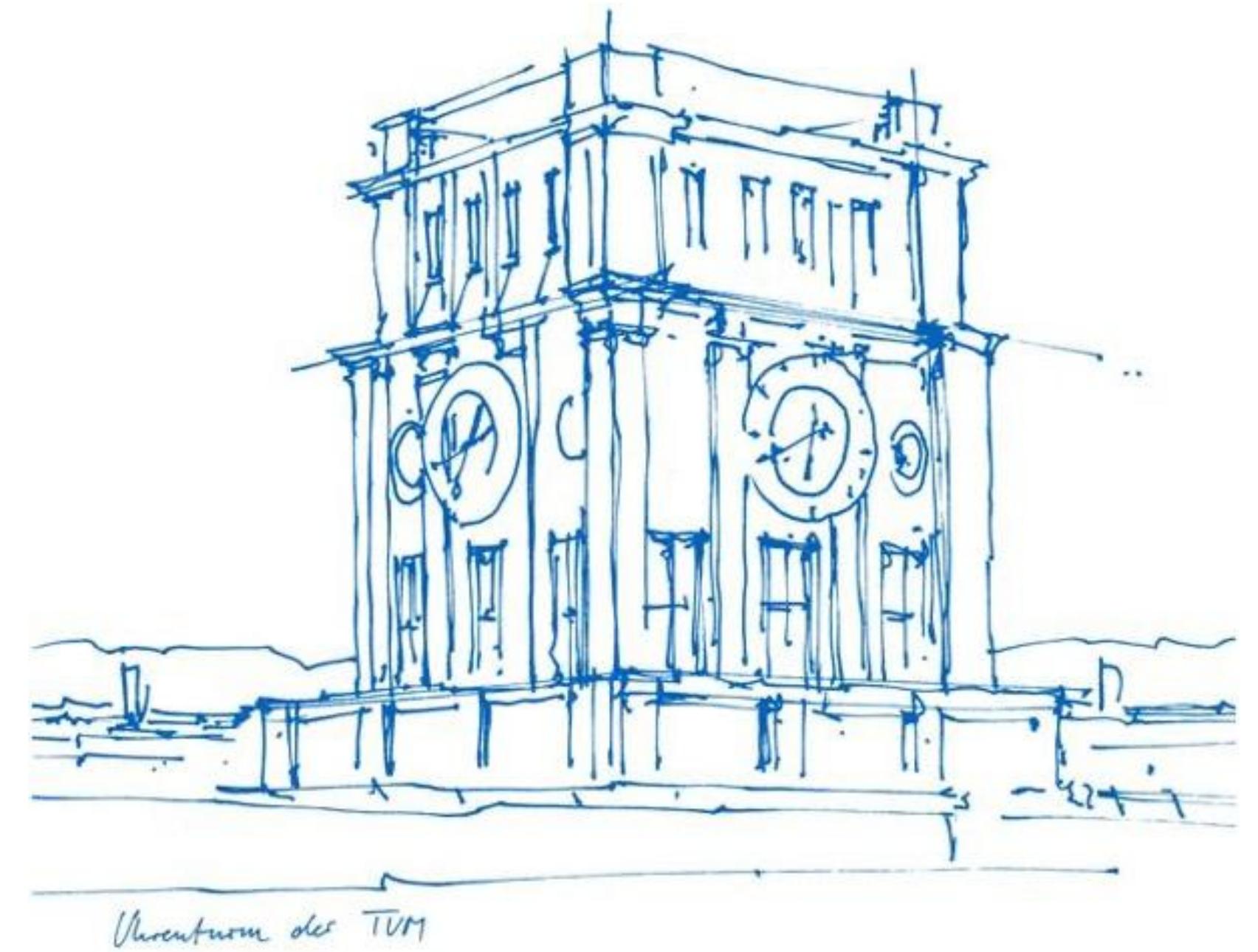


Master Seminar AI for Vision-Language Models in Medical Imaging (IN2107, IN45010)

Cosmin Bercea,
Jun Li,
Prof. Julia Schnabel



I32 – Chair for Computational Imaging and AI in Medicine - **CompAI**
Faculty of Informatics and Institute for Advanced Study

Outline

- Introduce ourselves
- Prerequisites
- Brief intro Vision-Language Models
- (Preliminary) semester outlook
- Deliverables
- Organisation and questions

L32: Computational Imaging and AI in Medicine



compai.io



Intro

Who we are



Cosmin I. Bercea

Postdoctoral Researcher

cosmin.bercea@tum.de



Julia A. Schnabel

Professor for Computational Imaging and
AI in Medicine, Director of the Institute of
Machine Learning in Biomedical Imaging



Jun Li

PhD Student

june.li@tum.de

Prerequisites



Who you are

- You are a **masters student in informatics** or a related program
- You have programming experience in Python and **PyTorch**
- You ideally took the **AI in Medicine I** lecture (IN2403) already (or related ML /DL courses)
- You want to start your academic career in ML and AI for Medicine
(A good grade in this seminar will help you find **IDPs, GRs, or MAs** at our chair)

Vision-Language Models

Why?

