



Filipp Trigub

Ingénieur systèmes IA

Adresse postale

Friedrich-Wöhler-Str 1

53117 Bonn

Téléphone: +49 178 404 2916

+33 75 300 9502

E-mail: filipp@trigub.tech

[LinkedIn](#) | [Web](#) | [GitHub](#)

Specialization

Architecture de solution de bout en bout pour les applications alimentées par l'IA agentique

Langue de Codage

Proficient: Python, TS, SQL

Familiar: R, C, C++, Flutter

Outils

Agentic: Claude Code, Cline, BMAD

LLM: LangGraph, Haystack, MCP, Vertex, Anthropic

ML: Tensorflow, Keras, SkLearn, PyTorch

Infra: Docker, Terraform, Azure CLI, GCP CLI, GitHub Actions, uv, poetry, GitLab

Cloud: Azure, AWS, TerraForm

Database: Oracle, PostgreSQL, MySQL, MongoDB, Weaviate, Qdrant

Langues

Bilingue: Anglais, Allemand, Français

Courant: Ukrainian

L'année dernière, j'ai apporté **4 419 contributions sur GitHub** à des projets que j'ai menés jusqu'à leur mise en production. Mon travail se concentre sur les systèmes d'IA et les intégrations via des clients, des API et des MCP, mais il couvre aussi, et surtout, le backend, le frontend, le cloud, les tests, le CI/CD et le mobile. Je crée des solutions fiables et je serais ravi de le faire avec vous.

Physicien théoricien avec une expérience dans **le conseil en technologie et en stratégie**. Il a fait ses preuves en convertissant les besoins des clients en propositions d'architecture technologique, en les améliorant de manière itérative et en mettant en œuvre la solution finale.

Parmi mes anciens clients, je compte des startups, des cabinets de conseil, de grandes banques, une **banque centrale internationale** et des multinationales dans le secteur pharmaceutique et de l'hébergement cloud.

Spécialisé dans le **développement d'applications d'IA agentique**, je fais passer des systèmes multi-agents complexes du stade de preuve de concept à celui de la production. Au-delà de cela, je suis un développeur expérimenté en **Python (9 ans)**, **TS (3 ans)** et **SQL (5 ans)** avec de vastes compétences en T, notamment dans le traitement des données, la formation et la mise en œuvre de modèles.

Je suis à l'aise pour travailler de manière indépendante ou en équipe, à distance ou sur site.

Formation

M.Sc. en physique théorique

FR 15 (GPA 3.6)

Université Humboldt de Berlin

10.2013 - 03.2019

Thèse

- "Mécanismes de dissipation makoviens et non-markoviens dans les forces de dispersion à l'équilibre".
- Modélisation mathématique des processus stochastiques avec effets de mémoire. Validation numérique du modèle à l'aide d'une implémentation en C.

Connaissances pertinentes

- Analyse des données statistiques : bases mathématiques et application pratique du traitement des données et de l'apprentissage automatique sur des données réelles.
- Mécanique statistique : statistiques théoriques de niveau avancé avec expérience en codage en Python, C et Fortran.
- Mécanique des fluctuations : compréhension approfondie des processus complexes régis par le hasard, favorisée par des exercices mathématiques et numériques avancés.

Expérience professionnelle



Titre : Plateforme collaborative d'IA communautaire multi-agents

Client : huuh.me

Durée : 6 mois (01.2025-07.2025)

Objectif :

- Permettre aux créateurs de contenu et aux responsables de communauté de monétiser leur contenu de manière native grâce à l'IA.

Solution :

- Création d'une base de connaissances avancée pour alimenter un système RAG avec segmentation sémantique, saisie multiformat, attribution de pages et filtrage des métadonnées.
- Création d'une interface réactive et de systèmes backend gérant l'authentification, la gestion des données et les services auxiliaires sous forme d'API RESTful.
- Activation de l'interaction avec la base alimentée par un système multi-agents.
- Activation d'un accès basé sur les rôles à la base et à l'agent afin de favoriser la collaboration grâce à des contributions de connaissances et d'outils d'IA.
- Activation d'une place de marché et de la monétisation via l'intégration de Stripe.



Titre : Bot de conseil fiscal

Client : colmo Ltd.

Durée : 3 mois (09.2024-12.2024)

Objectif :

- Permettre aux clients de la société de conseil d'obtenir des réponses générées par l'IA auxquelles ils peuvent se fier.

Solution :

- Compréhension des données propriétaires. Développement d'analyseurs syntaxiques personnalisés pour différentes structures de données.
- Développement d'approches de segmentation personnalisées afin de préserver la cohérence des segments de texte dans la mesure du possible et du raisonnable.
- Mise en œuvre d'une optimisation de l'alignement basée sur les thèmes identifiés et la recherche de base comme stratégie de récupération de secours.
- Déploiement avec interface utilisateur et backend dans le cloud.



Titre : Développement rapide d'un MVP pour une géodonnées start-up

Client : HB Partners

Durée : 2 mois (06.2025-08.2025)

Objectif :

- Développer en moins de 50 heures une application web prête à être présentée, interagissant de manière dynamique avec des fichiers KMZ locaux et Supabase.

Solution :

- Mise en place d'une coordination claire avec le chef de projet.
- Refactoring du code existant.
- Formulation d'exigences et d'étapes de mise en œuvre claires.
- Développement rapide avec Claude Code et des outils de codage agentique, avec un schéma de base de données et une structure de code en constante évolution.
- Garantie des fonctionnalités requises avant la date limite et dans le respect du budget de 50 heures.

