



COUR SUPÉRIEURE DU QUÉBEC

RAPPORT TECHNIQUE

Le fonctionnement des agents d'intelligence artificielle
dans le cadre du projet pilote de la Cour supérieure du
Québec

Rapport technique • septembre 2026

Document complémentaire

Rapport d'évaluation du projet pilote d'intelligence artificielle de la Cour supérieure du Québec (avril 2026)



AVERTISSEMENT

Le présent rapport est un document technique. Il décrit la configuration et le fonctionnement des agents d'intelligence artificielle de la Cour supérieure du Québec au terme de la première phase de son projet pilote. Il ne constitue ni un avis juridique, ni une directive, ni une position de la Cour sur une question susceptible de lui être soumise.

VERSION

Version 1.0, septembre 2026. Toute version ultérieure remplacera celle-ci.

MOT DE LA JUGE EN CHEF

La Cour supérieure du Québec a confié à des agents conversationnels d'intelligence artificielle une place mesurée dans le travail quotidien de ses juges. Elle l'a fait avec prudence, en commençant par un projet pilote encadré et évalué, dont les résultats ont été rendus publics. Le présent rapport technique complète cette évaluation en exposant ce qui se trouve sous l'outil : la plateforme retenue, les instructions données aux agents, le paramétrage choisi, la préparation des sources documentaires et les limites reconnues.

La Cour publie ce document pour deux raisons.

La première tient à la confiance du public. Une institution judiciaire qui introduit l'intelligence artificielle dans ses méthodes de travail doit être en mesure d'expliquer ce que ces outils font, ce qu'ils ne font pas et suivant quelles règles ils fonctionnent. Cette explication ne peut demeurer générale : elle doit être suffisamment précise pour être vérifiée. Un principe gouverne l'ensemble de la démarche et mérite d'être répété ici. Aucun agent de la Cour ne rend de décision, n'en rédige ni n'y contribue de manière autonome. La décision judiciaire, sa motivation et sa rédaction demeurent l'œuvre exclusive du juge.

La seconde raison tient à nos institutions sœurs. Plusieurs tribunaux, au Québec, au Canada et ailleurs, se posent aujourd'hui les questions que la Cour s'est posées en 2025, et peu de réponses documentées circulent. La Cour a donc choisi d'exposer ses choix techniques avec un degré de détail élevé, y compris lorsque ces choix ont été contraints plutôt que délibérés. Un tribunal qui souhaiterait reprendre la démarche doit pouvoir juger par lui-même de ce qui mérite d'être repris et de ce qui doit être écarté.

Ce rapport s'inscrit dans une démarche plus large. Les [lignes directrices sur l'utilisation de l'intelligence artificielle générative par les juges](#), signées le 31 août 2026 par la Cour d'appel du Québec, la Cour supérieure du Québec, la Cour du Québec et les cours municipales, en posent le cadre. Le présent document en illustre la mise en œuvre concrète.

Je remercie les vingt-deux juges qui ont accepté de participer au projet pilote et d'en documenter l'usage, ainsi que M^e Guillaume Bourgeois, adjoint exécutif juridique, qui a conçu les agents, dirigé le projet et rédigé ce rapport.

Marie-Anne Paquette

Juge en chef de la Cour supérieure du Québec

TABLE DES MATIÈRES

MOT DE LA JUGE EN CHEF	3
1. OBJET DU RAPPORT	5
2. LA PLATEFORME ET LES PRÉALABLES	5
3. LES COMPOSANTES D'UN AGENT	7
3.1 L'ancrage documentaire	7
4. LES AGENTS EN SERVICE	8
4.1 Les lois intelligentes	9
4.2 Le Livre bleu 2.0	9
4.3 Le Citateur	10
4.4 L'assistant rédacteur	10
4.5 Le Traducteur	10
4.6 Le Technicien informatique IA	11
5. LA RÉDACTION DES INSTRUCTIONS	11
6. LE PARAMÉTRAGE DES AGENTS	12
7. LA PRÉPARATION DES SOURCES DOCUMENTAIRES	13
7.1 Intégrer les notes de bas de page au corps du texte	14
7.2 Écarter les éléments qui nuisent à la recherche	14
7.3 Découper le texte en unités autoportantes	14
7.4 Rendre chaque fragment intelligible isolément	15
7.5 Renvoyer aux pages du document d'origine	15
7.6 Nommer, dater et tenir à jour	16
8. MISE À L'ESSAI, MESURE ET SURVEILLANCE	16
9. FEUILLE DE ROUTE À L'INTENTION DES TRIBUNAUX	17
10. LIMITES CONNUES ET ÉVOLUTIONS	18
11. CONCLUSION	20
ANNEXE A. LA PRÉPARATION DU LIVRE BLEU	21
A.1 Le document de départ	21
A.2 Le déroulement du prétraitement	21
A.3 L'alignement sur la pagination officielle	22
A.4 Les contrôles de qualité	22
A.5 Le résultat et sa mise à jour	22
A.6 Transposition à un autre document	23

1. OBJET DU RAPPORT

Ce rapport décrit la conception et le fonctionnement des agents conversationnels d'intelligence artificielle mis à la disposition des juges de la Cour supérieure du Québec dans le cadre de son projet pilote. Mené du 8 décembre 2025 au 16 mars 2026, ce projet a réuni 22 juges volontaires. Le présent document complète le [Rapport d'évaluation du projet pilote](#), publié en avril 2026, qui présente la démarche, les résultats et les recommandations.

Le rapport s'adresse à deux publics. Les justiciables, les avocats et les observateurs y trouveront une explication de ce que font les agents et, surtout, de ce qu'ils ne font pas, qu'il s'agisse de la plateforme utilisée, de la façon dont les réponses sont produites, des règles imposées aux agents ou de leurs limites. Ces éléments sont présentés aux sections 1 à 4, 10 et 11. Les sections 5 à 9 et l'annexe A, plus techniques, s'adressent aux tribunaux et aux équipes qui envisagent une démarche semblable. La Cour y documente son paramétrage et la préparation de ses sources avec suffisamment de détails pour qu'une autre institution puisse s'en inspirer, les adapter ou les remettre en question. La section 9 propose en outre une feuille de route.