- Humans understand the world by vision and language.
- Doctors analyze both patient-data and medical scans
- Obtain additional information on the images



NIC : a bus that is **sitting** in the street .
ours: a **red and white** bus driving down a street .



NIC: a close up of a **toaster on a wall**.
ours: a close up of a pair of scissors .



NIC: a white plate topped with a cut in half sandwich.
ours: a white plate topped with a sandwich and **salad** .



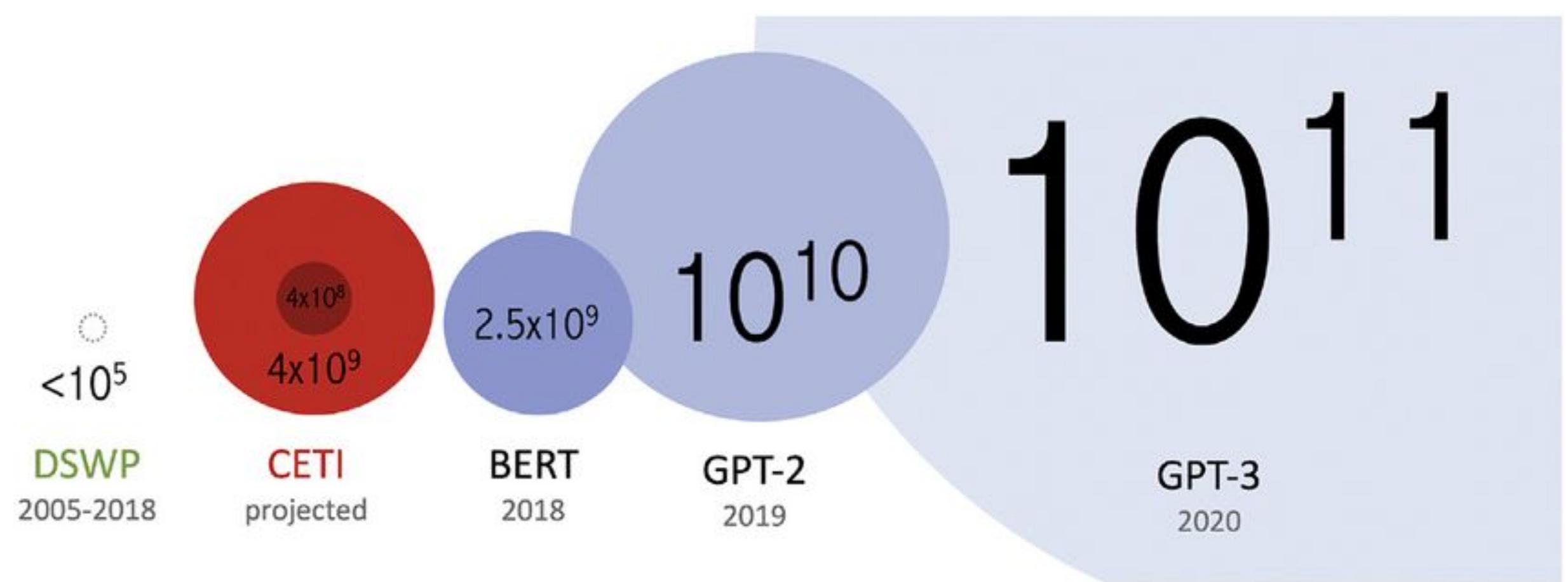
NIC: a group of people standing on top of a sandy beach.
ours: a group of people standing on a beach with **surfboards**.



NIC: a city street filled with lots of traffic.
ours: **a bus** driving down a street **next to a** traffic light.



