

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
École Supérieure en Informatique
- 08 mai 1945 - Sidi Bel Abbès



MÉMOIRE DE FIN D'ÉTUDES

Pour l'obtention du **diplôme de Master**

Filière : **Informatique**

Spécialité : **Intelligence Artificielle et Science des Données**

Thème :

**Système SaaS Agentique d'Assistance Juridique
avec Architecture Local-First**

Présentée par :

Boukharouba Souad

Soutenue en : **Juin 2026**

Devant le jury composé de :

Dr.	Président(e)
Dr.	Examineur(trice)
Dr. MEZRAR Samiha	Encadrante
Dr. BELMEKKI Ghizlene Amira	Encadrante

Année universitaire : **2025/2026**

Résumé

Les cabinets d’avocats belges font face à une tension structurelle entre l’adoption de l’intelligence artificielle et leurs obligations de confidentialité : le secret professionnel (article 458 du Code pénal belge) et le RGPD (UE 2016/679) interdisent l’envoi de données client vers des services cloud externes. Ce projet de fin d’études présente **Lexio**, un système SaaS agentique d’assistance juridique conçu selon une architecture local-first, où l’ensemble des modèles d’intelligence artificielle s’exécute localement sans aucune transmission de données vers l’extérieur.

Le système implémente cinq agents spécialisés orchestrés de façon asynchrone : un agent d’ingestion (détection de format, OCR Tesseract, classification documentaire à trois couches par heuristiques, patterns regex et fallback Mistral 7B), un agent d’extraction de faits (GLiNER pour la reconnaissance d’entités nommées, Mistral 7B pour l’extraction structurée de 24 types de faits juridiques, détection de contradictions), un agent de chronologie (classification d’événements, scoring de confiance HIGH/MEDIUM/LOW, déduplication à ± 3 jours, détection de conflits HARD/SOFT), un agent de recherche juridique combinant recherche dense (pgvector + nomic-embed-text, 768 dimensions), recherche lexicale (PostgreSQL FTS) et graphe de connaissances (Neo4j) au sein d’un pipeline RAG hybride trimodal, et un agent de rédaction générant des documents juridiques via des templates Jinja2 avec workflow de validation humaine obligatoire (DRAFT $\rightarrow$ REVIEWED $\rightarrow$ APPROVED $\rightarrow$ FINAL). Un sixième agent — simulateur de salle d’audience multi-agents — est documenté comme perspective de recherche.

L’évaluation empirique sur des documents représentatifs du droit belge démontre l’apport déterminant du graphe de connaissances : une comparaison de trois variantes du pipeline de recherche montre que les approches purement vectorielles échouent systématiquement sur les questions spécifiques à un dossier client, tandis que la variante hybride complète avec GraphRAG produit des réponses exactes et sourcées. L’évaluation couvre également la cascade de classification à trois couches — dont un test authentique forçant le déclenchement du fallback Mistral 7B sur un document conçu pour éviter les 35 patterns des couches 1 et 2 — le traitement de documents scannés par OCR (confiance Tesseract 0.92), la validation de 14 faits extraits (78% de confiance HIGH) et la génération de documents avec 9 variables auto-remplies sur 11. Un test

d'isolation inter-cabinet confirme empiriquement l'étanchéité des données entre tenants distincts. La conformité aux cinq obligations de supervision humaine de l'EU AI Act (Règlement UE 2024/1689) est vérifiée architecturalement dans le pipeline.

Mots-clés : SaaS juridique, Local-First, Multi-agents, RAG hybride, GraphRAG, Mistral 7B, GLiNER, Neo4j, pgvector, RGPD, EU AI Act, Supervision humaine, Droit belge.

Abstract

Belgian law firms face a structural tension between adopting artificial intelligence and meeting their confidentiality obligations : attorney-client privilege (Article 458 of the Belgian Penal Code) and the GDPR (EU 2016/679) prohibit sending client data to external cloud services. This final year project presents **Lexio**, an agentic legal assistance SaaS system designed under a local-first architecture, where all artificial intelligence models run entirely on-premises without any external data transmission.

The system implements five specialized agents orchestrated asynchronously : an ingestion agent (format detection, Tesseract OCR, three-layer document classification using heuristics, regex patterns, and Mistral 7B fallback), a fact extraction agent (GLiNER for named entity recognition, Mistral 7B for structured extraction of 24 legal fact types, contradiction detection), a timeline agent (event classification, HIGH/MEDIUM/LOW confidence scoring, ± 3 -day deduplication, HARD/SOFT conflict detection), a legal research agent combining dense retrieval (pgvector + nomic-embed-text, 768 dimensions), lexical search (PostgreSQL FTS), and a knowledge graph (Neo4j) within a trimodal hybrid RAG pipeline, and a drafting agent generating legal documents via Jinja2 templates with a mandatory human validation workflow (DRAFT $\rightarrow$ REVIEWED $\rightarrow$ APPROVED $\rightarrow$ FINAL). A sixth agent — a multi-agent courtroom simulator — is documented as a research perspective.