Il importe de préciser d'emblée qu'aucun agent ne rend de décision, n'en rédige ni n'y contribue de manière autonome. Les agents servent à la consultation documentaire, à la mise en forme et au soutien technique. La décision judiciaire, sa motivation et sa rédaction relèvent exclusivement du juge.

Le rapport décrit la configuration retenue pour la première phase du projet, amorcée en 2025. Certains choix s'expliquent par les capacités que la plateforme et les modèles permettaient à ce moment, et il faut noter que d'autres options existent aujourd'hui. La plupart demeurent néanmoins valables, parce qu'ils découlent d'exigences qui n'ont pas changé, notamment la sécurité des renseignements, l'indépendance judiciaire et la nécessité pour l'organisation de pouvoir entretenir elle-même ses outils. Ils rendent compte de l'expérience d'une cour ; il appartiendra à chaque tribunal de les adapter à sa propre réalité.

2. LA PLATEFORME ET LES PRÉALABLES

Les agents ont été construits dans Microsoft Copilot Studio. Les juges y accèdent par Microsoft Teams, dans l'environnement Microsoft 365 du ministère de la Justice du Québec.

La Cour n'a ni lancé d'appel d'offres ni comparé les solutions offertes sur le marché. Copilot Studio pouvait déjà être déployé dans l'environnement bureautique



en place, ce qui réglait d'emblée plusieurs questions. Les données demeuraient dans l'environnement gouvernemental, aucune nouvelle entente de traitement de renseignements n'était requise et aucun processus d'acquisition n'avait à être engagé. La plateforme a en outre permis de construire les neuf agents sans programmation et sans acquisition d'infrastructure.

Un tribunal dont l'environnement technologique est différent devra vérifier que la solution envisagée permet de créer des agents sans programmation, de limiter leurs réponses aux documents de référence, de choisir parmi des modèles approuvés par l'administrateur de l'organisation et d'héberger le tout dans son propre locataire infonuagique.

L'hébergement des agents dans l'environnement Microsoft 365 du ministère emporte trois conséquences. Les conversations et les documents traités demeurent dans l'environnement infonuagique sécurisé où se trouvent les services de la Cour. Les échanges d'un juge avec un agent constituent des renseignements de la magistrature, et non des données administratives de l'organisation ; ils échappent donc aux consultations et aux extractions administratives, comme il est précisé plus loin. Par ailleurs, les conditions de service de Microsoft 365 pour les clients organisationnels, et l'addenda sur la protection des données qui s'y rattache, prévoient que le contenu des conversations ne sert pas à entraîner les modèles. Enfin, les agents ne sont accessibles qu'aux utilisateurs autorisés du réseau et ne sont pas offerts au public.

Pour suivre la même voie, un tribunal aura besoin d'un environnement Microsoft 365 institutionnel administré par une autorité informatique compétente, et des licences qui donnent accès à Copilot Studio. L'utilisation de la plateforme se mesure en crédits, dont l'administrateur suit la consommation. À titre indicatif, le projet pilote a consommé environ 2 500 crédits en quatorze semaines pour 22 utilisateurs, selon la tarification et le mode de calcul en vigueur au premier trimestre de 2026, que le fournisseur peut modifier. La collaboration de l'administrateur du système est aussi nécessaire, notamment pour approuver les modèles offerts et partager les agents avec les groupes d'utilisateurs visés. Il faut enfin pouvoir compter sur une personne-ressource à l'interne pour configurer les agents et en assurer l'entretien. Le Rapport d'évaluation en fait une condition essentielle de pérennité, et l'expérience de la Cour le confirme, puisque les agents demandent des ajustements constants.

Chaque agent fonctionne à l'aide d'un grand modèle de langage. Pendant le projet, les agents ont d'abord utilisé GPT-4, de la société OpenAI, puis GPT-4.1 lorsque la plateforme l'a offert. Ce changement de modèle n'a pas obligé à reconstruire les agents, puisque les instructions, les bases de connaissances et les paramètres sont restés les mêmes. Au moment de choisir une plateforme, il importe donc

moins de viser un modèle en particulier que de s'assurer que l'on pourra en changer, dans un cadre approuvé par l'administrateur.

La confidentialité des échanges dépasse la seule question des contrôles d'accès, puisqu'elle touche à l'indépendance judiciaire. Pendant le projet pilote, les conversations d'un juge avec un agent sont conservées, mais elles ne sont accessibles ni aux autres utilisateurs, ni aux parties, ni au personnel administratif. Aux fins de gestion, l'organisation n'a accès qu'à des mesures d'utilisation agrégées, qui ne révèlent rien du contenu des échanges.

Un juge peut interroger un agent sur dans le cadre d'une affaire en délibéré. La trace de cette recherche révélerait sa démarche intellectuelle, qui doit rester à l'abri de tout regard extérieur. C'est pourquoi la Cour considère cette question comme un préalable. Avant tout déploiement, un tribunal devrait demander à son administrateur informatique de lui confirmer par écrit quels journaux d'utilisation la plateforme produit, qui peut les consulter et à quelles conditions une extraction peut être demandée. À défaut d'un tel engagement, les garanties décrites dans le présent rapport ne reposent que sur la configuration du moment.

3. LES COMPOSANTES D'UN AGENT

Un agent ne se programme pas comme un logiciel traditionnel. Il résulte de l'assemblage de quatre éléments que l'institution configure et contrôle.