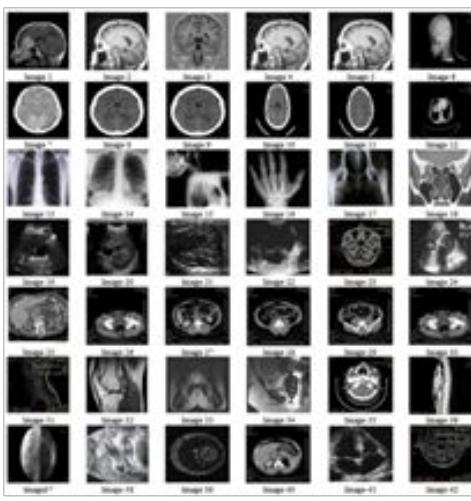
NIC: a **person** laying on a bed with a laptop.
ours: a dog laying on a bed in a bedroom.



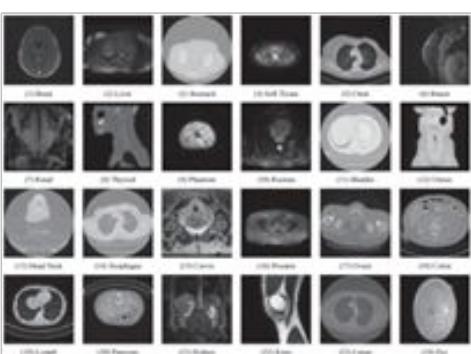
Vision-Language Models

How?

