



# LLM Induced Embedding Spaces

LLM performance and the internal representations they create

---

**Amanda Myntti**  
M.Sc. in Mathematics

**Supervised by**  
Filip Ginter, Veronika Laippala, and Ion Petre

# Summary of Research

---



## LLM pretraining data selection

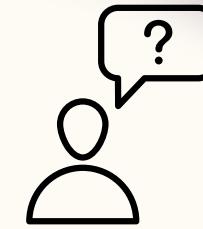
Effect of text qualities on  
LLM performance.

First publication: [here](#)



## Embedding space structure

Density and pathing in  
embedding spaces used  
to analyse models'  
hidden state and  
attention excitement.



## Hidden states of the model

Changes in hidden states  
with repeated training  
instances.

# Applications of my work

- Retrieval augmented generation
- LLM pretraining and finetuning
- Explainability methods

Contact me



[amanda.a.myntti@utu.fi](mailto:amanda.a.myntti@utu.fi)



[mmanteli.github.io](https://mmanteli.github.io)



[Amanda Myntti](https://www.linkedin.com/in/amanda-myntti)