



## Johannes Bertram

Born on August 7th, 2001  
Böblingen – Germany

🏠 Traifelbergplatz 3, 70597 Stuttgart

✉️ [johannes.bertram@yale.edu](mailto:johannes.bertram@yale.edu)

🔗 <https://johannesbertram.github.io>

## Education

### Current M.Sc. Machine Learning

University of Tübingen  
Current Grade: 1.0

### 2025 Student Exchange

Yale University

### 2023 B.Sc. Cognitive Science

University of Tübingen  
Final Grade: 1.07

### 2019 Abitur

Wilhelms-Gymnasium Degerloch  
Final Grade: 1.0

### 2017 Student Exchange

Kaiapoi High School  
New Zealand

## Scientific interests

- Interpretability and Explainability
- AI Safety
- LLMs and Foundation Models
- Agentic AI
- Neuro AI

## Further interests

- Game of Go: EGF 2 Dan
- Long-distance Running

## Experience

### Jul. 2025 – to date Research Assistant

ELLIS Institute and Max Planck Institute for Intelligent Systems  
Safety- & Efficiency-Aligned Learning Group  
Principled AI Safety Benchmarking

### Mar. 2025 – to date Engineering Lead

FulcrumCare - Yale Startup  
Full-stack Development of Front Desk Automation Tools

### Aug. 2024 – Jan. 2025 Research Project in Visual Saliency

NeuroAI - Bethge Lab  
Extending the MIT/Tübingen Saliency Benchmark

### May. 2024 – Dec. 2024 Research Assistant

Scalable Trustworthy AI Group  
Training Data Attribution and Concept Attribution, Teaching  
Assistant for Trustworthy Machine Learning

### Sep. 2023 – to date Board Member and Webmaster

Go-Verband Baden-Württemberg e.V.

### Oct. 2021 – Aug. 2022, Oct. 2023 – Mar. 2024 Research Assistant

Neurocognitive Modeling Group  
Benchmarking, Social Inference, Tutor for Cognitive Modeling

### Apr. 2023 – Sep. 2023 Deep Learning and Robotics Internship

FANUC Deutschland GmbH  
Detection of external Influences applied to a Robot

### Sep. 2020 – Aug. 2021 Tutor for Mathematics of Computer Science

University of Tübingen

## Publications

Bertram, J., Dyballa, L., Keller, T. A., Kinger, S., Zucker, S. (2025).  
Manifolds And Modules: How Function Develops In a Neural  
Foundation Model. In Data on the Brain and Mind Workshop  
at the 39th Conference on Neural Information Processing  
Systems (NeurIPS 2025).

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

Scholz, F., Ayari, E., Bertram, J., & Butz, M. V. (2024). Quick and  
Accurate Affordance Learning. Proceedings of the Annual  
Meeting of the Cognitive Science Society, 46.

## Awards and Scholarships

### Jan. 2025 – May. 2025 Yale International Exchange Fellowship

### Apr. 2020 – Mar. 2026 Deutschlandstipendium

### Jul. 2019 1st Place HFT Mathe-Cup

Tübingen, Sep. 2025

*Johannes Bertram*