

# Halil Erdogan

Senior AI Engineer | Software Engineer

✉ herdoganturkey@gmail.com | 🌐 herdogan.dev | 📄 github.com/halil-erdogan | 📄 halil-erdogan | 📞 +352691131586

## Technical Skills

- **Programming Languages:** Python, C++, SQL, Zig, Shell Scripting
- **GenAI Applications:** Conversational AI systems, LLM-powered content generation, RAG Systems, Voice Agents, Text Agents, Video Agents, Automated Copywriting, Semantic Search, AI-powered Analytics, Multi Agents
- **ML & GenAI Frameworks:** PyTorch, FastAPI, MLflow, scikit-learn, VAPI, Tavus, LangChain, OpenAI API, Langfuse, LlamaIndex, Transformers, Anthropic Claude API, Pinecone, ChromaDB, FAISS
- **Big Data:** Apache Airflow, Apache Spark, PostgreSQL, MongoDB, DuckDB, Redis, Pandas
- **DevOps & Cloud:** Docker, Kubernetes, AWS, GCP, Azure, Cloudflare, Terraform, Git, CI/CD
- **Monitoring & Analytics:** Grafana, Prometheus, ELK Stack, PowerBI, Looker, Tableau

## Professional Experience

<b>INFINIT</b> Senior AI Engineer	February 2025 - Present Luxembourg
<ul style="list-style-type: none"><li>• Engineered LangChain-based multimodal agent for loan application processing, integrating text and video analysis to automate document verification and applicant interviews, reducing processing time by 80% and improving approval accuracy by 35%</li><li>• Developed conversational AI voice agent using VAPI and OpenAI that reduced human dependency by 70% for routine inquiries while maintaining 95% customer satisfaction</li><li>• Architected ML-based vehicle evaluation system reducing inventory review time by 50%, enabling efficient prioritization of high-value listings</li></ul>	
<b>Aiva Technologies</b> Senior AI Engineer	January 2019 - September 2024 Luxembourg & Remote
<ul style="list-style-type: none"><li>• Progressed from intern to technical team lead, driving growth to \$1M ARR and securing investment at \$30M+ valuation</li><li>• Architected high-performance ML pipeline processing 1M+ monthly compositions with 99.9% reliability and 90% reduced latency</li><li>• Led migration from traditional ML to transformer-based ML resulting in 300% YoY user growth and scaling to 50,000+ MAU</li><li>• Built real-time music generation APIs achieving 50ms response time and designed fault-tolerant microservices architecture using Docker/Kubernetes</li></ul>	
<b>Fraunhofer IAIS</b> Research Intern	June 2018 - September 2018 Sankt Augustin, Germany
<ul style="list-style-type: none"><li>• Engineered first-of-its-kind speech segmentation system reducing manual annotation time from hours to minutes, leading to widespread adoption across organizational projects</li></ul>	
<b>Telecom ParisTech</b> Research Intern	June 2017 - September 2017 Paris, France
<ul style="list-style-type: none"><li>• Developed novel probabilistic model for speech enhancement, published in ICASSP 2018</li></ul>	

## Education

<b>Friedrich-Alexander-Universitat Erlangen-Nurnberg</b> MSc. in Communication and Multimedia Engineering	September 2019 - December 2021 Erlangen, Germany
<ul style="list-style-type: none"><li>• GPA: 1.9/1.0 (German grading system, equivalent to distinction)</li><li>• Thesis: 📄 <u>Data-Driven Singing Activity Detection for Music Recordings</u></li></ul>	
<b>Sabanci University</b> BSc. in Electronics Engineering	September 2014 - January 2019 İstanbul, Türkiye
<ul style="list-style-type: none"><li>• GPA: 3.78/4.00</li></ul>	