

CONFIDENTIEL — USAGE INTERNE SEULEMENT

Coûts, Profits & Architecture

ticket911.ca — Plateforme AI Multi-Juridiction

Document: Interne Croige

Date: 7 février 2026

NE PAS PARTAGER AU CLIENT

Coûts Réels vs Prix Client — Développement

Ce que ça nous coûte réellement vs ce qu'on charge au client. Marge cible: 55-65%.

MODULE	NOTRE COÛT	PRIX CLIENT	PROFIT	MARGE
Portail Client Trilingue + Avocat ~120h dev @ ~—— coût réel	—	—	—	54%
AI Premium Multi-Juridiction ~140h dev + config LangChain/RAG	—	—	—	56%
Bases Jurisprudence AI (3 juridictions) ~80h scraping + indexation + pgvector	—	—	—	50%
Infrastructure & Hébergement Fort ~40h setup + achat 1er mois serveurs ~——	—	—	—	60%
Paiement Multi-Devises & Facturation ~50h Stripe + Lago + taxes	—	—	—	58%
Communications (3 marchés) ~40h Novu + Twilio config	—	—	—	60%
Documents, Signatures & Traduction ~40h DocuSeal + WeasyPrint + DeepL	—	—	—	60%
CRM & Rendez-vous ~40h Twenty CRM + Cal.com	—	—	—	60%
SEO Multi-Marché & Google Ads ~50h SEO + config + contenu initial	—	—	—	58%
Automatisation & Workflows ~40h n8n workflows	—	—	—	60%
Tests, Sécurité & Lancement ~30h QA + pen test + formation	—	—	—	63%
TOTAL DÉVELOPPEMENT	—	—	—	57%

— profit

Sur le développement — marge de 57%

Coûts Réels vs Prix Client — Mensuel Récurrent

C'est ici que le vrai argent se fait. Revenus récurrents chaque mois, indéfiniment.

Scénario: 650 dossiers/mois (objectif réaliste)

SERVICE	NOTRE COÛT	PRIX CLIENT	PROFIT	MARGE
Hébergement Dédié + GPU OVH Canada: principal ~— + GPU ~—	—	—	—	20%
Claude API (Anthropic) Opus: ~0.— × 650 = — Sonnet: ~0.— × 650 = — Haiku: ~0.— × 650 = —	—	—	inclus dans AI Premium	
OpenAI API (backup + embeddings) text-embedding-3: ~30— GPT-4 fallback: ~—	—	—	inclus dans AI Premium	
Mindee OCR API ~650 pages × 0.— = —	—	—	inclus dans AI Premium	
DeepL Pro API Traduction FR/EN/ES	—	—	inclus dans AI Premium	
= Total AI Premium (regroupé pour client)	—	—	—	-31%
Twilio SMS (3 marchés) ~3000 SMS × 0.— + 3 numéros × —	—	—	—	80%
Maintenance & Support ~8h/mois × — = —	—	—	—	73%
AI SEO Stratégie (3 marchés) ~10h/mois × — + outils ~—	—	—	—	67%
Gestion Google Ads (3 marchés) ~6h/mois × — (budget ads payé par client séparément)	—	—	—	63%
Évolutions & Fonctionnalités ~8h/mois × —	—	—	—	60%
TOTAL MENSUEL	—	—	—	46%

Note sur l'AI Premium: on perd — dessus

Les coûts AI réels (~— à 650 dossiers) dépassent ce qu'on charge (—). C'est **voulu**. Le client voit un prix bas pour l'AI, et on récupère largement sur les services à haute marge (maintenance 73%, SEO 67%, communications 80%). **Le profit global reste à 46%**. Si le volume monte beaucoup, on ajuste le prix client "AI Premium" dans le contrat avec clause de volume.

— de profit récurrent

= — de revenus récurrents nets — CHAQUE année

PROFIT TOTAL ANNÉE 1

— + — = —

Développement + 12 mois de récurrent

Scalabilité: Coûts AI Selon le Volume

Si le client grossit beaucoup, voici comment les coûts AI évoluent et comment on protège nos marges.

VOLUME/MOIS	COÛT CLAUDE API	COÛT OCR	COÛT TOTAL AI	PRIX CLIENT AI	ÉCART
100 dossiers	—	—	—	—	—
300 dossiers	—	—	—	—	—
650 dossiers	—	—	—	—	—
1 500 dossiers	—	—	—	—	—
5 000 dossiers	—	—	—	—	—

IMPORTANT: Clause de Volume dans le Contrat

Le prix "AI Premium —" couvre jusqu'à **500 dossiers/mois**. Au-delà, facturer un supplément de **1.— additionnel**. Ça nous coûte ~0.—, on fait 0.— de profit par dossier en plus. Mettre ça clairement dans le contrat.

