ALAN BLANCHET

Ingénieur IA — Agents, systèmes LLM & ML

Motivé à rejoindre Mistral pour renforcer son offre souveraine.

Français — natif **English** — bilingue (3 ans R-U)

Grenoble, France
contact@alan-blanchet.fr
07 81 86 62 87
alan-blanchet.fr
github.com/AlanBlanchet
linkedin.com/in/alanblanchet

5+ ans en IA / ML · **~900 k€+ de projets livrés** · **Construit & opère la stack d'agents IA de l'équipe**

PROFIL

Ingénieur IA/ML surtout autodidacte, 5+ ans — systèmes de vision, de parole et de LLM en production, désormais centré sur les agents IA et les applis LLM de bout en bout, avec le réflexe de creuser le fonctionnement réel des modèles. J'aime un code propre, fortement généralisé et typé — construit avec des agents, mais pas « vibe-codé ». Je veux une IA fiable et bien encadrée, et une Europe qui ne soit pas prise en étau entre les États-Unis et la Chine : d'où une volonté forte de contribuer à une offre souveraine de premier plan. Pragmatique pour autant — valider d'abord avec les modèles les plus performants pour lever le risque, puis basculer vers le souverain là où il tient.

STACK PRINCIPAL

Agents & LLMOps	LiveKit Agents · LangGraph · OpenAI Agents SDK · MCP (crée des serveurs) · tool-calling · LLM-as-judge · Langfuse · LiteLLM · Ollama · LM Studio · AGENTS.md / CLAUDE.md
LLMs	Qwen · Llama · DeepSeek · Mistral · GPT-OSS · Claude / GPT / Gemini (API) · RAG · fine-tuning / LoRA · MoE · KV-cache · quantification · vLLM
Voix — STT	Whisper · faster-whisper · NVIDIA Parakeet · Kyutai · Silero · AssemblyAI · Deepgram
Voix — TTS	ElevenLabs · Cartesia · Deepgram Aura · Kokoro · Piper · Voxtral · Kyutai / Moshi
Audio & VAD	Silero VAD · EBEN · HiFi-GAN / ++ · MP-SENet · xLSTM-SENet · VoiceFilter · DTLN · RNNoise · SpeechBrain · PESQ / STOI / SI-SDR
Vision	RT-DETR · D-FINE · YOLO (v8/NAS) · DETR · SAM · SegFormer · U-Net · DINOv2/v3 · iBOT · SwAV · BYOL · EquiMod · ViT / Swin · CLIP / SigLIP · Mask R-CNN
Appariement d'images	GIM · RoMa · DKM · LoFTR · LightGlue · SuperPoint · OmniGlue · RANSAC / MAGSAC · homographie / TPS
Inférence · RL	TensorRT · ONNX / graph surgeon · quant. INT8 / FP8 · CUDA / Triton · batching · einops · REINFORCE / VPG · DQN family · R2D2 · PPO · RND · NGU · PER
Systèmes & cloud	Python · Rust · Scaleway (GPU) · AWS · Linux / systemd · Docker / K8s · SLURM (Jean Zay) · Nginx / Caddy / Traefik · MLflow / W&B

Jauge IA — comment chaque élément ci-dessous a été construit : écrit à la main
construit avec mes agents IA orchestrés

Les deux modes en jeu — et les agents eux-mêmes construits en interne.

EXPÉRIENCE — INGÉNIEUR R&D · NEOVISION, GRENOBLE · 2022–PRÉSENT (ALTERNANCE 2022–24, PUIS CDI)

IA / ML au cœur du poste, avec un appui transverse régulier — infra & réseau, full-stack, administration des comptes (admin).

Agent vocal conversationnel — téléphonie temps réel

LEAD TECH · projet 300 k€+ · 2025–présent · IA

Agent téléphonique temps réel, prise de rendez-vous médicaux · pipeline construit de zéro → migré sur LiveKit (avec l'équipe) · LLM/STT/TTS par appel avec repli · tool-calling · vérification de sûreté LLM · observabilité complète · en production · démo interne de toute la gamme, dont le clonage vocal en direct pendant un appel.

