

Manuel Kansy

✉ manuel.kansy@gmail.com

in manuel-kansy

🌐 mkansy.github.io

Professional Summary

Deep learning researcher with a PhD focused on generative models for image/video synthesis and editing. Experienced in diffusion / flow matching, multimodal generation, large-scale data and training, and controllable video generation, with growing focus on motion/action representations, world models, and physical AI.

Employment History

- 01/2026 – now ♦ **Research Scientist.** Luma AI. Remote (Switzerland), then Redwood City, USA from 10/2026.
 - ♦ Developed video captioning and SFT data rebalancing, contributing to Ray3.2.
 - ♦ Investigated Ray3.2 training instability across code, dependencies, and infrastructure.
 - ♦ Led *Creative Threads* project, defining research priorities for future video models.
 - ♦ Expanding research into world models and physical AI, with a focus on motion representations and goal-conditioned robotics.
 - ♦ Developed large-scale video data pipelines for captioning and human tracking.
- 01/2025 – 12/2025 ♦ **Associate Research Scientist.** DisneyResearch|Studios / Facial VFX Group. Zurich, Switzerland.
 - ♦ Led *Face Foundation Model* project (with diffusion / flow matching).
 - ♦ Collaborated with internal teams to align research with production goals.
 - ♦ Designed and trained diffusion transformer architectures.
 - ♦ Supervised students (production-ready face swapping, data pre-processing).
- 05/2021 – 12/2024 ♦ **Computer Science Ph.D. Student.** ETH Zurich / Computer Graphics Laboratory & DisneyResearch|Studios / Machine Intelligence and Data Science Group. Zurich, Switzerland.
 - ♦ Contributed to face swapping technology used in several movie productions.
 - ♦ (Co-)authored 4 peer-reviewed publications and 7 patent applications.
 - ♦ Supervised 10+ students on various topics (e.g., face swapping, audio-conditioned video generation, visual dubbing, text-to-speech, voice conversion).
 - ♦ Served as TA for 5 undergrad computer science C++ courses, incl. 2x Head TA.
- 06/2019 – 10/2019 ♦ **Research Intern.** Technical University of Munich / Dynamic Vision and Learning Group. Garching, Germany.
 - ♦ Integrated C++ graph algorithms for multiple object tracking into Python.
- 10/2015 – 09/2018 ♦ **Cooperative Student.** Robert Bosch GmbH. Stuttgart (area), Germany and Thailand.

Education

- 05/2021 – 12/2024 ♦ **Ph.D. (Dr. sc. ETH), Computer Science.** ETH Zurich. Zurich, Switzerland.
Thesis title: *From Still Faces to Dynamic Reenactment: Human-Centric Visual Analysis and Generation.*
- 10/2018 – 03/2021 ♦ **M.Sc., Robotics, Cognition, Intelligence (Computer Science).** Technische Universität München. Garching, Germany.
Final grade: 1.3 (where 1.0 is best and 4.0 worst)
- 10/2015 – 09/2018 ♦ **B.Eng., IT-Automotive (Computer Science).** Duale Hochschule Baden-Württemberg Stuttgart. Stuttgart, Germany.
Final grade: 1.1 (where 1.0 is best and 4.0 worst)
- 09/2011 – 06/2015 ♦ **High School Diploma.** Northville High School. Northville, Michigan, USA.
Final grade: 4.0 (where 4.0 is best and 0.0 worst)

Research Publications

Conference Proceedings

- 1 X. Ji, S. Weiss, **M. Kansy**, J. Naruniec, X. Cao, B. Solenthaler, and D. Bradley, “FastGHA: Generalized few-shot 3d gaussian head avatars with real-time animation,” in *The Fourteenth International Conference on Learning Representations*, 2026.
- 2 **M. Kansy**, J. Naruniec, C. Schroers, M. Gross, and R. M. Weber, “Reenact anything: Semantic video motion transfer using motion-textual inversion,” in *Proceedings of the Special Interest Group on Computer Graphics and Interactive Techniques Conference Conference Papers*, 2025, pp. 1–12.
- 3 S. Sadat, **M. Kansy**, O. Hilliges, and R. M. Weber, “No training, no problem: Rethinking classifier-free guidance for diffusion models,” in *The Thirteenth International Conference on Learning Representations*, 2025.
- 4 **M. Kansy**, J. Balletshofer, J. Naruniec, C. Schroers, G. Mignone, M. Gross, and R. M. Weber, “Self-supervised effective resolution estimation with adversarial augmentations,” in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2023, pp. 573–582.
- 5 **M. Kansy**, A. Raël, G. Mignone, J. Naruniec, C. Schroers, M. Gross, and R. M. Weber, “Controllable inversion of black-box face recognition models via diffusion,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023, pp. 3167–3177.

Patents and Patent Applications

- 1 J. K. Naruniec, **M. J. Kansy**, G. Mignone, C. R. Schroers, and R. M. Weber, *Generating an image including a source individual*, US Patent 12,555,273, Feb. 2026.
- 2 **M. J. Kansy**, J. K. Naruniec, C. R. Schroers, and R. M. Weber, *Semantic video motion transfer using motion-textual inversion*, US Patent App. 19/210,425, 2025.
- 3 **M. J. Kansy**, J. Neuner-Jehle, J. K. Naruniec, and R. M. Weber, *Subject-agnostic face swapping with low-rank adaptation*, US Patent App. 19/094,056, 2025.
- 4 J. K. Naruniec, A. Bionda, **M. J. Kansy**, C. R. Schroers, and R. M. Weber, *Visual dubbing of an audiovisual sequence*, US Patent App. 18/642,219, 2025.
- 5 S. Sadat, **M. J. Kansy**, and R. M. Weber, *Independent condition guidance for diffusion models*, US Patent App. 19/034,441, 2025.
- 6 S. Sadat, **M. J. Kansy**, and R. M. Weber, *Time step guidance for diffusion models*, US Patent App. 19/034,446, 2025.
- 7 **M. J. Kansy**, A. J. Raël, J. K. Naruniec, C. R. Schroers, and R. M. Weber, *Identity-preserving image generation using diffusion models*, US Patent App. 18/320,857, 2023.

Skills

Generative Modeling	◇ World models, unified models, diffusion / flow matching models (DiT, U-Net; latent space, pixel-space), GANs, autoencoders.
Programming & Workflows	◇ Python, C++, AI-assisted development, large-scale training and evaluation.
Frameworks & Infrastructure	◇ PyTorch, TensorFlow, HF Diffusers, Linux, Git, SLURM, ONNX, ffmpeg.
Languages	◇ German (native), English (fluent), Spanish (intermediate), Mandarin (basic).