Travis Seng, Ph.D

Ph.D - Engineer in Vision - NLP - AI - Computer Science

@ seng.travis@gmail.com   https://travis-seng.fr   @travisseng   travisseng   Toulouse, France

Experience

Applied AI Engineer, **Yuzzit**, Paris   November 2024 – Present
- Integrating AI into company products, spanning computer vision, NLP, and LLM applications.
- Delivered features including smart video cropping, live TV banner detection, highlight extraction, live transcription, and semantic search.

Ph.D. Student – Multimodal Analysis of Video Lectures, IRIT & IPAL (Singapore)   Oct 2021 – Feb 2025
- Research focused on multimodal analysis of video lectures by combining visual (slide layout, content) and textual (transcripts, OCR) information for segmentation, classification, and summarization tasks.
- Built and released an open-source toolkit that adds interactive features to video lectures using slide recognition, transcript alignment, RAG, and timeline segmentation for better understanding (github.com/travisseng/svla-toolkit)
- Created a synthetic slide generator to train visual models that can handle diverse styles and layouts, improving performance of visual analysis of slides
- Designed both unsupervised and supervised temporal segmentation pipelines: one based on aligning visual and language cues, the other by finetuning large language models to identify lecture structure.

Research Engineer, **IRIT**, Toulouse   Oct 2020 – Oct 2021
- Conducted research on the automatic detection and segmentation of defects (scratches, blotches) in old films using deep learning and synthetic data generation.
- Applied correlation networks to segmentation models to improve performance.

Projects

Video Lecture Analysis Toolkit (Accepted at ACM Multimedia)   github.com/travisseng/svla-toolkit
- Interface leveraging LLMs, vision and language models, OCR, and transcript parsing for multimodal interactive navigation of lecture videos.

Rewardly - Habit Tracker, Mobile App   apps.apple.com/us/app/rewardly-habit-tracker
- React Native app with personalized reward loops.

Quizzy - Online Quizz Game, Full-stack Web App   quizz.travis-seng.fr
- Multiplayer quiz platform built with React, Node.js, and WebSocket backend for live game experiences.

Publications

Multimodal Temporal Video Lecture Segmentation via LLM, **T. Seng**, A. Carlier, W. T. Ooi, **WACV 2026** (Submitted)

Video Lecture Analysis Toolkit, **T. Seng**, A. Carlier, W. T. Ooi, **ACM Multimedia Open Source 2025** (Accepted)

SlideCraft: Synthetic Slides Generation for Robust Slide Analysis, **T. Seng**, A. Carlier, T. Forgione, V. Charvillat, W. T. Ooi, **ICDAR 2024**

Correl-Net: Defect Segmentation in Old Films Using Correlation Networks, A. Renaudeau, **T. Seng**, A. Carlier, J. D. Durou, **SCIA 2023**

Enriching Existing Educational Video Datasets to Improve Slide Classification and Analysis, **T. Seng**, **ACM Multimedia 2022**

Learning Defects in Old Movies from Manually Assisted Restoration, A. Renaudeau, **T. Seng**, A. Carlier, F. Pierre, F. Lauze, JF. Aujol, JD. Durou, **ICPR 2020**

Skills

LLMs, NLP & Multimodal AI: Fine-tuning (LoRA, PEFT), RAG, prompt engineering, LangChain, vector databases (FAISS, Qdrant)

Deep Learning: PyTorch, Transformers, HuggingFace, LoRA, Quantization, synthetic data generation

Engineering & Infra: Python, C++, JavaScript, React, Node.js, Git, REST APIs, Docker

Teaching & Communication: Probability, Deep Learning, Mobile Programming (TA experience at ENSEEIHT)

French (native), **English (fluent)**

Education

Ph.D in Computer Science, INP Toulouse   2021 – 2024
Diplôme d'ingénieur (M.Sc) in Computer Science, ENSEEIHT   2017 – 2020
Exchange in Computer Science, University of Birmingham   2019 – 2020