

Octavian Cioaca

INGÉNIEUR SENIOR / STAFF - PRODUIT & PLATEFORMES IA

Paris, France · work@selardi.com · +33 6 95 44 00 87 · linkedin.com/in/selardi · selardi.com

PROFIL PROFESSIONNEL

J'écris du CSS et du JavaScript depuis 2006, et je code encore toutes les semaines. Je suis paresseux dans le bon sens : j'apprends vite parce que je déteste refaire deux fois la même chose. J'ai commencé dans le navigateur, puis les API, puis les plateformes en dessous. Aujourd'hui je construis l'infrastructure qui fait tourner les modèles, dans des startups de 20 personnes comme sur des systèmes à 145 000 utilisateurs. Donnez-moi un problème mal défini et je le porte de bout en bout : contrats, première version qui marche, production. Les agents de code font l'essentiel de la frappe. Je lis tout, et c'est moi qui décide ce qui part.

AXES D'INGÉNIERIE

- **Backend et plateforme** : systèmes distribués, contrats d'API, isolation multi-tenant, sécurité, observabilité, CI/CD.
- **Systèmes d'IA agentique** : orchestration d'agents, retrieval et embeddings, passerelles et catalogues de modèles, garde-fous, evals avec jeux de référence versionnés et seuils bloquants, paquets de preuve pour l'audit de modèles.
- **Plateformes pour développeurs** : schémas d'API, OpenAPI 3, GraphQL, ingestion idempotente, webhooks signés, SDK typés, monorepos, développement avec des agents de code IA.
- **Spécifications et standards** : rédiger les documents sur lesquels une équipe s'appuie - contrats d'API, règles d'ingénierie, modèles de menaces, listes de contrôles de sécurité avec leurs sources - et le processus de revue qui les fait respecter.
- **Modernisation** : legacy vers cloud, due diligence, résorption des backlogs, réécriture par étapes.
- **Qualité et exploitation** : TDD/BDD (Mocha, Chai, Jasmine, Karma, Playwright, Protractor) ; durcissement Linux et web, conteneurs, interfaces accessibles (WCAG/WAI-ARIA), SEO, débogage en production.
- **Leadership** : mentorat, revue de code, prise de parole, redressement des livraisons en retard. En ce moment : Rust, Databricks, Omnigent.

EXPÉRIENCE PROFESSIONNELLE

Directeur de l'ingénierie logicielle (opérationnel) - DOMSELARDI

PARIS, FRANCE | FÉVR. 2024 - PRÉSENT

Je porte la plateforme chez DOMSELARDI de bout en bout - architecture, livraison et sécurité - et j'écris encore le code.

- **0-to-1** : plateforme d'agents IA pour un fonds de private equity, d'un seul fichier spec.md à la production - des agents qui travaillaient à partir de la spec, des tests de contrat et des revues qui décidaient de ce qui partait.
- Construction des hooks d'événements REMIS, taxonomies versionnées et hachages d'événements, scellés en paquets de preuve horodatés, vérifiables hors ligne sans exposer les données métier.
- Conception de l'API REST de Level Zero (trames downlink, métadonnées EO/SAR, suivi en temps réel) qui permet de détecter l'usurpation de signal des drones : les entrées sont fusionnées en un seul produit GeoJSON signé, avec un vérificateur public.
- API pour développeurs - contrats définis par schéma, routes RBAC, idempotence, webhooks signés, SDK typés - et SSO d'entreprise : OIDC avec PKCE S256, JWKS et validation stricte de l'ID token ; SAML 2.0 avec XML validé et détection de rejeu ; jetons liés au tenant ; refresh tokens hachés, révoqués en cas de réutilisation.
- Définition du cadre de travail des agents de code IA sur nos dépôts de production : un document de 800 lignes (quelle instruction prime, ce que chaque tâche peut modifier, règles contre les raccourcis, revue indépendante avant tout commit) et une configuration où plusieurs agents travaillent côte à côte, chacun dans son propre worktree git.
- Gouvernance IA intégrée au produit : chaque chiffre du texte généré est comparé aux données sources, et l'approbation reste bloquée tant qu'un seul diffère ; chaque approbation enregistre le modèle, la version du prompt, le hash d'entrée et l'ampleur des retouches humaines ; les modèles tournent uniquement en local, et c'est vérifié à chaque appel.
- Passerelle LLM et catalogue de modèles sur les workflows agents et les principaux SDK, avec garde-fous et évaluations comparant qualité, sûreté et coût avant déploiement.
- Construction de l'infrastructure multi-cloud de l'offre payante de une plateforme agentique open source de premier plan : 12 000 VM sur quatre régions cloud, en Infrastructure as Code avec cloud-init.
- Passerelle LLM et catalogue de modèles avec garde-fous sur les principaux SDK, et harnais d'evals de quatre fonctionnalités IA en production : jeux de référence versionnés, seuils bloquants (WAPE contre une baseline naïve, rappel, extraction exacte, taux de détection de bugs introduits exprès) et tableau des scores. Si un score baisse, rien n'est livré.