Le premier est le modèle de langage, décrit à la section précédente. Les instructions constituent le deuxième. Il s'agit d'un texte rédigé par la Cour qui fixe le rôle de l'agent, ses règles de travail et ce qui lui est interdit, et qui s'applique à chaque échange (section 5). La base de connaissances regroupe les documents de référence versés dans l'agent, par exemple le texte d'une loi, un ouvrage de référence, un guide de citation ou un lexique (section 7). Les paramètres de la plateforme règlent notamment l'ancrage documentaire, la modération du contenu, l'accès au web et le traitement des fichiers (section 6).

3.1 L'ancrage documentaire

Les agents documentaires de la Cour fonctionnent selon la technique de la génération augmentée par la recherche (retrieval-augmented generation, ou RAG). Au moment où un document est versé dans la base de connaissances, la plateforme le découpe en fragments et les indexe selon leur sens. Lorsqu'un juge pose une question, elle cherche dans cet index les fragments les plus pertinents, même si les mots utilisés ne sont pas les mêmes, et les transmet au modèle avec la question. Les instructions obligent ensuite le modèle à bâtir sa réponse à partir de ces

fragments, en les citant au besoin, et non à partir de ce qu'il a appris lors de son entraînement.

Cet ancrage réduit considérablement le risque le plus connu des modèles de langage, celui de produire un contenu plausible mais inexact, ce qu'on appelle couramment une hallucination. Le risque ne disparaît pas pour autant, car le modèle peut mal rendre un fragment ou en combiner deux de façon trompeuse. Surtout, la qualité de la réponse dépend de celle de la recherche qui la précède. Si le bon fragment n'est pas trouvé, parce que le document est mal structuré, qu'une note importante est coupée du texte qu'elle annote ou que la question est formulée dans des termes trop éloignés de ceux du texte, le modèle répondra à partir d'une base incomplète. C'est ce qui explique l'importance accordée à la préparation des sources, décrite à la section 7. L'expérience de la Cour montre d'ailleurs que ce travail a davantage contribué à la qualité des réponses que les réglages de la plateforme.

Pour les agents qui doivent restituer fidèlement un texte, soit les lois intelligentes et le Livre bleu, la Cour a désactivé les réponses non ancrées, c'est-à-dire celles qui ne s'appuient pas sur les documents de référence. Lorsque la réponse ne se trouve pas dans ses sources, l'agent doit donc le dire.

4. LES AGENTS EN SERVICE

Neuf agents ont été mis en service pendant le projet pilote. Quatre d'entre eux sont des « lois intelligentes » qui partagent le même paramétrage (Code civil du Québec, Code de procédure civile, Code criminel, Loi sur la faillite et l'insolvabilité). Les cinq autres remplissent chacun une fonction particulière. Le tableau suivant les présente sommairement.

Agent	Rôle	Sources de connaissances	Accès au web
Lois intelligentes (4 agents : C.c.Q., C.p.c., C.cr., LFI)	Consultation des lois et codes en langage naturel, avec citation exacte des articles	Texte officiel de la loi visée (un agent par loi)	Aucun
Livre bleu 2.0	Recherche en droit de la famille dans un ouvrage de référence interne	Ouvrage de référence en droit de la famille, notes intégrées	Aucun
Le Citateur	Correction des références et citations juridiques selon le guide de la Cour	Aide-mémoire sur la citation des références juridiques	Contrôlé

Agent	Rôle	Sources de connaissances	Accès au web
L'assistant rédacteur	Amélioration de la lisibilité et de la structure de textes rédigés par les juges	Aucune (travaille sur le texte soumis)	Aucun
Le Traducteur	Traduction juridique français-anglais à équivalence fonctionnelle	Lexique bilingue québécois ; banque de décisions bilingues	Aucun
Technicien informatique IA	Soutien informatique de premier niveau (poste de travail et suite Microsoft 365)	Guides techniques du poste de travail et de la suite bureautique	Aucun

4.1 Les lois intelligentes

Chaque loi intelligente donne accès à une loi ou à un code, que l'utilisateur peut interroger en langage naturel. Sa base de connaissances se limite au texte de la loi visée. Les instructions exigent que l'agent cite les dispositions pertinentes intégralement et textuellement, en commençant par le numéro d'article et en respectant la structure du texte. Il ne peut pas les résumer. Il lui est aussi interdit, comme à tous les agents juridiques de la Cour, de donner des conseils juridiques qui exigeraient une consultation individuelle. L'interdiction peut surprendre, s'agissant d'un outil réservé aux juges. Elle vise à empêcher l'agent de se prononcer sur une situation concrète, par exemple en indiquant ou en suggérant comment trancher un litige décrit dans ses faits, et à le maintenir dans son rôle, qui est de restituer le texte applicable. Cette formule est la plus simple à reproduire, et c'est celle que la Cour recommande comme point de départ.

4.2 Le Livre bleu 2.0

Le Livre bleu est un ouvrage de référence interne en droit de la famille. L'agent qui porte son nom obéit aux mêmes règles de citation que les lois intelligentes. Il reproduit les extraits pertinents tels quels, avec leur référence, pour que le juge lise le texte de l'ouvrage et non une paraphrase. Les notes de bas de page ont été intégrées au texte lors de la préparation de la source (section 7), de sorte que la recherche les retrouve avec les passages auxquels elles se rapportent.

4.3 Le Citateur

Le Citateur corrige les références juridiques (jurisprudence, doctrine, lois et règlements) selon le guide de citation versé dans sa base de connaissances. Ses instructions précisent les règles typographiques à suivre, comme l'italique pour les noms des parties dans une référence jurisprudentielle, sauf pour les séparateurs. Elles lui interdisent d'inventer un élément manquant, le prénom d'un auteur par exemple, et l'obligent à signaler toute référence dont il ne peut confirmer l'existence. C'est le seul agent qui a accès à la recherche web, et cet accès est limité à la vérification des références.