Expérience professionnelle



Titre : Co-pilote LLM multicanal à usage universitaire

Employeur : infolab.ai

Client : Startup

Durée : 14 mois (10.2023-12.2024)

Objectif :

- Construire un copilote d'IA sophistiqué pour améliorer l'expérience d'apprentissage des étudiants et fournir des informations dérivées de l'IA aux professeurs.

Solution :

- Construire un système LLM multi-chaînes s'adaptant aux besoins de l'utilisateur avec un système RAG basé sur Weaviate vector DB et l'évaluer avec Ragas.
- Construire un frontend responsive react et des systèmes backend gérant l'authentification, la gestion des données et les services auxiliaires sous forme d'API RESTful.
- Déployer et gérer l'application dans le nuage dans un environnement de production incluant le CICD via un déploiement en plusieurs étapes.



Titre : Conception et développement d'applications complètes

Employeur : d-fine GmbH

Client : Banque centrale internationale

Durée : 24 mois (01.2021-12.2022)

Objectif :

- Concevoir et développer une application pour l'analyse de données statistiques de manière holistique et soutenir les utilisateurs à la demande.

Solution :

- Conception et développement d'extensions architecturales basées sur les demandes et les limites du client.
- Implémentation réussie de plus de 500 stories dans une base de code de plus de 300k LoC, qui incluait un backend Python hébergé dans AWS avec PostgreSQL et Oracle DB, ainsi qu'un frontend JS, tous deux livrés via CI/CD.
- Il a assuré une stabilité continue grâce à des tests unitaires et de bout en bout approfondis.
- J'ai été Scrum Master pour une équipe de 7 développeurs pendant plus d'un an.



Titre : GPT-Assistant du personnel de service alimenté

Employeur : d-fine GmbH

Client : Fournisseur international de services d'hébergement

Durée : 2 mois (04.2023-06.2023)

Objectif :

- Fournir un chatbot alimenté par GPT utilisant des documents internes pour soutenir le personnel de service et interagir avec les clients directement.

Solution :

- A dirigé la conception et le développement d'un chatbot léger alimenté par GPT pour le soutien du personnel de service d'un client international.
- Il a contribué de manière significative à la conception initiale, à la sélection de la technologie et à l'architecture.
- Mise en œuvre d'un pipeline haystack avec intégration d'OpenAI pour optimiser l'utilisation de GPT et intégration de CI/CD et DevOps pour un développement rapide et collaboratif.
- A dirigé le développement technique et introduit des pratiques de développement agile, en établissant des lignes directrices et des conventions pour l'équipe.

Expérience en matière de projets PdC



Titre : Telegram Search&Chat

Client : PME

Durée : 1 mois (07.2023)

Objectif :

- Rendre Telegram consultable. Permettre l'exploration automatisée d'une multitude de canaux Telegram via un GPT-bot alimenté par des documents vectorisés trouvés par un mécanisme de recherche.

Solution :

- Récupération du contenu des chats de Telegram avec Telethon et chargement dans une base de données FAISS avec des embeddings OpenAI hébergés dans un backend FastAPI.
- Implémentation d'un pipeline de questions-réponses avec Haystack, utilisant Dense Passage Retrieval et la génération OpenAI.
- Fourniture d'une interface utilisateur légère avec authentification de l'utilisateur à Telegram et fenêtre de chat ultérieure.
- Déploiement du bot sur Azure avec Docker.



Titre : Sommelier IA

Client : Détailleur en ligne

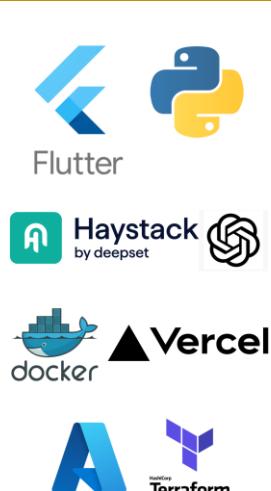
Durée : 1 mois (08.2023)

Objectif :

- Construire un bot de vente pour un magasin de vin avec une identification et une mémorisation cohérentes des préférences du client afin de lui offrir une expérience similaire à celle d'un sommelier.

Solution :

- Construction d'un agent LLM avec un outil de recommandation basé sur un modèle général d'agent haystack avec une interface utilisateur Streamlit.
- Mise en œuvre d'un méta-prompt spécifié pour identifier correctement et de manière cohérente les préférences du client et les mémoriser.
- Extraction programmatique de ces préférences pour forcer l'utilisation de l'outil après l'identification d'un certain nombre de préférences.
- Dockisation et déploiement sur Azure avec Terraform.



Titre : Audio-Summarizer

Client : Startup

Durée : 1 mois (09.2023)

Objectif :

- Résumer des enregistrements audio, en particulier des livres, en points concis.

Solution :

- Transformation d'un enregistrement en texte à l'aide de STT, ici les modèles Whisper et Huggingface.
- Développement d'un méta-prompt et d'une boucle de résumé qui divise l'enregistrement en parties (chapitres) définies par l'utilisateur et en extrait un résumé concis.
- Création d'une interface utilisateur basée sur Flutter, qui permet aux utilisateurs de résumer leurs enregistrements et de recevoir le résumé par e-mail.
- Dockerisation et déploiement sur Vercel et Azure avec Terraform et CI/CD.