Alva Markelius - Curriculum Vitae

ajkm4@cam.ac.uk | linkedin.com/in/alvamarkelius | alvamarkelius.github.io

Highly driven and ambitious Cambridge University Computer Science PhD researcher, Top 100 Brilliant Women in AI Ethics™ 2024 and Chair Centre for Human Inspired AI Early Career Committee. Deeply fascinated by ethical considerations, and design of AI, specifically researching social robotics and their intersection with large language models.

RESEARCH ACTIVITIES:

University of Cambridge, King's College **Cambridge, UK (2024-2028)**
PhD Researcher, Cambridge Trust Scholar
- Department of Computer Science & Technology, Affective Intelligence and Robotics Laboratory. Supervisor: Professor Hatice Gunes.

Brain+/DICE Lab, University of Gothenburg **Gothenburg, Sweden (2022 - 2024)**
Research Engineer
- Researching the use of LLMs in personalised and embodied contexts, with a main focus on ethical use and output control. Specific application development with LLM supervisor for driver monitoring system and biometric data.
- Part of STINT initiation grant funded project with Koç University developing and researching humanoid social robotics, human-robot interaction, affective computing and digitalised gamified cognitive interventions.

EU Joint Programme – Neurodegenerative Disease Research Network **Remote/Gothenburg (2022 - 2024)**
Network Organiser and Researcher
- Forte-funded project which includes a research network aimed at exploring the possibilities of using digital technology (e.g. robotics, VR) and for implementation of cognitive training including differential outcomes training.

Leverhulme Centre for the Future of Intelligence **Cambridge, UK (2022 - 2023)**
Research Collaborator
- Research assistant in the DARPA-funded Kinds of Intelligence project RECOG-AI specifically aimed at a robust conceptual and methodological framework for benchmarking evaluation of cognitive capabilities and generality in AI.
- Specifically worked with task development for training AI agents in a novel and unique cognitive approach for learning affordances and object understanding.

Human Factors Psychology Lab, Seoul National University **Seoul, S. Korea (2021 - 2022)**
Robotics Research Assistant
- Collaborated with an international and interdisciplinary cohort and assisted research in affective human/robot interaction, applied cognitive psychology, UI/UX, social robotics and VR.
- Active member of academic skill workshop. Demonstrated usage of Furhat robot design social interface and acquired field-relevant academic presentation and discussion skills.

ETIS Laboratory CY Cergy Paris Université **Paris, France (2022)**
Research Assistant
- Recipient of the EUTOPIA EURSS research scholarship to visit and conduct research at human/robot interaction lab.
- Conceptualised and implemented research project *A Human-Robot Mutual Learning System with Affect-Grounded Language Acquisition and Differential Outcomes Training* which was finalist for best paper award at ICSR 2023.

EDUCATION:

University of Cambridge, Lucy Cavendish College **Cambridge, UK (2022 - 2024)**
AI Ethics & Society Master of Studies
- First class degree.
- Part time master studies with a main interest in socio-technical systems, social robotics, cognitive capabilities in AI, emotion recognition AI, technology policy, intercultural AI narratives and human rights impacted by AI.

Seoul National University **Seoul, S. Korea (2021 - 2022)**
Exchange year (Artificial Intelligence)
- GPA 4.12 (Equivalence of First Class Honours classification).
- Relevant coursework: Explainable AI and Data Science, Affective Science, Psychology and Human-AI Interaction, Interdisciplinary approaches to AI, AI and Labour Economics and Cognitive Neuroscience.

University of Gothenburg **Gothenburg, Sweden (2019 - 2022)**
Cognitive Science Bachelor Degree
- GPA 3.8 (Equivalence of First Class Honours classification).

- Achieved highest grade and highest overall exam mark of all students on two AI courses.
- Board member Cognitive Science Student Association.
- Conducted thesis as part of international research project with Koç University in Türkyie; *Differential Outcomes Training of Visuospatial Memory: A Gamified Approach Using a Socially Assistive Robot*, published in International Journal of Social Robotics.

WORK EXPERIENCE:

EthicAI Advisory Ltd. **London, UK (2023 - present)**
Founder & Consultant
- Founding team member, senior consultant and lead product developer at AI ethics consultancy startup.

University of Gothenburg **Gothenburg, Sweden (2022 - 2023)**
Teacher assistant and lecturer: AI/Machine Learning/Affective Science
- Taught and lectured on e.g. explainable AI at the course Machine Learning and Data Analysis. Practical assistance with programming and marking and computer science related parts of the cognitive science course.
- Teaching assistant and responsible for organising and carrying out inter-university course Mechanism of Emotion, focusing on theory of emotion and affective computing methodology with University of Warwick.

