

Federico Betti

📍 Zurich, Switzerland
✉ federico.betti@outlook.com
🌐 fbetti.com
in federico_betti
🔗 federicoBetti



Professional Summary

Senior ML Engineer and PhD with 5+ years building and shipping production NLP and multimodal systems. Lead the NLP/LLM roadmap at Y Combinator-backed Typewise, designing retrieval-augmented agents and real-time text prediction for thousands of users. Strong track record of turning state-of-the-art research in multimodal LLM-based evaluation and generative models into robust production systems, with publications at ICCV, ECCV, ACM Multimedia, and IJCV.

Professional Experience

Senior ML Engineer

May 2022 – Present

Typewise - ETH Zurich

- Own NLP/LLM roadmap for Typewise (YC S22), shipping real-time multilingual text prediction, grammar correction, and conversational agents to 4000+ users.
- Design and implement retrieval-augmented, hierarchical AI agent architectures for customer service automation with human-in-the-loop constraints using LangGraph, achieving over 90% automated resolution rates in specific flows.
- Drove adoption of TDD and tooling, increasing automated test coverage from 15% to 80% across core NLP services.
- Build and maintain production Python services for low-latency inference (<250ms), evaluation, and data pipelines on AWS (testing, CI/CD, observability) to balance accuracy, latency, and cost.
- Collaborate with ETH Zurich research lab to translate AI research into product features; patent pending.
- Work closely with Typewise founders on product strategy, driving a full product pivot and supporting hiring (5+ ML/engineering hires) and internal process optimization.

AI Research Fellow

Jan 2021 – May 2022

Leonardo Labs

- NLP research and development on the davinci-1 HPC to deliver transformer-based AI models for semantic similarity, retrieval, and sentiment analysis using PyTorch. Implemented parallel training over 64+ A100 GPUs.
- Managed cross-functional AI initiatives and projects across different branches of the company, aligning timelines, technical documentation, and deliverables.

AI Engineer

Jan 2020 – Jan 2021

Meta-liquid

- Developed production-grade multimodal AI models in Scala, adapting transformer and generative architectures to strict latency and cost targets.
- Worked on low-level implementations in CUDA to optimize the performance of specific kernels and models.

Data Scientist

Aug 2019 – Dec 2019

EnergyWay

- Delivered NLP analytics for a leading media client by combining Python pipelines (Pandas, PyTorch, scikit-learn) with automated AWS training and deployment workflows.
- Partnered with Innovation Center stakeholders to scope requirements, ship PoCs, and hand over scalable solutions to operations teams.

Education

PhD in Computer Science

Nov 2022 – Jan 2026

University of Trento

- Research on making diffusion-based image generation models more efficient and reliable, with a focus on reducing wasted compute and detecting hallucinated content early.
- Designed and evaluated methods for instruction-guided image editing, including multimodal LLM-based evaluators and metrics that better match human judgments of image quality and alignment.

Alta Scuola Politecnica

2017 – 2019

Politecnico di Milano

- Selected for the two-year 150-students honors program jointly run by Politecnico di Milano and Politecnico di Torino, focusing on engineering execution, business design, entrepreneurship, and leadership.

Visiting Period

2018

École polytechnique fédérale de Lausanne (EPFL), Switzerland

Double Degree MSc in Computer Science and Engineering

2017 – 2019

Politecnico di Milano / Politecnico di Torino, 110L/110

BSc in Computer Engineering

2014 – 2017

Politecnico Di Milano, 110/110

Main Research Publications

What Changed? Detecting and Evaluating Instruction-Guided Image Edits with Multimodal Large Language Models

L Baraldi, D Bucciarelli*, F Betti*, M Cornia, L Baraldi, N Sebe, R Cucchiara, ICCV 2025*

Hallucination Early Detection in Diffusion Models

F Betti, L Baraldi, L Baraldi, R Cucchiara, N Sebe, IJCV 2025

Optimizing Resources in Diffusion Models through Hallucination Early Detection

F Betti, L Baraldi, L Baraldi, R Cucchiara, N Sebe, ECCV 2024

Let's ViCE! Mimicking Human Cognitive Behavior in Image Generation Evaluation

F Betti, J Staiano, L Baraldi, L Baraldi, R Cucchiara, N Sebe, ACM Multimedia 2023

Controlled Text Generation using Adversarial Networks

F Betti, G Ramponi, M Piccardi, INLG 2020

Technical Skills

AI/ML Frameworks: PyTorch, TensorFlow, Hugging Face, LangGraph, LangChain

Programming Languages: Python, Scala, Java, C, SQL, HTML, CSS, CUDA

Cloud & Infrastructure: AWS (EC2, S3, Lambda), MLOps, Docker, Kubernetes, HPC Systems

AI Specializations: Natural Language Processing, Computer Vision, Large Language Models, Generative AI, AI Agents

Development & Tools: Software Development, Web Development, Linux OS, Git, Agile Frameworks

Leadership & Management: Project Management, Team Leadership, Product Development, Strategic Planning, Cross-functional Collaboration

Additional Experiences

SEI Pioneer

2018 – 2019

Fondazione Agnelli, Alta Scuola Politecnica

Selected for the ASP/SEI Pioneer entrepreneurship track, co-designing a tech venture with industry mentors and refining investor-ready product strategy.

Knowai: Knowledge Extraction from Conversations using NLP September 2018 – June 2019
Fondazione Agnelli

Co-founded a proto-startup that built NLP pipelines and dashboards to surface institutional knowledge from conversations, delivering pilots that attracted investor and enterprise interest.

Polimi Data Scientists - Board Member 2018 – 2019
Politecnico di Milano

Coordinated industry-backed data science events and partnerships that connected students, researchers, and workers across the Politecnico di Milano community.

Languages

Italian: Mother Tongue

English: Professional Knowledge *IELTS Test 2016*

Interests

Innovation and Startup: I have completed MBA-like courses with ASP and Fondazione Agnelli, cultivating an entrepreneurial mindset that, when combined with my technical expertise, inspires my pursuit of innovative yet tangible solutions. Currently working within a Y Combinator-backed startup, my interests in emerging technologies continue to grow. I also maintain my knowledge base and stay current with industry trends through extensive reading.

Technology: Since a young age, I have been fascinated by new technology trends. More recently, I've focused on non-professional drone flight.

Sport: I have played football and basketball since a young age, and I still play them with friends. I consider them a fantastic way to meet new people in different contexts.