

Introduction de l'IA dans l'édition juridique : rationalités limitées et prudence généralisée ?

Christophe Dubois, Cris/IRSS

Dans nos recherches portant sur le marché de l'édition juridique en Belgique (Dubois & Delgoffe, 2025), nous avons rendu compte de la situation dans laquelle se situent toutes sortes d'utilisateurs (magistrats, avocats, étudiants, chercheurs et enseignants, autres professionnels du droit, etc.) désireux d'accéder à une grande partie de l'information juridique en Belgique, le plus souvent de manière payante. La position de force des éditeurs juridiques (Wolters Kluwer, Lefebvre-Sarrut et Legitech), fournisseurs uniques de contenus non-substituables, est inscrite dans les contrats d'abonnement auxquels les clients captifs n'ont d'autre choix que de souscrire pour obtenir et prolonger un accès aux informations. Nos premiers résultats indiquent notamment que,

- 1) la position dominante des éditeurs dépend dans une certaine mesure de la **contribution – souvent gratuite** – des auteurs, directeurs de collection et des directeurs des revues ;
- 2) le **prix** des abonnements aux revues constitue une **contrainte financière** pour les clients davantage qu'un vecteur de **choix** ou qu'un signal marchand de **qualité** ;
- 3) les informations juridiques peuvent devenir « **hors de prix** » pour les clients dotés des plus faibles ressources financières, renforçant le **sentiment d'iniquités** tarifaires vécues par ces clients captifs d'un **marché opaque** ;
- 4) les **prix** sont essentiellement **supportés, en bout de chaîne**, par les clients des avocats et par les contribuables dont les impôts financent les abonnements de l'ordre judiciaire et des universités ;
- 5) les **technologies numériques** ont jusqu'ici permis aux éditeurs de renforcer les verrous techniques de captation des clients vis-à-vis des bases de données et moteurs de recherche tels que Jura, Strada et LexNow (via les « Big deals », ou les pratiques d'archivages rendues plus compliquées pour les clients ce qui renforce leur dépendance aux éditeurs, etc.)

Depuis deux ans, ces mêmes éditeurs juridiques ont amorcé un nouveau virage technologique pour intégrer l'IA générative à leur plateforme. C'est ainsi que *GenIA-L* est disponible dans *Strada lex* depuis juin 2024 ; que *LexGAIN* a été intégrée à *LexNow* en mars 2025 ; et que le passage de *Jura* à *InView Legal* est annoncé par Wolters Kluwer pour décembre 2025. Ces annonces concomitantes présentent ces développements comme la suite logique dans le développement des services successivement offerts via des revues et des ouvrages « papiers », puis via des bases de données numériques et des moteurs de recherche. Les éditeurs présentent l'introduction de l'IA générative comme un moyen de « révolutionner l'accès aux textes juridiques » via des outils rapides, fiables, intègres, sécurisés, avec lesquels l'utilisateur peut dialoguer.

« Depuis 25 ans, Jura est une valeur sûre dans le monde juridique. Elle est sans conteste la base de données qui excelle en termes de fiabilité et d'exhaustivité, qui a accompagné des générations de juristes. Il est aujourd'hui temps de passer à un nouveau chapitre : Jura devient InView Legal, une plateforme prête pour l'avenir. [...] Il ne suffit plus d'avoir simplement accès aux documents appropriés, mais aussi de disposer d'un **outil rapide et fiable** pour une **réponse concluante et**

utile. » (<https://www.wolterskluwer.com/fr-be/expert-insights/from-jura-to-inview-legal-ready-for-the-future-of-legal-research>)

« Selon nous, l'intelligence artificielle va **révolutionner** les méthodes d'accès à l'information juridique. Jusqu'à présent, cet accès s'effectuait via des mots-clés, permettant d'identifier des documents parfois très complexes. Désormais, l'utilisateur peut **poser une question directement à l'IA et engager un dialogue avec elle, sous forme de chat**. La solution, afin d'assister l'utilisateur, génère une réponse en s'appuyant sur l'ensemble des données disponibles sur la plateforme ». (Marc-Olivier Lifrange, CEO de Legitech, <https://www.lexnow.lu/blog/articles-4/lia-au-service-de-linformation-juridique-3329>)

« Chez Larcier-Intersentia, notre engagement en matière d'IA repose sur les piliers de **l'intégrité, de la sécurité et de la fiabilité**. C'est pourquoi, nous avons développé GenIA-L, notre propre solution pionnière d'IA générative basée sur l'excellence de nos contenus éditoriaux. » (<https://ai-hub-fr.larcier-intersentia.com/the-future-of-knowledge>)

Nous pouvons raisonnablement supposer que l'emploi d'un tel registre sémantique « côté cour » vise à inspirer confiance aux utilisateurs, certains d'entre eux pouvant se montrer sceptiques ou critiques face aux développements de l'IA dans leur champ d'expertise professionnelle. En coulisses, toutefois, c'est un autre discours qui prévaut : les acteurs impliqués dans le développement de ces solutions d'IA génératives décrivent en effet les limites et inhérentes à ces solutions et les incertitudes qu'elles génèrent. Relevons ici trois types de limites fréquemment invoquées.