Empirical evaluation on documents representative of Belgian law demonstrates the decisive contribution of the knowledge graph : a comparison of three retrieval pipeline variants shows that purely vector-based approaches systematically fail on client-specific questions, while the full hybrid variant with GraphRAG produces accurate, sourced answers. The evaluation also covers the three-layer classification cascade — including an authentic test forcing Mistral 7B fallback on a document designed to avoid all 35 patterns in layers 1 and 2 — scanned document processing via OCR (Tesseract confidence 0.92), validation of 14 extracted facts (78% HIGH confidence), and document generation with 9 out of 11 variables automatically filled. A cross-tenant isolation test empirically confirms data segregation between distinct tenants. Compliance with the five EU AI Act (Regulation EU 2024/1689) human oversight obligations is verified architecturally throughout the pipeline.

Keywords : Legal SaaS, Local-First, Multi-agent Systems, Hybrid RAG, GraphRAG, Mistral 7B, GLiNER, Neo4j, pgvector, GDPR, EU AI Act, Human Oversight, Belgian Law.

يُعدّ الذكاء الاصطناعي فرصةً واعدةً لتحديث الممارسة القانونية، غير أنه يصطدم في السياق الأوروبي بعائق جوهري : تحظر قواعد السرية المهنية (المادة 458 من قانون العقوبات البلجيكي) واللائحة العامة لحماية البيانات (الاتحاد الأوروبي 2016/679) نقل بيانات الموكلين إلى أي خادِم خارجي. ومن هذا التناقض الجوهري انبثق مشروع التخرج هذا، الذي يُقدّم نظام **Lexio** : منصة SaaS قانونية متعددة المستأجرين، مبنية على مبدأ المعالجة المحلية الكاملة، (local-first) حيث تعمل جميع نماذج الذكاء الاصطناعي داخل البنية التحتية للمكتب دون أي إرسال للبيانات إلى الخارج. يتكون النظام من خمسة وكلاء ذكاء اصطناعي متخصصين تعمل بشكل غير متزامن ومتكامل : وكيل الاستيعاب المسؤول عن معالجة الوثائق الواردة (كشف الصيغة، التعرف الضوئي على الحروف، التصنيف التلقائي بثلاث طبقات)، ووكيل استخراج الحقائق القانونية (باستخدام نموذج GLiNER للتعرف على الكيانات و 7B Mistral لاستخراج الحقائق المنظمة مع كشف التناقضات)، ووكيل إدارة الجدول الزمني للدوسيه (تصنيف الأحداث، تقييم الثقة، رفع الازدواجية، كشف التعارضات)، ووكيل البحث القانوني المُدمج لثلاثة محركات بحث متكاملة (البحث الدلالي الكثيف عبر pgvector + nomic-embed-text والبحث المعجمي عبر PostgreSQL، FTS والبحث الرسم المعرفي عبر Neo4j) ووكيل الصياغة الذي يُولّد الوثائق القانونية انطلاقاً من قوالب Jinja2 ضمن مسار اعتماد إلزامي من المحامي (مسودة ← مراجعة ← اعتماد ← نسخة نهائية). أما الوكيل السادس، المتعلق بمحاكاة جلسات المرافعة، فيُوثّق كمنظور بحثي مستقبلي.

أثبت التقييم التجريبي على وثائق تمثيلية للقانون البلجيكي الدور المحوري لرسم المعرفة (GraphRAG) : فبينما تفشل المقاربات المعتمدة على البحث الدلالي وحده في الإجابة عن الأسئلة الخاصة بملف موكل بعينه، تُنتج نسخة النظام الهجينة الكاملة إجابات دقيقة مدعومة بالمصادر. كذلك شمل التقييم اختبار التصنيف بالطبقات الثلاث (بما فيه الطبقة الثالثة القائمة على Mistral، 7B) ومعالجة الوثائق الممسوحة ضوئياً (معامل ثقة 0.92)، والتحقق من 14 حقيقة مستخرجة (78% ذات ثقة عالية)، وتوليد الوثائق مع ملء 9 متغيرات من أصل 11 تلقائياً. وأكّد اختبار العزل بين مستأجرين مختلفين إحصاءات البيانات فعلياً وتجريبياً. كما جرى التحقق من الامتثال لمتطلبات الإشراف البشري المنصوص عليها في قانون الذكاء الاصطناعي الأوروبي

(اللائحة 2024/1689) عبر خمس بوابات تحقق معمارية داخل خط الأنابيب.

الكلمات المفتاحية : منصة SaaS قانونية، معالجة محلية كاملة، وكلاء ذكاء اصطناعي متخصصون، RAG هجين ثلاثي الوسائط، رسم المعرفة القانونية، Mistral، 7B، GLiNER، Neo4j، pgvector، حماية البيانات الشخصية، قانون الذكاء الاصطناعي الأوروبي، إشراف بشري إلزامي، القانون البلجيكي.