

Elisa Nguyen

PhD student in Trustworthy Machine Learning

✉ elisa.nguyen@live.de | 🏠 <http://elisanguyen.github.io> | 🌐 ElisaNguyen

Education

International Max Planck Research School for Intelligent Systems

Tübingen, DE

PhD student in STAI group led by Prof. Dr. Seong Joon Oh

May 2022 - present

- Thesis (tentative): Towards human-centered data attribution
- Advisor: Prof. Dr. Seong Joon Oh

University of Twente

Enschede, NL

MSc in Interaction Technology, GPA: 8.6 (cum laude)

Sep. 2019 - Dec. 2021

- Thesis: Temporal Spike Attribution - A local feature-based explanation for spiking neural networks
- Thesis advisors: Prof. Dr. Christin Seifert, Dr. Gwenn Englebienne and Dr. Meike Nauta

Nordakademie

Elmshorn, DE

BSc in Computer Science and Management, GPA: 1.8

Oct. 2013 - Mar. 2017

- Thesis in collaboration with Airbus: Requirements development for a PLM system for specific processes in an Industry 4.0 context
- Thesis advisors: Prof. Hinrich Schröder and Alcibiades Gonzalez, M.Sc. (Airbus).
- Erasmus Semester at ebs Paris with courses in international management and marketing.

Professional Experience

Microsoft Research

Redmond, WA, USA

Research internship on data provenance

Jun 2025 - Sep 2025

Interning in MSR's Deep Learning group under the supervision of Dr. Andrzej Bankurski-Fahey and Jaron Lanier.

Vector Institute

Toronto, CA

Research internship on influence functions

Jan 2025 - May 2025

Visiting Prof. Dr. Roger Grosse's group in the frame of Vector's research internship program.

University of Twente

Enschede, NL

Literature survey on evaluation methods in explainable AI

May 2021 - Dec 2021

Student research assistant in a team across the University of Twente and the University of Duisburg-Essen led by Meike Nauta.

Airbus

Hamburg, DE

IT Project Manager in Technical Data Management department

Mar 2017 - Sep 2019

Scrum Master for legacy retirement of CAD Software, Project support (Risk, Budget and Stakeholder Management) in PLM projects, Technical and financial management of bundle contracts with suppliers.

Teaching experience

Supervision:

Philipp Davydov

Master Thesis, Apr. 2025 - Dec. 2025, University of Tübingen

Anastasia Alekseeva

Master Thesis, Mar. 2024 - Sep. 2024, University of Tübingen

Albert Catalan

Research project and Master Thesis, Nov. 2023 - Mar. 2025, University of Tübingen

Evgenii Kortukov

Master Thesis, Jun. 2023 - Apr. 2024, University of Tübingen

Teaching:

Tutor Trustworthy ML

WS 2022/2023, University of Tübingen

TA Data Science

SS 2021, University of Twente

TA NLP

WS 2020/2021, University of Twente

Tutor Discrete Mathematics 2

Q3 2016, Nordakademie

Tutor Discrete Mathematics 1

Q1 2015, Nordakademie

Publications

* Equal contribution

Conference and Journal papers:

[1] Wang, A., **Nguyen, E.**, Yang, R., Bae, J., McIlraith, S. & Grosse, R. (2025). Better Training Data Attribution via Better Inverse Hessian-Vector Products. arXiv preprint arXiv:2507.14740. *Accepted at NeurIPS 2025*.

[2] Kortukov, E., Rubinstein, A., **Nguyen, E.**, Oh, S. J. (2024). Studying Large Language Model Behaviors Under Realistic Knowledge Conflicts. arXiv preprint arXiv: 2404.16032. *Accepted at COLM 2024*.

[3] **Nguyen, E.**, Seo, M., & Oh, S. J. (2024). A Bayesian approach to analysing training data attribution in deep learning. *Advances in Neural Information Processing Systems*, 36.

