# Ninell Oldenburg



## Education

2024 - current PhD Philosophy of AI, University of Copenhagen

Supervisor: Anders Søgaard

2021 - 2023 MSc IT & Cognition, University of Copenhagen & Hong Kong University of Science and Technology

Focus: Machine Learning, AI Ethics

2017 - 2020 BSc Computational Linguistics, University of Potsdam & University of Trento

Focus: Theory of Computer Science, Formal Semantics

### Academic Experience

#### Research Fellow, PIBBSS Summer Research Fellowship, Oxford, UK/Prague, CZ

Jun 2023 - Sep 2023

Modeled an environment for learning and sustaining normative systems in multi-agent Markov games via Bayesian rule induction. *Mentor: Tan Zhi-Xuan (MIT)*.

Visiting Graduate Research Assistant, MIT, Cambridge, US / Remote

Feb 2023 - Jun 2023

Bayesian multi-agent modelling for symbolic norm learning.

Supervisors: Tan Zhi-Xuan (MIT), Dr. Nora Hollenstein (University of Copenhagen).

### Graduate Research Assistant, IT University Copenhagen, DK

Oct 2021 - Dec 2022

Meta-study on the state of the art of bias research in NLP.

Supervisor: Prof. Barbara Plank (LMU Munich/IT University Copenhagen).

#### Undergraduate Research Assistant, University of Dramatic Art, Berlin, DE

Sep 2019 - Feb 2020

End-to-end system for computer-aided pronunciation training.

Supervisors: Prof. Andree Gubisch (University of Dramatic Art), Prof. David Schlangen (University of Potsdam).

### Software Engineering Intern, German Research Center for AI, Berlin, DE

May - Aug 2017

Multi-modal scheduling app. Supervisor: Dr. Kinga Schumacher (DFKI).

### **Publications**

**Ninell Oldenburg** and Gleb Papyshev. 2026. The Stories We Govern By: Al, Risk, and the Power of Imaginaries. (Under review.)

**Ninell Oldenburg**, Ruchira Dhar and Anders Søgaard. 2025. *Two Conceptions of Intelligence: Why AI Researchers Disagree About Scaling Laws and Future Risk*. (Under review.)

Iwan Williams, **Ninell Oldenburg**, Ruchira Dhar, Joshua Hatherley, Constanza Fierro, Nina Rajcic, Sandrine R. Schiller, Filippos Stamatiou, and Anders Søgaard. 2025. *Mechanistic Interpretability Needs Philosophy*. (Under review).

Ruchira Dhar, Ninell Oldenburg, and Anders Søgaard. 2025. Measure Generality, Not Intelligence. (Under review).

Ruchira Dhar, **Ninell Oldenburg**, and Stephanie Brandl. 2025. *Beyond Technocratic XAI: The Who, What & How in Explanation Design*. (Under review).

Ninell Oldenburg and Anders Søgaard. 2025. Navigating the Informativeness-Compression Trade-Off in XAI. AI & Ethics.

**Ninell Oldenburg** and Tan Zhi-Xuan. 2024. *Learning and Sustaining Shared Normative Systems via Bayesian Rule Induction in Markov Games*. In Proceedings of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS).

### Honours, Awards, Fellowships, & Summer Schools

- 2025 German Academic Scholarship Foundation (German: Studienstiftung des deutschen Volkes): \$28,000
- 2025 Machine Intelligence and Normative Theory Lab Affiliate, Australian National University
- 2025 Al & Humanity Lab Research Visit, Hong Kong University
- 2024 Philosophy of Al Summer School, University of London, Paris
- 2024 ABC Summer School "Neuro-Al: Understanding computation in mind and brain", University of Amsterdam
- 2023 Training for Good EU Tech Policy Fellowship: \$9,000 (declined)
- 2022 FTX Future Fund Grant for an Al Reading Group Berkeley, CA & Al Self-Study: \$17,500
- 2019 31st European Summer School for Language, Logic, and Information (ESSLLI): Travel Grant
- 2018 Erasmus+ Study Grant University of Trento: \$3,600

### **Talks**

Do Scaling Laws Suggest Universal Intelligence? Examining Philosophical Assumptions of AI Evaluations.. Artificial Intelligence and New Human-Technology Relations as a Challenge for the Philosophy of Science Workshop. University of Münster. 2025.

Philosophical Commitments to LLM Evaluations: The Problem of Moving Goalposts and Observational Relativity. Al & Humanity Lab Hong Kong, Seminar Series 2025.

The Completeness Problem: Beyond Human Metrics in Assessing Cognitive Systems. 4th Advanced Course on Artificial Intelligence & Neuroscience 2024.

Navigating the Informativeness-Compression Trade-Off in Explainability. International Association for Computing and Philosophy (IACAP) 2024 Conference.

# Teaching/Mentoring

#### Research Mentor at Lumiere Education

May 2025 - ongoing

Mentoring high school students to get their first research experiences.

### Lecturer "Datamining", University of Copenhagen, DK

Apr - May 2025

Master's Level Language Processing and AI Ethics

Lecturer "Empirical Methods and Theory of Science", University of Copenhagen, DK

Oct 2024 - Nov 2024

Bachelor's Level Statistics

TA "Introduction to Computer Science", Hong Kong University of Science and Technology, HK Sep 2022 - Dec 2022 Bachelor's Level C++

### Media

Is there someone behind the screen? Researchers are divided on AI consciousness. ScienceNordic (Danish Version on Videnskab.dk). 2025.

# Non-Academic Experience

#### Email Marketing Student Assistant, SOUNDBOKS, Copenhagen, DK

Oct 2021 - Jun 2022

Hold the end-to-end ownership of the email marketing strategy and all email needs (planning, creation, tracking, and data analysis); guided the English and German copywriting and localization.

#### Associate Linguist, Google Germany (via Adecco), Hamburg, DE

Aug 2020 - Aug 2021

Natural language generation & understanding for the Google Assistant.

### Integration Mentor, Huckepack e.V., Berlin, DE

Sep 2016 - Sep 2018

Mental health and integration mentor for a refugee teenager.

### Miscellaneous

**Programming** Python (incl. NLTK, SciPy), C++, Lua, JavaScript, Java

Languages German (native), English (fluent), Italian, French, Danish (intermediate)

Journalism Writer & Editor (2016 - 2020) Kiez und Kneipe, Berlin

Youth Journalist (2010 - 2014) Märkische Oderzeitung, Frankfurt (Oder)

Blogging 12 Months at 12 Places (2020 - 2021) Zwoelfmaldeutschland, Germany

Books (2015 - 2017) Yet Another Bookreview

Last updated: June 19, 2025