PROFESSIONAL ACTIVITIES

Centre for Human Inspired AI (CHIA) Early Career Community **Cambridge, UK (2023 - present)**
Community Committee Chair
- Responsible for organising research events for a wide community of Cambridge postgraduate students in AI.

Cambridge University AI Ethics Society **Cambridge, UK (2022 - 2024)**
President (previously Secretary and Events Coordinator)
- President of a student run society aimed at spreading awareness on the social and political effects of technology.

AI Ethics & Society Cambridge Delegation to Nairobi **Nairobi, Kenya (2023)**
Delegate
- Part of delegation of Cambridge AI Ethics students to learn about the promises and potential pitfalls of AI in Kenya and East Africa by engaging with i.e. Kenyan government, Microsoft, IBM, British Commission and Lawyer´s Hub.

Cambridge AI Safety Hub Fellowship **Cambridge, UK (2022)**
Fellowship Participant
- Fellow in technology focused reading/discussion group on e.g. interpretability, AI capabilities and generality.

TEACHING EXPERIENCE

University of Cambridge - Machine Learning and Real World Data, Interaction Design (first year undergraduate Computer Science). Generative AI In Business (Cambridge University Press and Assessment professional course).
University of Gothenburg - Machine Learning and Data Analysis, Explainable AI (second to third year undergraduate Cognitive Science, Systems and Communication)
EUTOPIA - Mechanism of Emotion, Affective Computing (undergraduate Psychology and Cognitive Science)

AWARDS

King's College Cambridge Ferris Travel Award - Awarded for travel to Paris Conference on AI and Digital Ethics
Cambridge Trust International Scholar - Fully funded PhD scholarship for outstanding academic achievements.
100 Brilliant Women in AI Ethics™ - Women in AI Ethics annual list of rising stars and pioneers in the AI Ethics space.
Nova 111 Student List 2023 - one of the 111 young people with most potential in Sweden (Mathematics, Data, and Analytics)
Career Stipend Sweden 2022 - prestigious stipend awarded to promising early career students in Sweden.
Lucy Cavendish College Travel Award (x2) - granted for research related travels to Nairobi, Kenya and Pasadena, USA.

SELECTED INVITED TALKS

International Conf. Affective Computing & Intelligent Interaction – *Workshop keynote/panellist* **Glasgow (2024)**
KTH Royal Institute of Technology – *Lecture on AI Ethics for energy engineering students* **Stockholm (2024)**
Raoul Wallenberg Academy Conference – *Lecture on AI Ethics, human rights and democracy* **Stockholm (2024)**
Cambridge Centre for Human Inspired AI - *Social Robotics Research* **Cambridge (2023)**
Microsoft Share & Evolve Sweden – *AI Ethics and Responsibility* **Stockholm (2023)**
Oxbridge Robotics Summer School – *Robotics and Ethics* **Cambridge (2023)**
EQUALS EU – *Whose Digital Future? Gender, Rights, Inclusion, and Exclusion* **Geneva (2023)**
Nordic China Startup Forum – *AI Trends, Business and Female Leadership* **Online (2022)**
Science Festival Gothenburg – *AI, Robotics and Science Fiction* **Gothenburg (2020)**
Climate Parliament, Stockholm University – *AI, Climate Change and Science Fiction* **Stockholm (2018)**

MISCELLANEOUS

Software abilities: Computer programming (Python, R, Matlab, HTML, LaTex), machine learning and explainable AI.
Languages: English (Fluent, IELTS 8.5), Swedish (Native), German (Fluent), Korean (Beginner).
Societies: President Cambridge University Polo Club, Member of Cambridge University Scandinavian Society, Chair Centre for Human Inspired Technology Early Career Committee.

PUBLICATIONS

Markelius, A. & Gunes, H. (2025). Social Robotics and Large Language Models for Disability: A Scoping Review. *PREPRINT available at Research Square*. https://doi.org/10.21203/rs.3.rs-7260703/v1

Laban, G., Hough, J., Lee, M., **Markelius, A**., Foster, M. E., Stuart-Smith, J., & Ahmad, M. I. (2025). Bias and Fairness in Conversational User Interfaces. In *Proceedings of the 7th ACM Conference on Conversational User Interfaces* (pp. 1-8). https://doi.org/10.1145/3719160.3728629

Ravandi, B. S., Khan, I., **Markelius, A.**, Bergström, M., Gander, P., Erzin, E., & Lowe, R. (2025). Exploring task and social engagement in companion social robots: a comparative analysis of feedback types. *Advanced Robotics*, 1-16. https://doi.org/10.1080/01691864.2025.2526668

Markelius, A., Lou, Y., Galazka, M., Lundgren, S., Zemblys, R., Lind, H., & Lowe, R. (2025). Investigating the Mitigation of Stress in Autonomous and Non-autonomous Vehicles Using LLM Feedback. In *International Conference on Human-Computer Interaction* (pp. 108-127). Cham: Springer Nature Switzerland.