LiveKit · LangGraph · STT/TTS multi-fournisseurs · téléphonie SIP · Langfuse

Appui IA interne & outillage d'agents 2024–présent IA

*J'épaule les collègues qui veulent tester une solution ou expérimenter avec une IA via API · outillage construit en interne — serveurs MCP · routage multi-fournisseurs · serveur **LM Studio** auto-hébergé (2× RTX 3090, on-prem/confidentiel) · un système de prompts/standards qui régit la façon dont les agents écrivent le code · automatisation partielle de la génération du **Crédit Impôt Recherche (CIR)** de l'entreprise (2025, réutilisée en 2026).*

serveurs MCP | serveur LLM local (on-prem) | Ollama / LM Studio | enablement IA | standards en prompts

Système de classification d'images multi-têtes LEAD TECH projet 65 k€ 2026 IA

*Classification d'images multi-têtes (DINOv2 + VLM), livrée pour un client · le plus important : le **système agentique maison** qui a permis de la construire.*

DINOv2 | classification multi-têtes | construit par agents | FastAPI / Gradio

Framework de rehaussement & reconstruction de la parole LEAD TECH projet 85 k€ 2025 IA

Reconstruit une parole propre à partir de micros intra-auriculaires · plusieurs modèles de rehaussement (EBEN · HiFi-GAN/++ · MP-SENet · xLSTM-SENet · VoiceFilter) derrière un pipeline · simulateur de transfert acoustique bouche-oreille · PESQ/STOI/SI-SDR · mis en production (FastAPI · Gradio · Docker).

PyTorch Lightning | torchaudio | GANs | débruitage audio | FastAPI / Gradio

Mentorat technique & enseignement — encadrement de stage & coaching d'étudiants 2025

*Encadrement d'un stage de recherche de Master sur les **agents à base de LLM** (soutenu, Grenoble INP/UGA) · agent d'extraction & validation de données d'entreprises B2B · benchmark LLMs × outils de recherche MCP · coaching de **5 équipes d'étudiants ingénieurs à l'ESISAR** (Grenoble INP) sur des projets industriels d'IA — dont un outil de transcription vocale temps réel exigeant des performances sur du vocabulaire métier.*

encadrement de stage | coaching étudiants | agents LLM | recherche MCP | ASR / vocabulaire métier

Appariement & recalage d'images — matching visuel / template LEAD TECH projet ~80 k€ 2025 IA

Recalage d'images / template matching pour l'authentification de produits · matching de points-clés (GIM-DKM · RoMa · LoFTR · LightGlue · SuperPoint · OmniGlue) → RANSAC/MAGSAC → homographie/TPS · robuste aux captures smartphone (lumière · flou · perspective) · plugin d'annotation CVAT de keypoint-matching sur mesure · métriques de précision GTE/CTE.

matching de keypoints | RANSAC / homographie | plugin CVAT | GTE / CTE | PyTorch

Plateforme interne de démonstrateurs ML — démos & solutions internes LEAD TECH 2024–26 IA

Héberge les démonstrateurs clients + solutions internes · 25+ applis (classification · détection · prévision · recommandation · photogrammétrie) · chacune déployée/gérée à partir d'images Docker, liens partageables · orchestration par démo · reverse proxy · accès par rôle.

Next.js | PostgreSQL / Prisma | orchestration Docker | Scaleway

Détection d'objets temps réel — inspection industrielle 2024 IA

*Étude approfondie + ablation des détecteurs SOTA (RT-DETR · YOLO-NAS · YOLOv8) → sélection de **RT-DETR** (alors SOTA) · déploiement edge (ONNX · TensorRT · quant. INT8/FP8).*

RT-DETR | TensorRT / ONNX | quantification | MLflow / DVC

DATAWISE — benchmark d'apprentissage auto-supervisé — R&D, sur supercalculateur national LEAD TECH programme ~400 k€ 2024–25 IA

*Benchmark de modèles de vision auto-supervisés modernes (DINOv2 · DINO · iBOT · SwAV · Barlow Twins) sur ImageNet-1K/22K · CIFAR · Food-101 · SUN397 · COCO · architecture modulaire backbone/neck/head propre · entraîné sur le supercalculateur national **Jean Zay** (SLURM · sweeps Hydra-Submitit · Lightning DDP) · EquiMod (BYOL + LARS) également reproduit. Programme de **~400 k€** financé par la **Région Auvergne-Rhône-Alpes** — actuellement inachevé et en pause.*

SSL | PyTorch Lightning (DDP) | SLURM · Jean Zay

Plateforme de recherche image + texte — catalogue de motifs projet 500 k€ 2022–24 IA