- Orchestration LLM multi-agents avec RAG, et retrieval hybride sur PostgreSQL : pgvector avec index HNSW et recherche plein texte, combinés par reciprocal-rank fusion ; sans résultat pertinent, la réponse est “aucune correspondance” plutôt que du remplissage, et les droits d'accès sont vérifiés dans les deux recherches.
- Passage progressif de JavaScript annoté Flow à TypeScript (web responsive, backend, CLI), sans geler les livraisons.

Technologies : Node.js, TypeScript, Bun, React 19, Vue.js, SolidJS, Next.js 16 (App Router), Express, Vite, shadcn/ui, Tailwind CSS, Python avec FastAPI et Celery, PostgreSQL/PostGIS, ClickHouse, MongoDB, SQL, Supabase, Drizzle ORM, TypeORM, Snowflake, Databricks, Redis, BullMQ, Kafka, SQLite.

Outils et infrastructure : Docker, Nginx, Let's Encrypt, Fail2ban, D3.js, ECharts, GSAP, Kubernetes avec Helm, Terraform, SST, GitHub Actions, pnpm/Turborepo, cloud-init, AWS, Azure, GCP, Sentry, Datadog, Prometheus/Grafana, OpenTelemetry, Vercel, Playwright, GraphQL, OpenAPI 3, API REST, JSON Schema, GeoJSON, MQTT, gRPC/protobuf, RFC 7807, Ed25519, SHA-2/3, RFC 3161, API Salesforce, Magnolia CMS, intégrations de paiement, comptage d'API.

IA et modèles : contribution Flowise open source et SaaS Cloud, LangChain Core, RAG, recherche vectorielle, embeddings, Qdrant, Pinecone, FAISS/faiss-node, passerelle LLM, catalogue de modèles, garde-fous et évaluations, AWS Bedrock, SDK Anthropic Claude, SDK OpenAI, Google Generative AI/Gemini, modèles open-weight auto-hébergés.

Directeur, transformation numérique industrielle et ingénierie logicielle - Inetum

PARIS, FRANCE | SEPT. 2017 - JANV. 2024

J'ai dirigé des équipes de 10 à 30 personnes, réparties sur plusieurs pays, sur des plateformes financières et publiques. Migration cloud, analytique, IA appliquée et DevSecOps, en gardant les mains dans le code.

- Migration de la suite financière legacy d'une banque, 145 000 utilisateurs, vers des microservices cloud Azure, avec gestion de sources, analyse statique de code, tests automatisés et CI/CD.
- Relance de programmes réglementés bloqués, en consolidant la migration par lots et en vidant le backlog de défauts.
- Contribution à 120+ évaluations techniques de recrutement et de profils seniors (ingénierie, support, leadership).
- Déploiement de plateformes publiques hautement sécurisées (finances publiques, signature électronique, santé d'urgence), adoptées à l'échelle nationale.
- KPI de livraison améliorés, budget d'innovation +-4 M EUR, OPEX -~2 M EUR sur un produit d'assurance, centres d'excellence RPA (UiPath, Blue Prism) et travaux IA/ML via Addstones.