Avec Clause de Volume Appliquée

VOLUME/MOIS	NOTRE COÛT AI	CLIENT PAIE AI	PROFIT AI	PROFIT TOTAL MENSUEL
100 dossiers	—	—	—	—
500 dossiers	—	—	—	—
650 dossiers (objectif)	—	500 + 150 = —	—	—
1 500 dossiers	—	500 + 1 000 = —	—	—
5 000 dossiers	—	500 + 4 500 = —	—	—

Résultat: On ne perd jamais d'argent

Avec la clause de volume, notre profit mensuel reste entre — et — peu importe le volume. Plus le client grossit, plus on gagne. Le modèle est durable et scalable.

Stratégie d'Optimisation des Coûts AI

- **Opus 4.6 seulement pour l'analyse juridique + plaidoyer** — les 2 tâches critiques qui justifient le prix premium
- **Sonnet 4.5 pour tout le reste** — scoring, devis, RAG, résumés — 5x moins cher qu'Opus, presque aussi bon
- **Haiku 4.5 pour les micro-tâches** — classification, chat FAQ, validation OCR — 30x moins cher qu'Opus
- **Cache de prompts** — Les system prompts juridiques (lois, procédures) sont cachés par Anthropic = -90% sur ces tokens
- **Ollama local pour le backup** — Si API down OU pour tâches simples en batch la nuit = coût —

Architecture Complète des Bases de Données

Chaque base qu'on construit et utilise, par catégorie.

A. Bases de Jurisprudence (Sources Externes → Notre Index)

1. CanLII — Canadian Legal Information Institute

QC

ON

Source: canlii.ca — API disponible

- Jurisprudence complète du Canada (QC + ON)
- Décisions des tribunaux provinciaux et fédéraux
- Textes de lois (Code sécurité routière QC, Highway Traffic Act ON)
- **On indexe:** toutes les décisions liées aux infractions routières dans PostgreSQL + pgvector
- **Coût:** Gratuit (API publique, respect rate limits)

2. SOQUIJ — Société québécoise d'information juridique

QC

Source: soquij.qc.ca — Abonnement requis

- Résumés analytiques de jurisprudence québécoise
- Annotations et commentaires par des juristes
- Recherche par type d'infraction, montant, résultat
- **On indexe:** les résumés pertinents pour enrichir les prompts AI
- **Coût:** ~200\$ — abonnement

3. LégisQuébec

QC

Source: legisquebec.gouv.qc.ca — Gratuit

- Code de la sécurité routière du Québec (complet, à jour)
- Règlements municipaux (Montréal, Québec, etc.)
- **On indexe:** chaque article de loi avec son contexte pour le RAG
- **Coût:** Gratuit

4. e-Laws Ontario

ON

Source: ontario.ca/laws — Gratuit

- Highway Traffic Act (HTA) complet
- Provincial Offences Act
- Règlements municipaux Toronto, Ottawa, etc.
- **On indexe:** articles HTA + amendes + points de démerite

- **Coût:** Gratuit

5. CourtListener + RECAP NY

Source: courtlistener.com — API open source (Free Law Project)

- Millions de décisions de tribunaux américains
- Filtrable par État de New York, par type d'infraction
- **On indexe:** décisions NY Traffic Violations Bureau (TVB) + tribunaux
- **Coût:** Gratuit (API publique, donation recommandée)

6. NY Vehicle & Traffic Law NY

Source: nysenate.gov/legislation/laws/VAT — Gratuit

- Toutes les lois NY sur les infractions routières
- Points système, amendes, surcharges
- Règles spécifiques NYC (parking, red light cameras)
- **On indexe:** chaque section avec montants et pénalités
- **Coût:** Gratuit

7. NYSCEF — NY State Courts Electronic Filing NY

Source: nycourts.gov/efile — Gratuit

- Dossiers électroniques des tribunaux NY
- Décisions récentes, tendances des juges
- **On indexe:** patterns par juge, par tribunal, par infraction
- **Coût:** Gratuit

B. Nos Bases de Données Internes (PostgreSQL)

Ce qu'on construit nous-mêmes dans PostgreSQL sur le serveur dédié:

8. DB Clients TOUTES

- **clients** — id, nom, email, téléphone, langue, juridiction, date inscription
- **adresses** — adresse, ville, province/état, code postal, pays
- **auth** — hash mot de passe, 2FA, sessions, tokens JWT
- **preferences** — langue, notifications SMS/email, fuseau horaire

