Bot Steals Your Job, It Will Steal Your Name. ***The Atlantic***. August 11 2023.

Lee. D. Conference report: Taking gender pay gap to task. ***The Scotsman***. October 6 2022.

Invited talks

2024   **Scottish Informatics & Computer Science Alliance PhD Conference Keynote**, Aberdeen
2024   **Alan Turing Institute**, London
2023   **School of Computing**, **Napier University**, Edinburgh
2023   **The National Robotarium**, Edinburgh
2023   **Sociotechnical AI Research**, **Google Deep Mind**, London
2023   **School of Mathematical & Computer Sciences**, **Heriot-Watt University**
2023   **Institute for Language, Cognition, and Communication**, **University of Edinburgh**
2021   **The Content-centered Computing Group**, **University of Turin**
2021   **Milan Natural Language Processing Group**, **Bocconi University**, Milan

2020   **House of Commons Library, United Kingdom Parliament**, London
2019   **Information Retrieval Group**, **University of Glasgow**
2019   **PEAK Ensemble**, **peak.ai**, Manchester
2018   **NLP Group, Data & Web Science Group**, **University of Mannheim**
2018   **Social Science Data Group (MZES)**, **University of Mannheim**

Academic service & activities

Grant award panels

UKRI EPSRC 2023

Conferences & workshops

Organising Committee member :
Perspectivist Approaches to Disagreement in NLP (NLPerspectives) workshop (at LREC-COLING) 2024
CounterSpeech for Online Abuse (CS4OA) at INLGxSIGDial 2023
Perspectivist Approaches to Disagreement in NLP (and Beyond) (NLPerspectives+) workshop (at ECAI) 2023
Perspectivist Approaches to Disagreement in NLP (NLPerspectives) workshop (at LREC) 2022
SIGDIAL Special Session on Safety for E2E Conversational AI 2021

Area Chair
Association for Computational Linguistics ARR Rolling Review 2023, 2024

Programme Committee member:
AAAI Conference on Artificial Intelligence 2024
ACM Conference on Fairness, Accountability, and Transparency (ACM FAccT) 2024
Association for Computational Linguistics (ACL) 2023
Conference on Computational Linguistics (COLING) 2022
Empirical Methods in Natural Language Processing (EMNLP) 2023
European Association for Computational Linguistics (EACL) 2023, 2024
European Conference on Artificial Intelligence 2024
Gender Bias in Natural Language Processing (GeBNLP) 2021, 2022, 2023
Language Resources and Evaluation Conference (LREC) 2022
Language Resources and Evaluation Conference - International Conference on Computational Linguistics (LREC-COLING) 2024
SEMEVAL Task 10 Explainable Sexism 2023
Special Interest Group on Dialogue and Discourse (SIGDIAL) 2022, 2023
Widening NLP (WiNLP) 2020, 2021, 2022, 2023
Workshop on Online Harms and Abuse (WOAH) 2023
Workshop on Safety in Conversational AI 2024

Ethics Committee member:
Association for Computational Linguistics ARR Rolling Review 2023, 2024
European Association for Computational Linguistics (EACL) 2023

Journal reviews
Artificial Intelligence

European Language Resources …
IEEE Transactions on Cognitive and Developmental Systems
Harvard Data Science Review
Journal of Cultural Analytics
Language Resources and Evaluation
Online Social Networks and Media
Political Research Exchange
Transactions on Affective Computing

PhD examination

Marco Antonio Stranisci, University of Turin. External examiner

Outreach and science communication

2024	Universidad Complutense de Madrid	*Conferencia sobre campañas de desinformación y promoción de discursos de odio*. Panel organised by the Cabinet of the Presidency of the Spanish government.
2022	The Scotsman Data Conference	*Using Data and AI to Shine a Light on Inequality and Bias*. Panel.
2022	Edinburgh Science Festival	*Datasphere* at the National Museum of Scotland. Activities for children introducing robotics, programming and design.
2021	West Lothian Libraries	*Understanding Online Abuse: an AI Challenge*. Invited talk as part of 'Lothian Lugs: Borrow a Researcher' Engaging Libraries series. Heriot-Watt University Societal Engagement Award winner.
2021	Edinburgh Science Festival	*Gender Bias in AI: the Case of Conversational Assistants* Producer, Editor, Presenter.
2019	Blue Dot Festival, Cheshire	Instructor for the University of Manchester Department of Computer Science stall.
2017	TuringLab, Manchester	Instructor delivering programming classes to children.