4.4 L'assistant rédacteur

L'assistant rédacteur aide les juges à rendre leurs textes plus clairs et mieux structurés. Il peut proposer de scinder un paragraphe trop long, d'uniformiser la terminologie, d'améliorer les transitions entre les citations et le texte ou d'ajuster le ton. Il ne travaille que sur le texte que le juge lui soumet. C'est l'agent qui a été le plus utilisé pendant le projet. Ses instructions prévoient notamment ce qui suit :

Instruction de l'assistant rédacteur

Tu ne dois jamais accepter de rédiger un jugement ou de générer un jugement, à partir d'instructions ou de documents.

Il ne peut pas non plus ajouter de références que l'utilisateur n'a pas fournies, ce qui écarte le risque de références inventées. Son rôle s'apparente ainsi à celui d'un réviseur, qui intervient sur un texte conçu et rédigé par le juge ; c'est le juge qui arrête la version finale.

Il faut toutefois reconnaître que l'agent le plus utilisé est aussi le moins protégé sur le plan technique. Comme il n'a pas de base de connaissances, il doit pouvoir produire des réponses non ancrées, et ses interdictions ne reposent donc que sur ses instructions. Or une instruction influence fortement le comportement du modèle, mais ne le garantit pas. La Cour en a tenu compte. Le banc d'essai de cet agent comprend davantage de requêtes conçues pour contourner ses consignes, les commentaires qui le concernent sont examinés en priorité et la vérification par le juge s'y applique avec la même rigueur que pour les agents documentaires.

4.5 Le Traducteur

Le Traducteur traduit des textes juridiques du français vers l'anglais et inversement, en fonction du droit applicable au Québec et au Canada. Ses instructions lui demandent de rechercher l'équivalence fonctionnelle, c'est-à-dire de rendre le

sens et l'effet juridique du texte dans la langue qu'emploient les tribunaux canadiens, et d'éviter la traduction littérale. Il s'appuie sur un lexique bilingue des termes juridiques en usage au Québec et sur une banque de formulations tirées de décisions publiées dans les deux langues. Quand une expression consacrée s'y trouve, il doit l'utiliser telle quelle. La traduction qu'il produit reste une ébauche, que l'utilisateur doit réviser.

4.6 Le Technicien informatique IA

Le Technicien informatique IA répond aux questions simples des juges sur leur poste de travail et la suite Microsoft 365, à partir de guides techniques versés dans sa base de connaissances. Son architecture est la même que celle des autres agents, pour un usage plus modeste, le dépannage de premier niveau. Ses instructions lui demandent d'expliquer sans jargon et de procéder étape par étape.

5. LA RÉDACTION DES INSTRUCTIONS

Ce sont les instructions qui font d'un modèle généraliste un outil adapté à un usage précis. Au fil du projet, la Cour a dégagé une structure type, que les instructions de ses agents suivent à des degrés divers et qu'un autre tribunal peut reprendre.

Cette structure compte cinq volets. Le premier définit le rôle de l'agent en une ou deux phrases, en précisant son domaine et ses utilisateurs (« Tu es un traducteur juridique bilingue spécialisé en droit applicable au Québec et au Canada... »). Le deuxième désigne les sources que l'agent peut utiliser parmi les documents de sa base de connaissances, et la manière de s'en servir. Viennent ensuite les règles de travail, formulées comme des consignes concrètes et vérifiables : citer l'article en indiquant d'abord son numéro, conserver la structure du texte, mettre les noms des parties en italique. Les interdictions forment un volet distinct. Elles sont rédigées de façon absolue, puisque ce sont elles qui doivent tenir quand la conversation s'éloigne de l'usage prévu (ne jamais rédiger de jugement, ne jamais inventer de référence, ne pas donner de conseil juridique individuel). Le dernier volet indique à l'agent quoi faire lorsqu'il ne trouve pas la réponse dans ses sources. Il doit alors le dire, sans compléter de mémoire.

Quelques leçons se sont dégagées de la mise au point. Une interdiction importante doit être isolée et formulée sans nuance. Noyée parmi des préférences de style, elle a moins de poids qu'une phrase distincte, au besoin en majuscules. Les instructions doivent aussi être mises à l'épreuve. La Cour a soumis à ses agents des requêtes conçues pour les faire dévier, par exemple une demande de rédaction de jugement présentée comme un simple exemple ou une demande de conseil juridique déguisée en hypothèse, et a resserré les instructions jusqu'à ce que les

refus soient systématiques. Il faut enfin dater et archiver chaque version des instructions. C'est la seule manière d'expliquer, après coup, pourquoi un agent a répondu comme il l'a fait à un moment donné.

6. LE PARAMÉTRAGE DES AGENTS

Copilot Studio offre, pour chaque agent, une série de réglages qui déterminent ce qu'il peut faire. Le tableau suivant les présente, avec la position retenue par la Cour. Un tribunal qui configure ses propres agents aura à se prononcer sur chacun d'eux.

Paramètre	Ce qu'il contrôle	Position retenue par la Cour
Orchestration générale	Mode de fonctionnement : réponses composées dynamiquement à partir des connaissances, ou dialogue scénarisé par rubriques prédéfinies	Activée pour tous les agents : la valeur ajoutée vient de la composition dynamique ancrée dans les sources
Réponses non ancrées	Autorise ou interdit les réponses composées à partir des connaissances générales du modèle, sans passer par les sources versées	Interdites pour les agents documentaires (lois intelligentes, Livre bleu) ; autorisées pour les agents qui travaillent sur le texte soumis par l'utilisateur (rédacteur, traducteur, technicien)
Recherche sur le web	Permet à l'agent de consulter des sites web publics pour compléter ses réponses	Désactivée partout, sauf accès contrôlé du Citateur pour la vérification de l'existence des références
Chargements de fichiers	Permet aux utilisateurs de soumettre des documents (limite de 15 Mo par fichier) dans la conversation	Activés pour les agents qui traitent le texte de l'utilisateur ; désactivés pour les agents purement documentaires, avec consigne de ne pas y déposer de pièces de dossiers
Interpréteur de code	Autorise l'agent à générer et exécuter du code pour traiter des fichiers ou produire des rapports	Désactivé pour tous les agents : inutile à leur mission et source de risque
Assistants connectés	Permet à d'autres agents de se connecter à celui-ci pour former des chaînes de travail	Fermé par défaut ; ouvert uniquement lorsqu'un flux inter-agents est défini et documenté
Modération du contenu	Niveau du filtre appliqué aux réponses (Basse / Modérée / Élevée)	Réglée agent par agent, après essais sur des contenus judiciaires sensibles ; le niveau et sa justification sont consignés dans un registre interne