Large Dataset



Target

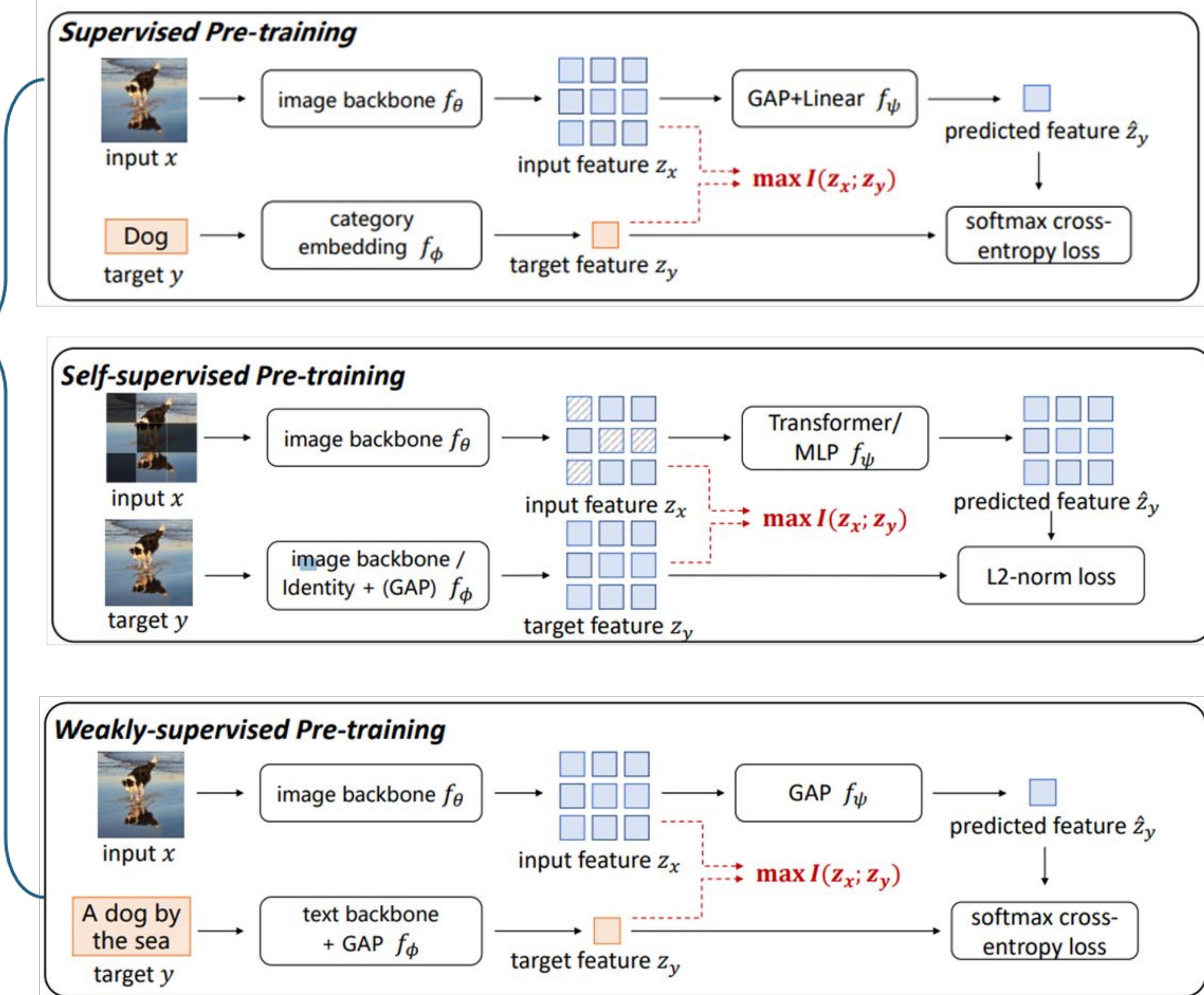


The paradigm of Transfer learning

Pretext
Task

Knowledge
Transfer

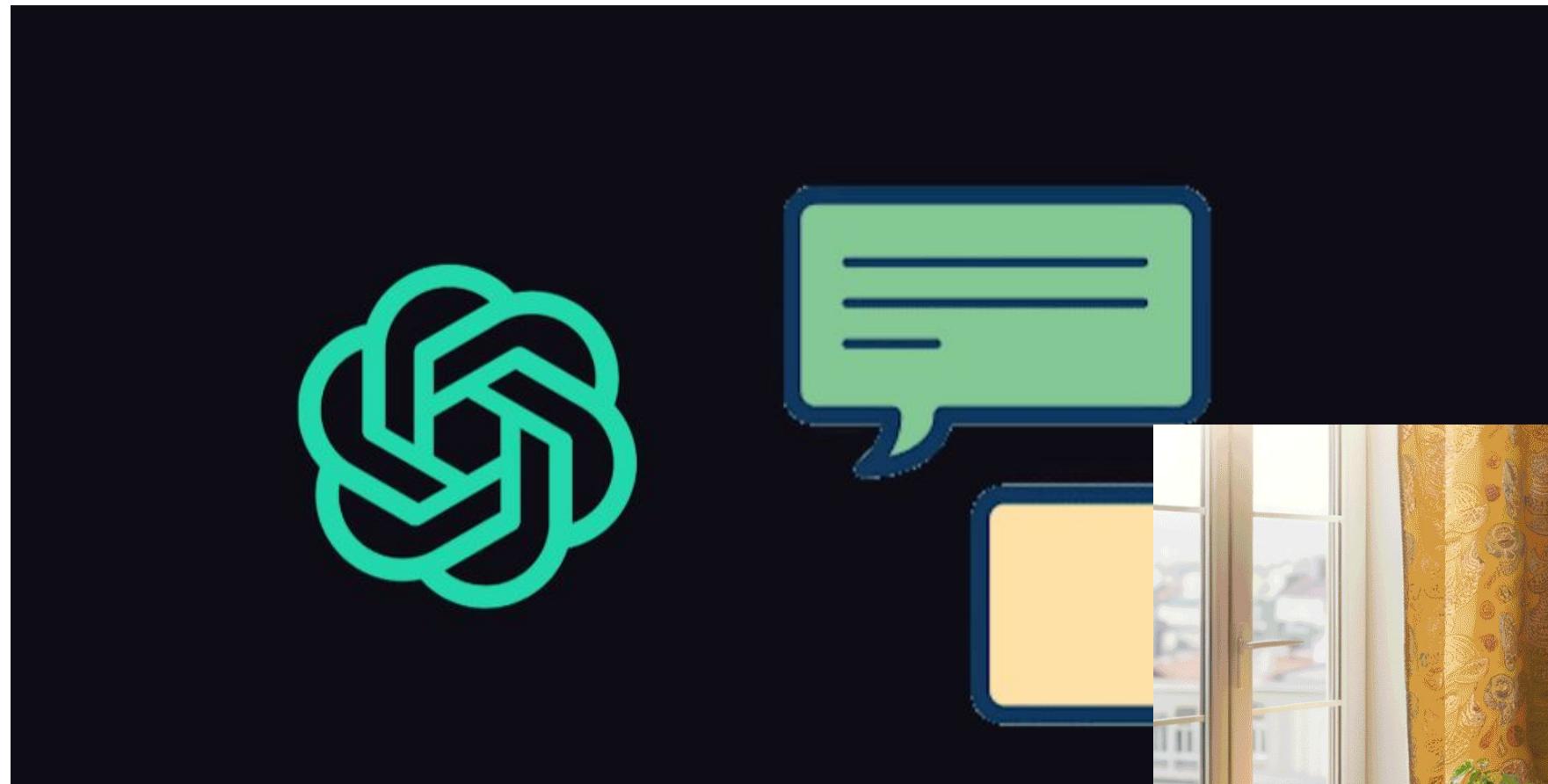
Target
Task



Vision-Language Models



Applications



GPT-4



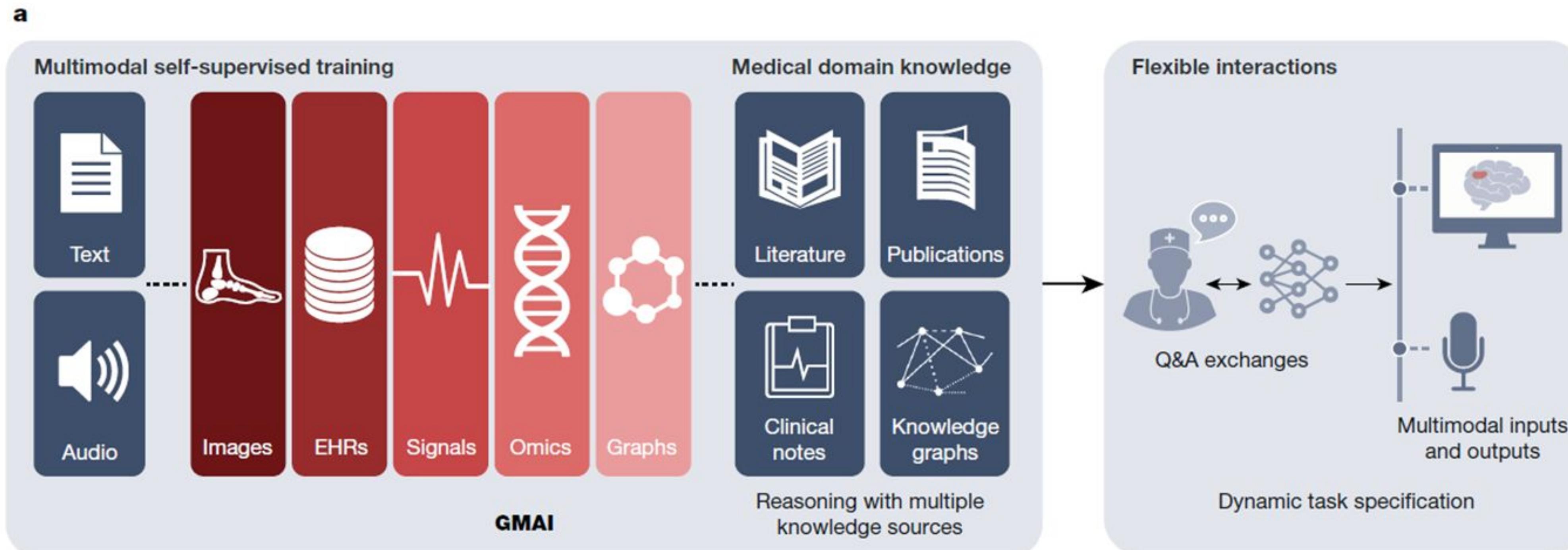
DALL•E 3



Sora

Vision-Language Models

Applications



Seminar structure (Preliminary)

Outlook

16.10.	Welcome and Introduction to VLMs
23.10.	VLMs part II
30.10.	How to read papers present make posters
06.11.	Guest talk: Prof. Weidi Xie (Oxford Shanghai Jiao Tong)
13.11.	Guest talk: Prof. Benedikt Wiestler (TUM KRI)
20.11.	Student Presentations
27.11.	Student Presentations
04.12.	Guest talk: Che Liu (Imperial College London, UK)
11.12.	Student Presentations
18.12.	Guest talk: Prof. Qi Dou (Chinese University of Hong Kong)
08.01.	Student Presentations
15.01.	Student Presentations
22.01.	Student Presentations
29.01.	Student Presentations
05.02.	Poster Session (All groups)



Deadline Paper Selection

Deadline Poster Submission

Deliverables

Your tasks - The academic cycle.



READ.

- Individual work
- Choose a topic & paper
- Present the paper
 - How the method works
 - Critical evaluation of the results
 - **15 min talk + 5 min Q/A**

CODE.