Markelius, A., Bailey, J., Gibson, J. L., & Gunes, H. (2025). Stakeholder Perspectives on Whether and How Social Robots Can Support Mediation and Advocacy for Higher Education Students with Disabilities. *arXiv preprint arXiv:2503.16499.* https://arxiv.org/abs/2503.16499

Rutar, D., **Markelius, A.,** Cheke, L. G., Hernández-Orallo, J. (2025). Cognitive Science-Inspired Evaluation of Core Capabilities for Object Understanding in AI. https://doi.org/10.48550/arXiv.2503.21668

Markelius, A. (2025). An Empirical Design Justice Approach to Identifying Ethical Considerations in the Intersection of Large Language Models and Social Robotics. in Philipp Hacker (ed.), *Oxford Intersections: AI in Society* (Oxford, online edn, Oxford Academic, 20 Mar. 2025), https://doi.org/10.1093/9780198945215.003.0013

Rutar, D., **Markelius, A**., Schellaert, W., Hernández-Orallo, J., & Cheke, L. (2025). General interaction battery: Simple object navigation and affordances (GIBSONA). *Cognitive Systems Research*, 101411. https://doi.org/10.1016/j.cogsys.2025.101411

Zokaei, N., Kjeldsen, P., Havild, N., Dong, M., **Markelius, A**., Bergstrom, M., … Blicher, J. (2024). ACTnow: Computerised cognitive training for Mild Cognitive Impairment. osf.io/preprints/psyarxiv/v3g84

Vivas, A. B., Estévez, A. F., Khan, I., Roldán-Tapia, L., **Markelius, A.**, Nielsen, S., & Lowe, R. (2024). DigiDOP: A framework for applying digital technology to the Differential Outcomes Procedure (DOP) for cognitive interventions in persons with neurocognitive disorders. *Neuroscience & Biobehavioral Reviews, 105838.* https://doi.org/10.1016/j.neubiorev.2024.105838

Markelius, A., Wright, C., Kuiper, J., Delille, N., Kuo, Y.T. (2024) The Mechanisms of AI Hype and its Planetary and Social Costs, *Topical Collection on The Ethical Implications of AI Hype: Examining the overinflation and misrepresentation of AI capabilities and performance, Springer Nature AI & Ethics:* https://doi.org/10.1007/s43681-024-00461-2

Markelius, A., Sjöberg, S., Bergström, M., Ravandi, B.S., Vivas, A.B., Khan, I., Lowe, R.: (2023), Differential Outcomes Training of Visuospatial Memory: A Gamified Approach Using a Socially Assistive Robot. *International Journal of Social Robotics.* https://doi.org/10.1007/s12369-023-01083-0

Markelius, A., Sjöberg, S., Lemhauori, Z., Cohen, L., Bergström, M., Lowe, R., Cañamero, L., (2023), A Human-Robot Mutual Learning System with Affect-Grounded Language Acquisition and Differential Outcomes Training. *International Conference on Social Robotics 2023.* ***Finalist for Best Paper Award.*** https://doi.org/10.1007/978-981-99-8718-4_10

Chan, A., Salganik, R., **Markelius, A.**, Pang, C., Rajkumar, N., Krasheninnikov, D., ... & Maharaj, T. (2023). Harms from Increasingly Agentic Algorithmic Systems. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency* (pp. 651-666). https://doi.org/10.1145/3593013.3594033

CONFERENCE ACTIVITIES

2025 Conference on Robot Learning CoRL **Seoul, Korea (2025)**
Workshop Main Organiser
- Collaboration with Google DeepMind, lead organiser for workshop titled Robot Learning Done Right: Responsibly Developing Foundation Models for Robotics

ACM International Conference on Conversational User Interfaces **Waterloo, Canada (2025)**
Workshop Organiser
- Bias and Fairness in Conversational User Interfaces

IEEE International Conference on Robot and Human Interactive Communication **Eindhoven, NL (2025)**
Associate Editor

IEEE International Conference on Robot and Human Interactive Communication **Pasadena, US (2024)**
Workshop Main Organiser
- Large Language Models in the RoMan Age: Exploring Implications on Design, Ethics and Social Impact

International Conference on Social Robotics **Doha, Qatar (2023)**
Workshop Co-Organiser and Co-Chair, Robot Competition Judge
- Co-Researching with Humans-in-the-Loop: Using participatory methods, research and co-design in HRI. Nomination for Best Paper Award

IEEE International Conference on Robot and Human Interactive Communication **Busan, S. Korea (2023)**
Special Session Organiser, Chair and Reviewer
- SARCHA: Socially-Assistive Robots in Clinical and Healthcare Applications.