Paramètre	Ce qu'il contrôle	Position retenue par la Cour
Recherche sémantique	Indexation sémantique des sources pour un appariement question-passages plus précis	Activée pour tous les agents dotés de sources de connaissances
Rétroaction des utilisateurs	Recueil des évaluations (positives ou négatives) et des commentaires sur chaque réponse	Activée partout ; les commentaires sont transmis à l'organisation et revus périodiquement
Outils, déclencheurs et connecteurs	Raccordement de l'agent à des systèmes externes, des automatisms ou des sources dynamiques	Aucun raccordement : les agents sont volontairement clos sur leurs bases documentaires

Deux de ces réglages méritent quelques explications. Le premier, et le plus important, concerne les réponses non ancrées. C'est lui qui distingue les agents documentaires, qui ne doivent jamais sortir de leurs sources, des agents qui travaillent sur ce que l'utilisateur leur soumet et qui ont besoin des capacités générales du modèle. Il faut trancher cette question agent par agent. À défaut, un agent documentaire risque de compléter ses sources par ses connaissances générales sans le signaler, avec toute l'autorité que lui confère sa spécialisation.

Le second concerne la modération du contenu. Les filtres des plateformes commerciales sont pensés pour un usage général. Or plusieurs matières judiciaires, comme la violence conjugale, les agressions ou l'exploitation, tombent précisément dans les catégories qu'ils surveillent. Un niveau de modération trop élevé peut bloquer des réponses légitimes dans les dossiers mêmes où l'outil serait le plus utile. La Cour recommande donc de tester chaque agent avec des contenus sensibles représentatifs de sa matière avant de fixer le niveau, puis de consigner ce choix. Il faut savoir aussi que la modération n'est pas offerte dans toutes les langues. Pendant le projet pilote, elle ne l'était pas pour les agents dont la langue principale était le français. Les instructions et la formation des utilisateurs doivent alors prendre le relais.

7. LA PRÉPARATION DES SOURCES DOCUMENTAIRES

Un document long et complexe versé tel quel dans un agent donne des résultats médiocres. La préparation des sources repose sur une idée simple : chaque fragment que la plateforme peut retrouver doit pouvoir être compris seul. Un fragment qui, lu isolément, ne dit pas de quel texte il vient, à quelle disposition il se rattache ou sur quelle autorité il s'appuie produira une réponse incomplète, même si la recherche a bien fonctionné.



La Cour soumet donc ses sources à un prétraitement en six opérations, effectué au moyen d'un outil développé à l'interne. Elles sont décrites ci-dessous avec suffisamment de détails pour être reprises avec d'autres documents. Deux types de sources sont en cause. Les textes législatifs ont une structure régulière et une numérotation stable. Les ouvrages de référence internes sont plus volumineux, plus hiérarchisés et abondamment annotés. L'annexe A décrit en détail le traitement d'un ouvrage de ce second type.

7.1 Intégrer les notes de bas de page au corps du texte

Les notes d'un texte législatif ou d'un ouvrage de doctrine contiennent souvent de l'information essentielle, comme l'historique d'une disposition, la source d'une affirmation ou des renvois à la jurisprudence. Dans un fichier Word, elles sont toutefois enregistrées à part du corps du texte. À l'indexation, elles se retrouvent donc dans d'autres fragments que les passages qu'elles annotent, et la recherche ne les réunit plus. Le prétraitement règle le problème en insérant chaque note dans le texte, à l'endroit de son appel, sous une forme balisée comme « [Note 12 : 2016, c. 30, a. 15] ». Le modèle peut ainsi distinguer la disposition de son annotation et citer l'une sans la confondre avec l'autre.

Les notes des ouvrages de doctrine posent une difficulté supplémentaire. Une note qui se limite à « Id. » ou à « Ibid. » perd son sens une fois isolée, puisque la note à laquelle elle renvoie peut se trouver dans un autre fragment. Ces abréviations sont donc remplacées par la référence complète avant l'insertion. Cette opération est effectuée dans le fichier Word, avant le découpage.

7.2 Écarter les éléments qui nuisent à la recherche

La table des matières est utile au lecteur, mais elle nuit à la recherche. Ses centaines de lignes, composées d'un intitulé et d'un numéro de page, reprennent justement le vocabulaire des questions des utilisateurs. Le moteur de recherche les fait donc remonter en tête des résultats, au détriment de fragments utiles, alors qu'elles ne contiennent aucune réponse. Le prétraitement retire la table des matières et les pages liminaires sans contenu normatif. Elles sont remplacées par un fichier d'index unique, qui décrit la composition du corpus et sert de plan à l'agent.

7.3 Découper le texte en unités autoportantes

Un document de plusieurs centaines de pages n'est pas versé en un seul fichier. Il est découpé selon une méthode qui varie en fonction de sa nature.



Les textes législatifs sont découpés en blocs d'une vingtaine de pages. Comme un fichier Word n'a pas de pagination fiable, la taille est estimée en mots, à raison d'environ 450 mots par page pour un texte législatif dense, ce qui donne des blocs d'environ 9 000 mots. Une fois cette taille atteinte, le bloc se termine au titre, au chapitre ou à la section qui suit. Aucun bloc ne commence ainsi au milieu d'une disposition, et aucun titre ne se retrouve seul à la fin d'un bloc.

