

Lab session

À la carte

abc [BERT](#) Named-entity recognition (NER)

abc [nanoGPT](#) training from scratch on a dataset of your choice



[CLIP](#) embedding to cluster/index images from a dataset of your choice



[VisionEncoderDecoder](#) model:

- small pretrained [encoder](#) & random [decoder](#)
- Pretrain decoder
- Find and prepare dataset for OCR (streetview, [IAM](#)...) — ⚠️ tokenization
- finetune X-attn and rest using LoRa-like if needed
- viz attn (cross & self)
- Probably requires to resize (h) and pad (w) images for the encoder.