



Yannis Tevissen

HEAD OF RESEARCH

Video Understanding · Multimodality

CONTACT

- ✉ yannis.tevissen@orange.fr
- ☎ +33 6 46 61 67 04
- 🌐 yannistevisse.fr
- 🐙 GitHub
- 👤 Hugging Face
- 🎓 Google Scholar
- 🌐 LinkedIn

LOCATION

📍 Paris, France

LANGUAGES

- French:** Native 🇫🇷
- English:** Bilingual 🇬🇧
- German:** Basic 🇩🇪

RESEARCH INTERESTS

Video understanding, world models, multimodal agents, AI fairness, and leading research organizations.

OTHER ACTIVITIES

- **TEDx Speaker** — Talk on disability biases in AI systems (Mar 2024)
- **OECD Panelist** — Invited speaker on AI and disability (Dec 2024)
- **Board Member** — Universal Wings Mobility (Jul 2025)
- **DiverseSpectrum** — Open-source dataset contribution (Jan 2026)
- **UpToCure** — AI platform for rare disease research reports (Mar 2025)
- **Mobility Project** — Installed a robotic open-source arm on my wheelchair

SKILLS

- Multimodal AI
- Video Understanding
- Vision-Language Models
- AI Systems
- Research Strategy
- MLOps
- Python
- Research Leadership
- Open-source AI
- Speaker Diarization
- AI Fairness
- Ethics Advocacy

EXPERIENCE

- **Head of Research** Dec 2023 – Present
Moments Lab · Paris, France
Defines the company's **research roadmap** for **multimodal video understanding**, spanning vision-language models, retrieval-augmented video systems, and production-scale AI workflows. Leads **model and system architecture** decisions for long-form video understanding, from research prototypes to scalable production deployment. Built strategic partnerships with top AI labs and major media corporations, and shipped state-of-the-art video understanding technologies (MXT-1, 1.5 & 2). MXT-2 powers workflows processing more than **10K long-form videos per month**.
- **Research Scientist** Aug 2020 – Dec 2023
Moments Lab · Paris, France
Industrial PhD on multimodal speaker diarization, advised by Prof. J. Boudy and Prof. Em. G. Chollet. Contributed to scaling product impact from zero to multimillion business value through research-driven innovation.
- **Visiting Lecturer** May 2022 – Present
Telecom SudParis · Evry, France
Teaching AI, NLP, and speech processing in the DATAPAC Master's program.
- **Co-founder & CTO** May 2019 – 2024
VocaCoach · Paris, France
Built a tech startup around proprietary voice AI technology and drove the company from zero to a 1M valuation.
- **Research Intern** Jul 2018 – Aug 2018
Renault Group · Guyancourt, France
Debugged the perception module of autonomous vehicles and studied V2X communications.

EDUCATION

- 🎓 **Institut Polytechnique de Paris**
Ph.D. in Computer Science, AI, Data
2021 – 2023
- 🎓 **Telecom SudParis**
M2 in General Engineering
2017 – 2020
- 🎓 **IMT Business School**
M2 in Strategy & Entrepreneurship
2019 – 2020

PATENTS

- F. Petitpont, P. Petitpont, **Y. Tevissen**, K. Guetari (2025). *Computer-based platforms and methods for efficient AI-based digital video shot indexing*. US 12,288,377
- F. Petitpont, **Y. Tevissen**, K. Guetari (2024). *Systems and methods for AI generation of image captions enriched with multiple AI modalities*. US 12,148,233

PUBLICATIONS

- 1 M. Brkic, A. Filali Razzouki, **Y. Tevissen**, K. Guetari, M. El Yacoubi (2025). Frame Sampling Strategies Matter: A Benchmark for Small Vision Language Models. *arXiv:2509.14769*. [\[Open-source benchmark\]](#)
- 2 K. Guetari, **Y. Tevissen**, F. Petitpont (2025). Vision and Language Models for Enhanced Archive Video Management. *NAB BEIT 2025*.
- 3 **Y. Tevissen**, K. Guetari, F. Petitpont (2024). Towards Retrieval Augmented Generation over Large Video Libraries. *HSI 2024*. [\[Best Presentation Award\]](#)
- 4 **Y. Tevissen** (2024). Disability Representations: Finding Biases in Automatic Image Generation. *CVPR AVA Workshop 2024*.
- 5 G. Chollet, H. Sansen, **Y. Tevissen** et al. (2024). Privacy Preserving Personal Assistant with On-Device Diarization and Spoken Dialogue System. *IHIET 2024*.
- 6 K. Guetari, **Y. Tevissen**, F. Petitpont (2024). Multimodal Chaptering for Long-Form TV Newscast Video. *arXiv:2406.17590*.
- 7 **Y. Tevissen**, K. Guetari, M. Tassel, E. Kerleroux, F. Petitpont (2024). Inserting Faces inside Captions: Image Captioning with Attention Guided Merging. *arXiv:2405.02305*. [\[Dataset published\]](#)
- 8 **Y. Tevissen** (2023). Multimodal diarization: towards robustness and fairness in the wild. *Ph.D. Thesis, IPP*.
- 9 **Y. Tevissen** et al. (2023). Home monitoring for frailty detection through sound and speaker diarization. *JETSAN 2023*. [\[Best Presentation Award\]](#)
- 10 **Y. Tevissen**, J. Boudy, G. Chollet, F. Petitpont (2023). Towards measuring and scoring speaker diarization fairness. *arXiv:2302.09991*.
- 11 **Y. Tevissen**, J. Boudy, G. Chollet, F. Petitpont (2023). Detection d'activite vocale multi-flux pour la diarisation du locuteur. *GRETSI 2023*.
- 12 **Y. Tevissen**, J. Boudy, G. Chollet (2022). Multi-stream voice activity detection for robust speaker diarization. *GDR ISIS 2022*.
- 13 **Y. Tevissen**, J. Boudy, F. Petitpont (2022). The Newsbridge–Telecom SudParis VoxCeleb Speaker Recognition Challenge 2022 System. *VoxSRC 2022*.