Technologies : Docker, Nginx, Java, Django, Node.js, TypeScript, Ruby on Rails, Celery, Kafka, GraphQL, PostgreSQL, Oracle, MongoDB, ClickHouse, Redis, AS/400, IBM Power 9, Kubernetes, Terraform, AWS, Azure, GCP, React, Angular, Vue.js, SolidJS, CI/CD, analyse statique, Splunk/SIEM, Datadog, Prometheus/Grafana, OpenTelemetry, RPA, SEDA, RGS, REU/INSEE, IBM i, Alfresco, SLO/SLA, incidents (N1/N2/N3), DevSecOps.

Architecte logiciel & Team Lead - CustomerMatrix

PARIS, FRANCE | JUIL. 2015 - JUIL. 2017

J'ai porté l'architecture frontend d'une plateforme B2B d'intelligence d'entreprise : architecture d'interface, fiabilité des builds et workflows appuyés par la recherche.

Technologies : Angular, RxJS, Python, Elasticsearch, Solr, concepts RDF/OWL, modèles de données orientés SPARQL, systèmes de build frontend, architecture d'interface d'entreprise, accessibilité WCAG/WAI-ARIA, SEO.

Technical Lead logiciel senior - lfeelgoods

PARIS, FRANCE | MAI 2014 - JUIN 2015

J'ai travaillé au sein de l'équipe front-end d'une plateforme SaaS B2B de récompenses, livrant des parcours utilisateurs, de l'outillage E2E et des intégrations avec de grandes marketplaces retail.

Technologies : AngularJS, JavaScript, Ruby on Rails, HTML/CSS, Jenkins, Protractor, outillage d'automatisation E2E, intégrations d'API de marketplaces retail, workflows CRM/récompenses.

Expériences antérieures

AOÛT 2007 - AVR. 2014

C'est là que j'ai appris à livrer, et à faire tourner des interfaces à 60fps quand ce n'était pas simple. Banque, e-commerce, édition, luxe, hôtellerie, medtech et campagnes pour Orange, Verseau Paris, Nemea Bank, eRepublik Labs, Yahoo, eMAG et Net Rom Business. Yahoo atteignait 8,2 M d'utilisateurs mensuels, eRepublik 6 M de joueurs dans 70 pays, et chez Nemea Bank j'ai aidé à bâtir une banque en ligne paneuropéenne et son paiement temps réel, en partant de zéro.

DISTINCTIONS & LANGUES

- Open source et écosystème : Splunk Visualization Toolbox depuis 2023 ; contributions antérieures à Angular UI Material et Benchpress.
- Auteur de questions du Java Challenge France 2019 ; mentor et jury du Hackathon IA/ML, Université Paris-Dauphine.
- Codility Golden Certificate (Omega, 2013) pour la résolution algorithmique de problèmes.

- Distinctions : Cannes Lions (sélection, 2010), Sabre European Awards (Or, 2010), Business Review Awards (Projet en ligne de l'année, 2010), Trophée d'Or IPRA (2009).
- Certifications et formation continue : diplômes UiPath RPA Academy - Solution Architect, Business Analyst, Developer Foundation, Infrastructure, Implementation Methodology, RPA Awareness (2018) ; AWS Solutions Architect (Professional & Associate), AWS Developer & SysOps Administrator (Associate) ; Oracle Database 12c SQL et conception de bases relationnelles ; Docker ; TOGAF / architecture d'entreprise ; ITIL 4 Foundation ; CISA ; sécurité cloud (2020-2021).

Langues : Courant en anglais, français et roumain.