9. DB Dossiers / Tickets TOUTES

- **tickets** — id, client_id, juridiction, type_infraction, montant_amende, date_infraction, lieu
- **ocr_results** — données extraites par OCR, image originale (ref MinIO), confiance
- **ai_analysis** — score contestation, raisons, jurisprudence trouvée, modèle utilisé, tokens consommés
- **dossiers** — id, ticket_id, statut (reçu/analyse/contestation/audience/résultat), avocat_notes
- **contestations** — brouillon AI, version finale avocat, date soumission, tribunal
- **audiences** — date, heure, tribunal, adresse, juge (si connu), résultat

10. DB Financière TOUTES

- **devis** — id, dossier_id, montant_cad, montant_usd, devise, statut, date_envoi
- **paiements** — id, stripe_id, montant, devise, méthode (carte/interac/ach), date, statut
- **factures** — id, dossier_id, montant, taxes (TPS/TVQ/HST/NY), PDF ref, envoyée
- **remboursements** — si contestation gagnée et remboursement applicable

11. DB Documents TOUTES

- **documents** — id, dossier_id, type (mandat/plaidoyer/facture/devis), url MinIO, signé
- **signatures** — id, document_id, client_id, date_signature, ip, hash_document
- Fichiers physiques stockés dans **MinIO** (S3-compatible), chiffrés AES-256

12. DB Jurisprudence Indexée (pgvector) TOUTES

- **jurisprudence_qc** — id, source (CanLII/SOQUIJ), titre, résumé, texte, embedding_vector, date, résultat
- **jurisprudence_on** — même structure, source Ontario
- **jurisprudence_ny** — même structure, source CourtListener/NYSCEF
- **lois_qc** — articles du Code sécurité routière, embedding_vector
- **lois_on** — articles Highway Traffic Act, embedding_vector
- **lois_ny** — sections Vehicle & Traffic Law, embedding_vector

- Recherche par **similarité sémantique** — l'AI trouve les précédents par sens, pas par mots-clés

13. DB Communications **TOUTES**

- **notifications** — id, client_id, canal (sms/email/push), template, statut (envoyé/lu), date
- **chat_messages** — id, dossier_id, sender (client/avocat/ai), message, date
- **audit_log** — id, user_id, action, ressource, date, ip — traçabilité complète

14. DB Analytics & Performance **TOUTES**

- **ai_usage** — modèle, tokens_in, tokens_out, coût, latence, date — suivi coûts AI en temps réel
- **stats_contestation** — par juridiction, par type infraction, taux succès, montant moyen
- **stats_revenus** — par mois, par juridiction, par source (SEO/Ads/direct)
- **stats_seo** — positions Google par mot-clé, par ville, évolution

C. Bases Complémentaires

15. Redis (Cache) **TOUTES**

- Cache des sessions utilisateurs (JWT)
- Cache des résultats AI fréquents (même type d'infraction = même analyse de base)
- Rate limiting par IP
- File d'attente Celery (tâches async: OCR, AI, emails)

16. Elasticsearch (Recherche) **TOUTES**

- Recherche full-text dans les dossiers, clients, jurisprudence
- Filtres par juridiction, date, statut, type d'infraction
- Autocomplete pour le portail avocat

TOTAL BASES DE DONNÉES

16 bases / index

7 sources juridiques externes + 7 bases internes + Redis + Elasticsearch

Résumé Exécutif — Chiffres Clés

Les chiffres qui comptent pour nous, en un coup d'oeil.

Développement

Notre coût réel	—
On charge au client	—
PROFIT développement	— (57%)

Récurrent Mensuel (à 650 dossiers/mois)

Notre coût mensuel réel	—
Client paie	—
PROFIT mensuel récurrent	— (46%)

Année 1 Complète

Profit développement	—
Profit récurrent (12 mois)	—
PROFIT TOTAL ANNÉE 1	—

Année 2+ (récurrent seulement)

PROFIT ANNUEL RÉCURRENT	— minimum
-------------------------	-----------

— — Année 1

Puis ~— de récurrent. Si le client grossit à 1 500+ dossiers/mois, encore plus avec la clause de volume.

Protection de nos Marges — Clauses Contrat

- **Clause volume AI:** — couvre jusqu'à 500 dossiers. Au-delà: +1.— additionnel
- **Clause SMS:** 3 000 SMS inclus/mois. Au-delà: +0.—
- **Révision annuelle:** Prix ajustables annuellement selon inflation et coûts API
- **Engagement minimum:** 12 mois pour le plan mensuel
- **Propriété du code:** On garde la propriété du code source. Le client a une licence d'utilisation

NE JAMAIS MONTRER CE DOCUMENT AU CLIENT

Ce document contient nos coûts réels, nos marges et notre stratégie de prix. Le client voit uniquement la proposition commerciale (proposition-ticket911.pdf).

Croige — Document Interne Confidentiel

7 février 2026