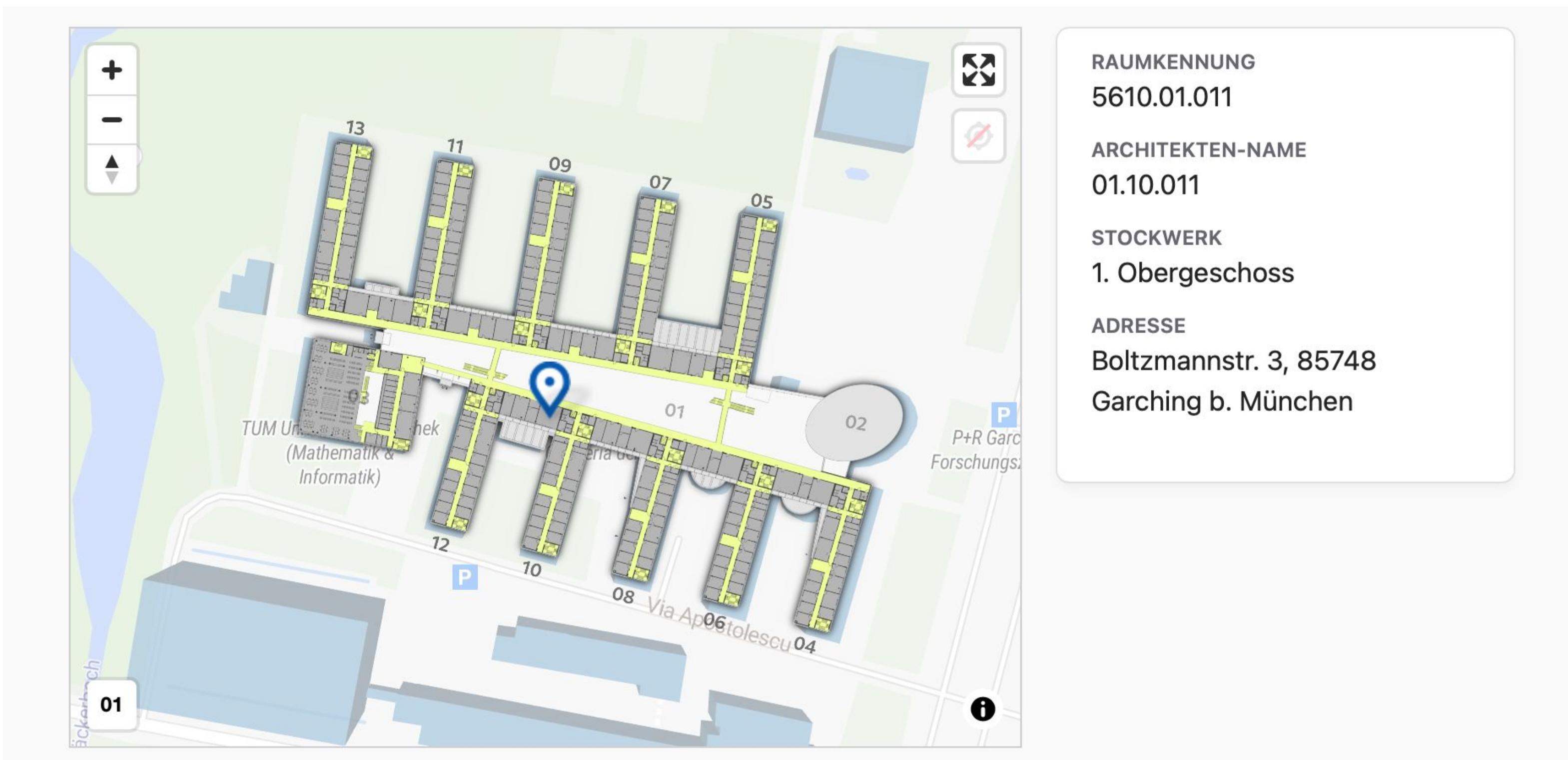
- Individual work
- Implement and test your selected method on our data
- Deliver: **Code**

PRESENT.

- Group work (2-3 students)
- Compare the strengths and weaknesses of the two methods
- Prepare **group poster**
- Individual **pitch: 3 min/person**

Where and When?

- Garching FMI - Seminar room [5610.01.011](#)
- Wednesday, 14:00-16:00



Throwback to last semesters



See you soon!



Cosmin I. Bercea
Postdoctoral Researcher
cosmin.bercea@tum.de



Jun Li
PhD Student
june.li@tum.de