[4] **Nguyen, E.**, Nauta, M., Englebienne, G., & Seifert, C. (2023). Feature Attribution Explanations for Spiking Neural Networks. In 2023 IEEE 5th International Conference on Cognitive Machine Intelligence (CogMI) (pp. 59-68). IEEE.

[5] Nauta, M., Trienes, J., Pathak, S., **Nguyen, E.**, Peters, M., Schmitt, Y., ... & Seifert, C. (2022). From anecdotal evidence to quantitative evaluation methods: A systematic review on evaluating explainable ai. *ACM Computing Surveys*.

[6] **Nguyen, E.***, Theodorakopoulos, D.*, Pathak, S., Geerdink, J., Vijlbrief, O., Van Keulen, M., & Seifert, C. (2020, October). A hybrid text classification and language generation model for automated summarization of dutch breast cancer radiology reports. In 2020 IEEE second international conference on cognitive machine intelligence (CogMI) (pp. 72-81). IEEE.

[7] Geissler, D.*, **Nguyen, E.***, Theodorakopoulos, D.*, & Gatti, L. (2020). Pokérator-Unveil your inner Pokémon. In ICCV (pp. 500-503).

Workshop papers:

[1] Davydov, P., Prabhu, A., Bethge, M., **Nguyen, E.***, & Oh, S. J.* (2025). LLM generation novelty through the lens of semantic similarity. arXiv preprint arXiv:2510.27313. *Accepted at ICML FoGen workshop 2026*.

[2] Davies, A., **Nguyen, E.**, Simeone, M., Johnston, E., & Gubri, M. (2024). Social Science Is Necessary for Operationalizing Socially Responsible Foundation Models. arXiv preprint arXiv:2412.16355. *Accepted at ICLR HAIC workshop 2025*.

[3] **Nguyen, E.**, Kortukov, E., Song, J. & Oh, S. J. (2023). Exploring Practitioner Perspectives On Training Data Attribution Explanations. arXiv preprint arXiv:2310.20477. *Accepted at NeurIPS XAI in Action workshop 2023*.

Textbook:

[1] Mucsányi, B., Kirchhof, M., **Nguyen, E.**, Rubinstein, A., Oh, S. J. (2023). Trustworthy Machine Learning. <https://trustworthyml.io/>.

Preprints:

[1] **Nguyen, E.**, Bertram, J., Kortukov, E., Song, J. & Oh, S. J. (2024). Towards User-Focused Research in Training Data Attribution for Human-Centered Explainable AI. arXiv preprint arXiv:2409.16978.

Talks

TDA: Promises, Challenges and Human Needs	Invited Talk at Guide Labs, May 2026
TDA as Explanations: Insights from User Studies	Data Foundations of AI Seminar Series, May 2026
An Introduction to TDA	Security Seminar, UC Berkeley, Dec 2025
Towards User-Focused Research in TDA for HCXAI	IMPRS-IS Think and Link, MPI-IS, Oct 2024
Towards Human-Centred Data Attribution	PUI Group Seminar, University of Stuttgart, Jun 2024

Academic Service

Reviewing	NeurIPS '24, '25 (Top Reviewer), '26; CHI '25; Workshops at ICCV, ECCV, CVPR, NeurIPS
Workshops	Co-organising ATTRIB@NeurIPS 2026, ATTRIB@NeurIPS 2024, Volunteering at WiML@ICML 2024
Representation	PhD representative in the Tübingen AI Center Council August 2024 - May 2026

Skills

Coding	Python, Django, Version control with Git
Organization	Classical project management, Agile project management (Scrum, SAFe)
Languages	German (Mother tongue), English (C2), French (C1), Vietnamese (Mother tongue), Dutch (B2)

Outreach and Volunteering

KI macht Schule Leadership local group Tübingen (starting Mar. 2026), AI outreach for high schools (students and teachers)	<i>Tübingen, DE</i> Nov. 2022 - Present
Viva con Agua Member, FUNdraising for water, sanitation and hygiene projects in the global south	<i>Tübingen, DE</i> Sep. 2022 - Present