Pour les ouvrages de référence, c'est d'abord la structure qui guide le découpage. Le découpage part des chapitres. Un chapitre qui tient sous le seuil retenu devient un fichier ; s'il le dépasse, il est remplacé par ses sous-sections, et l'opération est répétée jusqu'à ce que chaque unité passe sous le seuil. Les sections trop courtes pour former un fichier utile sont ensuite regroupées avec leurs voisines de même niveau. Le corpus reproduit ainsi la structure de l'ouvrage, à raison d'un fichier par sujet, et aucun développement n'est coupé en son milieu.

Dans un cas comme dans l'autre, les notes sont intégrées avant le découpage. Autrement, les appels de note d'un bloc pourraient renvoyer à des notes restées dans un autre.

7.4 Rendre chaque fragment intelligible isolément

Le découpage présente toutefois un inconvénient, puisqu'il sépare les passages des titres qui les situent. Un développement sur la date d'évaluation, par exemple, ne dit plus à lui seul de quel régime matrimonial il est question. Pour y remédier, chaque fichier commence par un en-tête qui indique le document source, sa date de mise à jour, le rang du fichier dans le corpus et le chemin complet de la section dans l'ouvrage (« Patrimoine familial › Partage du patrimoine familial › Valeur nette du patrimoine (art. 417 C.c.Q.) »). Dans le fichier, une ligne rappelant ce chemin suit chaque sous-titre, pour qu'un fragment tiré du milieu du fichier reste situé.

Les fichiers sont nommés d'après le même chemin. Ce nom a une incidence réelle. Il influe sur le classement des résultats, et c'est lui que l'agent cite comme source, puisque la plateforme attribue les extraits au fichier, et non à la page, pour les formats autres que le PDF.

7.5 Renvoyer aux pages du document d'origine

Un fichier Word ne contient pas de pagination. Celle-ci est calculée à l'affichage et varie selon la police, la version du logiciel et même l'imprimante configurée. Le fichier Word ne permet donc pas, à lui seul, de produire des renvois de page. Or le juge a besoin de ce renvoi pour vérifier le passage dans l'ouvrage, et cette vérification est une règle pour tous les agents de la Cour.

Les renvois de page sont donc obtenus en comparant le texte à la version PDF de l'ouvrage, qui fait foi de la pagination. Les deux textes sont d'abord uniformisés (minuscules, sans accents, chiffres ni ponctuation), ce qui élimine du même coup les appels de note que le PDF colle aux mots. Ils sont ensuite parcourus en parallèle afin de retrouver chaque paragraphe dans le PDF, et la page où il commence est inscrite dans le corpus sous la forme « [p. 458] ». L'agent doit citer cette marque avec chaque extrait.

La méthode a ses limites. La marque renvoie à la page où commence le paragraphe ; un passage qui s'étend sur deux pages est donc rattaché à la première. Plus important encore, les renvois ne valent que pour une version précise de l'ouvrage. À chaque mise à jour, la pagination change et l'alignement doit être refait avec le nouveau PDF. À défaut, l'agent renverrait à des pages qui ne correspondent plus au texte, ce qui serait plus préjudiciable que l'absence de renvoi.

7.6 Nommer, dater et tenir à jour

Parce qu'il intervient dans la recherche, le nom de chaque fichier doit décrire son contenu, soit l'identifiant du texte, la portion couverte et la date de version, par exemple « C-25.01, art. 1 à 160, v2026-07 ». Un registre interne indique, pour chaque agent, les sources versées, leur version et la date de leur versement. Les textes législatifs sont revérifiés chaque fois qu'une modification touchant les lois visées entre en vigueur, et les instructions peuvent demander à l'agent d'indiquer la version du texte au début de la conversation.

Lorsqu'une source est mise à jour, les fichiers de l'ancienne version sont retirés de l'agent avant l'ajout des nouveaux. Deux versions indexées en même temps donnent des réponses contradictoires, sans que l'utilisateur en soit averti. Le registre et le calendrier de mise à jour font partie de l'outil au même titre que les agents, et leur prise en charge doit être prévue dès le départ.

8. MISE À L'ESSAI, MESURE ET SURVEILLANCE

Quelques essais réussis ne suffisent pas pour mettre un agent en service. La Cour procède en trois temps, avec les outils qu'offre la plateforme.

Avant la mise en service, chaque agent est soumis à un banc d'essai, soit une série de questions types couvrant sa mission, avec les réponses attendues. La série comprend volontairement des pièges : un article qui n'existe pas, une référence inventée, une demande de conseil juridique, une demande de rédaction de jugement, un contenu sensible susceptible de déclencher la modération. Le banc d'essai est repris après chaque modification des instructions, des sources ou du

modèle, parce qu'un changement qui semble anodin peut modifier le comportement de l'agent.

Une fois l'agent en service, la plateforme fournit des mesures d'utilisation (sessions, engagement, réactions des utilisateurs) et transmet les commentaires des utilisateurs à l'organisation. Le projet pilote a enregistré 893 sessions, soit environ trois par juge par semaine. L'assistant rédacteur représentait à lui seul plus de la moitié des crédits consommés. Il s'agit d'un usage d'appoint, ce qui est cohérent avec un projet pilote de durée limitée fondé sur le volontariat. La répartition de la consommation entre les agents aide à prévoir la capacité nécessaire et à établir les priorités d'entretien. Les juges participants ont aussi évalué chaque agent, ce qui a permis de repérer les écarts de qualité et d'orienter les correctifs. Les agents qui travaillent sur le texte de l'utilisateur ont obtenu les meilleures notes, et les agents de recherche juridique spécialisée, les moins bonnes.

À première vue, ce résultat semble contredire l'importance que le rapport accorde à la préparation des sources, puisque les agents les mieux notés n'ont pas de base de connaissances et que les moins bien notés ont les plus élaborées. La Cour y voit surtout l'effet de la difficulté des tâches. Le juge peut vérifier immédiatement le travail d'un agent qui restructure son texte. L'agent de recherche juridique, lui, reçoit des questions dont une partie dépasse le contenu de sa base. La bonne réponse consiste alors à dire que la source ne contient pas l'information, ce que l'utilisateur perçoit naturellement comme un échec. La comparaison des notes attribuées à ces deux types d'agents est donc peu significative.