*Lead front-end, plateforme de production 500 k€ · extraction de motifs visuels (**Mask R-CNN**) · recherche par similarité d'image + texte libre (CLIP/OpenCLIP) · a aussi construit des parties du backend de recherche par embeddings.*

React / TypeScript | Mask R-CNN | CLIP / OpenCLIP | AWS | recherche sémantique

R&D classification de bactéries clinique — co-auteur de l'article 2023–24 IA

Co-pilotage ML · segmentation + classification de colonies bactériennes vs références classiques (SVM · K-fold) · rapport scientifique · co-auteur de l'article SPIE évalué par les pairs (ci-dessous).

segmentation | références SVM | validation croisée | rédaction scientifique

~200 **« R&D News » — veille modèles, à toute l'équipe**
Synthèse toutes les 2 semaines sur les nouveaux modèles, architectures & techniques notables. Suivi rapproché des SOTA — LMArena · ARC-AGI · SWE-bench…

PROJETS SÉLECTIONNÉS — surtout personnels & open-source ; plus sur GitHub

Système de prompts & standards multi-agents
Le projet auquel je tiens le plus : il compile mes standards d'ingénierie (code générique, flexible, soigné) en skills, sous-agents et prompts réutilisables sur Claude Code / Codex / Cursor.
personnel · le cœur de ma façon de construire

interact
Serveur MCP qui permet aux agents IA de voir & agir sur un écran — navigateur et bureau — par ce qui est visible, pas par des sélecteurs fragiles. Construit pour améliorer mes propres agents.
github.com/AlanBlanchet/interact

Framework de deep learning de zéro
ResNet / ViT / DETR (matching hongrois) et la famille DQN / PPO réimplémentés à la main, avec un moteur d'autograd écrit à la main — et plusieurs aussi reconstruits en Rust sur LibTorch.
github.com/AlanBlanchet/AI-4-Alan

App d'apprentissage multiplateforme
App Flutter + Rust avec un pipeline de contenu LLM — déduplication par embeddings locaux, inférence par lots et en cache.
personnel · full-stack

Moteur de calcul & viz Rust
Un moteur agnostique du framework avec noyaux GPU/SIMD et bindings Python / JS / WASM — construit presque entièrement avec mes agents IA.
github.com/AlanBlanchet/any-compute

Agent de tri de candidatures
Analyse CV & lettres, classe les candidats avec un LLM, les range et notifie l'équipe — le genre d'agent qui lit peut-être ce CV.
interne · agent planifié

Site CMS bilingue construit par agents
Un site de contenu bilingue qu'une équipe non-technique édite elle-même — échafaudé et rempli en environ une journée, presque entièrement avec des agents Claude. Un petit marqueur de mon gain de productivité.
client · construit par agents en ~1 jour

linux-commands
Mon outillage Linux publié — raccourcis auto-documentés, dont un rebase multi-branches interactif.
github.com/AlanBlanchet/linux-commands

APPRENTISSAGE & PUBLICATIONS

COMMENT J'APPRENDS

Surtout autodidacte — articles, cours en ligne, communautés techniques, certificats de code. Très concret : versions CUDA · ONNX & graph nodes · quantification · optimisation d'inférence · des milliers d'heures sur Linux, serveurs & déploiement. Bâtisseur d'outils compulsif, qui aime aider l'équipe.

PUBLICATIONS

Co-auteur (2e sur 6) d'un article SPIE évalué par les pairs — classification CNN / SVM en imagerie multispectrale biomédicale (2025). DOI : 10.1117/12.3097794.

FORMATION

- **Ingénieur ML** (alternance) — OpenClassrooms, 2022–24
- **Robotique orientée IA** (alternance) — IMERIR, 2021–22
- **DUT Informatique** — IUT Montpellier-Sète, 2019–21
- **Baccalauréat S, mention Bien** — 2019 · Cambridge B2

Diplôme ML en ligne choisi délibérément — les écoles d'ingénieur locales n'étaient pas assez orientées IA ; cette voie m'a laissé le temps d'expérimenter largement et de monter en compétence bien plus vite.

Ouvert à Paris — installation ou trajet hebdomadaire.

Alan Blanchet · Grenoble, France · contact@alan-blanchet.fr · github.com/AlanBlanchet · alan-blanchet.fr