La surveillance est confiée à une personne désignée, qui examine périodiquement les mesures et les commentaires. Le paramétrage de chaque agent (instructions, sources et réglages) est archivé à chaque modification, de façon à pouvoir expliquer, pour n'importe quelle date, pourquoi l'agent se comportait comme il le faisait. Un registre des incidents consigne aussi toute erreur importante signalée par un juge, qu'il s'agisse d'une référence inexacte, d'une citation altérée ou d'un blocage injustifié par la modération, avec la date, l'agent en cause et la correction apportée. La Cour reconnaît à la section 10 que ses outils peuvent se tromper. Ce registre lui permet de rendre compte des erreurs survenues et des mesures prises pour les corriger.

9. FEUILLE DE ROUTE À L'INTENTION DES TRIBUNAUX

Les étapes suivantes sont proposées aux tribunaux qui envisagent une démarche semblable. Elles s'appuient sur l'expérience de la Cour et sur les recommandations du Rapport d'évaluation.

Établir d'abord la gouvernance. La Cour a adopté et publié son [cadre de gouvernance en matière d'intelligence artificielle](#) avant que les juges aient accès aux agents. Ce cadre énonce les principes d'utilisation, les usages interdits et la responsabilité du juge à l'égard de tout ce que produisent les outils. C'est sur lui que repose l'acceptabilité des outils, auprès des juges comme du public.

Réunir ensuite les préalables : un environnement Microsoft 365 institutionnel ou l'équivalent, les licences et la capacité de la plateforme, la collaboration de l'administrateur du système et une personne-ressource technique à l'interne. À long terme, cette personne-ressource constitue l'investissement le plus déterminant.

Amorcer la démarche par un projet pilote de taille restreinte, fondé sur le volontariat, d'une durée définie et, s'il y a lieu, en groupes échelonnés. Le projet de la Cour a réuni 22 juges pendant quatorze semaines, ce qui a suffi pour obtenir des mesures utiles sans engager toute l'institution.

Choisir des premiers agents simples et à faible risque. Une loi intelligente est le meilleur point de départ, suivie d'un assistant de révision qui travaille uniquement sur le texte soumis par l'utilisateur. Il est préférable de reporter à une étape ultérieure les agents de recherche juridique spécialisée, puisque c'est pour eux que la Cour a eu le plus de difficulté à atteindre une qualité satisfaisante.

Préparer les sources avant de construire les agents (section 7), rédiger les instructions selon la structure type (section 5), trancher chaque réglage (section 6) et faire passer chaque agent au banc d'essai (section 8) avant de l'offrir aux utilisateurs.

Mesurer, recueillir les commentaires et ajuster, de façon continue. Les agents, les sources, les instructions et le registre des versions demandent un entretien constant, qu'il faut planifier et financer en conséquence. La Cour est prête à partager son expérience avec les tribunaux qui voudraient s'engager dans cette voie.

10. LIMITES CONNUES ET ÉVOLUTIONS

Même ancré dans des documents, un modèle de langage peut se tromper. Il peut rapprocher une question d'un passage qui n'y répond pas, omettre une disposition pertinente que la recherche n'a pas fait remonter ou citer un texte de façon imparfaite. Le Rapport d'évaluation l'a constaté. Les hallucinations, peu nombreuses, se sont concentrées dans les tâches de recherche juridique spécialisée, lorsque la question dépassait le contenu des bases de connaissances. La modération, réglée prudemment, peut aussi bloquer à l'occasion une réponse légitime portant sur des faits sensibles, ce qui n'a rien de rare dans les dossiers de la Cour. Quant aux bases de connaissances, elles ne sont à jour qu'à une date donnée et doivent être



entretenues avec rigueur. C'est en raison de ces limites que les réponses des agents ne constituent qu'un point de départ. Le juge vérifie toute citation dans les sources officielles avant de l'utiliser.

Le projet continuera d'évoluer dans le sens des recommandations du Rapport d'évaluation lorsque la Cour supérieure disposera d'une ressource dédiée, notamment par le suivi de l'utilisation et des commentaires, la réduction des erreurs documentées, l'enrichissement et l'entretien des bases de connaissances, l'interconnexion éventuelle des agents et l'adoption de nouveaux modèles. Le présent rapport sera mis à jour au besoin.

11. CONCLUSION

Les agents de la Cour reposent sur une architecture relativement simple : un grand modèle de langage, des instructions rédigées par la Cour, des sources qu'elle a choisies et préparées pour les agents documentaires, et des réglages qui privilégient la fidélité aux sources. Ils ne remplacent aucune des fonctions du juge. Leur but est de réduire le temps consacré à la recherche documentaire, à la mise en forme et à certaines tâches de soutien, pour en laisser davantage au travail de juger. Aucune de leurs composantes n'a exigé de développement sur mesure, ce qui les rend reproductibles. L'essentiel de l'effort réside dans la rigueur des instructions, la préparation des sources, les essais et l'entretien.

La Cour publie ce rapport en complément du [Rapport d'évaluation du projet pilote](#). Elle espère qu'il sera utile aux justiciables, à la communauté juridique et aux tribunaux qui voudront s'engager dans la même voie, et le mettra à jour à mesure que le projet évoluera.

M^e Guillaume Bourgeois

Adjoint exécutif juridique
Cour supérieure du Québec

ANNEXE A. LA PRÉPARATION DU LIVRE BLEU

Cette annexe décrit en détail le prétraitement du Livre bleu, l'ouvrage de référence de la Cour en droit de la famille. Elle montre, sur un cas réel, comment les principes de la section 7 ont été appliqués, avec les paramètres retenus et les contrôles effectués. Les seuils indiqués ont été établis pour cet ouvrage et devront être revus pour tout autre document.

A.1 Le document de départ

L'ouvrage se présente sous la forme d'un fichier Word unique d'environ 258 000 mots, soit 756 pages imprimées. Il compte 3 183 notes de bas de page, presque toutes des références à la jurisprudence ou à la législation, une cinquantaine de tableaux, sept niveaux de titres et une table des matières automatique d'environ 900 lignes. Chacune de ces caractéristiques nuit à l'indexation. Le fichier est trop long pour permettre des citations précises, les notes sont enregistrées à part, la table des matières encombre les résultats et la hiérarchie des titres, qui situe chaque passage, se perd au découpage.

A.2 Le déroulement du prétraitement

Le traitement est automatisé et se fait en une seule passe. Les notes sont d'abord extraites et réinsérées à l'endroit de leur appel, et les 36 renvois abrégés de l'ouvrage sont résolus au passage. La table des matières et les pages liminaires sont ensuite retirées, et les tableaux convertis en un format texte qui conserve les colonnes. Vient alors le découpage selon la hiérarchie de l'ouvrage. Une unité est retenue si elle compte moins de 60 000 caractères. Celles de moins de 20 000 caractères sont regroupées avec leurs voisines de même rang, et celles d'au plus 1 200 caractères, qui se résument à peu près à un titre, sont jointes à l'unité suivante. En dernier lieu, chaque fichier reçoit son en-tête de provenance, les rappels de chemin sous ses sous-titres et un nom tiré du chemin hiérarchique.

Ces opérations sont exécutées par un script Python d'un peu plus de cinq cents lignes, développé à l'interne à partir de bibliothèques courantes de lecture des fichiers Word et PDF. Le script prend en entrée le fichier Word de l'ouvrage et le PDF correspondant, qui sert à l'alignement des pages. Il produit les fichiers du corpus, l'index et le registre des descriptions. Le traitement complet du Livre bleu prend moins d'une minute, ce qui permet de le relancer au besoin plutôt que de corriger le corpus manuellement. Le code n'est pas reproduit ici, parce que c'est la méthode qui compte, mais la Cour pourrait le communiquer sur demande.

A.3 L'alignement sur la pagination officielle

L'alignement se fait sur le PDF que Word génère à partir de l'ouvrage. Il a d'abord été vérifié que le numéro imprimé de chaque page correspond à sa position réelle dans le fichier, pour les 721 pages numérotées. Les deux textes sont ensuite uniformisés et parcourus en parallèle. Pour chaque paragraphe, une chaîne de 45 caractères est recherchée dans le texte du PDF, en partant de la position courante et sur une vingtaine de pages. Le curseur ne revient jamais en arrière. Si un paragraphe n'est pas retrouvé, il prend simplement la page du paragraphe précédent, et l'alignement se poursuit.

Cette méthode a un point faible. Si le curseur avance trop loin par erreur, tout ce qui suit est perdu, et rien ne le signale. C'est ce qui s'est produit au premier essai, qui permettait de rechercher une chaîne plus courte à n'importe quelle distance. Seulement 54 % des paragraphes avaient alors été retrouvés. Une fois cette recherche de repli limitée au voisinage immédiat, le taux d'appariement est passé à 95 %.

A.4 Les contrôles de qualité

Un corpus mal aligné conserve une apparence correcte. Les renvois semblent plausibles, et seule une vérification révèle qu'ils sont faux. Cinq contrôles ont donc été effectués. Le taux d'appariement, d'abord, est de 95 % (4 973 paragraphes sur 5 219). La séquence des pages ne recule jamais, ce qui confirme que l'alignement n'a pas décroché. Chaque marque insérée a été recherchée dans le texte de la page indiquée, et 482 des 483 marques contrôlées étaient exactes ; la seule exception est une ligne de tableau dont la mise en page change à l'extraction. Des pages du PDF tirées au hasard ont aussi été vérifiées visuellement. Restaient les 246 paragraphes non retrouvés, qui ont pris la page du paragraphe précédent et que les contrôles précédents ne couvrent pas. C'est là que se trouve le risque résiduel, sous la forme d'un renvoi décalé d'une page, et un échantillon de ces paragraphes a donc été vérifié à part dans le PDF. Il faut refaire ces contrôles chaque fois que le corpus est régénéré. Si le taux d'appariement tombe sous 90 % environ, il est probable que le PDF ne correspond pas à la version traitée.

A.5 Le résultat et sa mise à jour

Le corpus final compte 75 fichiers, d'environ 2 000 à 88 000 caractères ; le plus long est une annexe sans sous-titres. Il est accompagné d'un index, qui décrit la composition du corpus et les pages couvertes par chaque fichier, et d'un registre des descriptions à utiliser au moment du versement dans l'agent. Toutes les notes de l'ouvrage s'y trouvent, intégrées au texte.



À chaque nouvelle édition de l'ouvrage, le traitement est relancé avec la nouvelle version et son PDF, et les fichiers de l'ancienne version sont retirés de l'agent avant le versement des nouveaux. Ni l'ouvrage d'origine ni son PDF ne sont versés en plus du corpus. Un contenu présent en double nuit à la recherche et amène l'agent à citer tantôt une version, tantôt l'autre.

A.6 Transposition à un autre document

Enfin, le traitement repose sur des conventions de mise en forme que rien ne garantit dans un fichier Word. Il suppose que l'ouvrage utilise les styles de titre standards ; avec des styles personnalisés, la hiérarchie serait ignorée et le traitement produirait un fichier unique. Il suppose aussi que la table des matières a été générée automatiquement et que chaque renvoi abrégé vise bien la note qui le précède. Ces trois points se vérifient en quelques minutes, et doivent l'être avant de lancer le traitement. S'ils ne sont pas respectés, aucune erreur n'apparaît, mais le corpus obtenu, correct en apparence, sera de moindre qualité.