Des solutions incertaines : limites techniques, juridiques et pratiques

Tout d'abord, il existe des limites techniques. L'une d'entre elles est relative à l'agentivité de l'IA, capable de sélectionner des sources et d'afficher des résultats de manière automatisée, mais moyennant la présence d'un indispensable travail de paramétrage humain en amont, ce qui contraste avec les discours et promesses d'automatisation totale.

[1] « extrait. » [A, employé d'une maison d'édition juridique]

[2] « extrait. » [B, employée d'une maison d'édition juridique]

Si le travail des développeurs consiste notamment à solutionner ce type de problèmes techniques, un autre type de limite concerne la conception des RAG¹ utilisés. Ainsi, l'un des éditeurs explique que le RAG « maison » - développé par la maison mère au sein du groupe international – a essentiellement été développé pour effectuer des recherches dans la doctrine. Lorsqu'il s'est agi d'étendre le RAG aux monographies puis à la législation, il est apparu que les formats (pdf pour les monographies) et les *structures* textuelles (importance des dates pour la législation, des notes infrapaginales pour les monographies) exigeaient de découper les ensembles de documents en segments plus petits, appelés *chunks*.

[3] « extrait. » [B, employée d'une maison d'édition juridique]

Ensuite, relevons des limites d'ordre documentaire. Certaines sont relatives à la qualité des sources. En effet, dans la mesure où il est impossible de proposer une hiérarchie juridictionnelle

¹ La *Retrieval-Augmented Generation* (ou génération augmentée par récupération) est une technique d'optimisation de réponses de modèle de langage (LLM, *Large Language Models* de type GPT, par exemple) en intelligence artificielle générative. Elle consiste à articuler des bases de données externes à un LLM de fondation.

valable pour toutes les branches du droit, certains résultats de requêtes tendent à visibiliser les décisions de certaines juridictions.

4 « extrait. »

D'autres limites documentaires sont davantage liées à la datation des sources recherchées. Ainsi, **5** l'un des éditeurs indique que le RAG « maison » part systématiquement de la législation en vigueur afin de proposer aux utilisateurs « des informations à jour. On surpondère ainsi la date dans les critères de recherche ». Ce paramétrage pose toutefois un problème aux usagers dont la requête porte sur une législation antérieure, ou sur certaines mesures transitoires, ou encore sur des lois promulguées dont on ignore l'entrée en vigueur ou pour lesquelles des mesures d'application sont nécessaires.

Relevons encore deux autres limites documentaires. L'une concerne l'impossibilité pour la solution de prendre en compte le nombre de consultations et de téléchargements des sources et des documents, notamment – selon un employé d'une maison d'édition juridique – en raison des contraintes relatives au RGPD. L'autre concerne les manuels qui composent la base de données. S'ils sont très nombreux, rares sont ceux qui disposent de mises à jour régulières. Cette absence de mise à jour tend à réduire le volume de la base de données pertinente pour répondre aux requêtes des utilisateurs. L'exigence de mises à jour régulière par les auteurs tend désormais à être inscrite dans les contrats d'édition.

Enfin, l'hétérogénéité des utilisateurs constitue un troisième type de limite pour les solutions basées sur l'IA générative. Cette hétérogénéité concerne notamment les compétences juridiques et techniques des utilisateurs (qui déterminent la précision et la pertinence des requêtes et des résultats), leurs attentes (certains recherchent surtout l'accès à une large base de données, d'autres attendent une analyse spécifique d'un corpus juridique), leurs usages (réguliers et entraînés vs occasionnels et profanes) ainsi que leur attitude (enthousiaste, critique ou raisonnée) vis-à-vis des outils proposés.

Rationalité limitée et prudence des éditeurs juridiques

Nous comprenons à présent que la « révolution » annoncée, si elle est en marche, s'inscrit actuellement dans un périmètre limité. Si divers acteurs s'emploient à repousser ces limites, leur action collective dépend de rationalités individuelles limitées, au sens défini par Herbert Simon (1972). On peut considérer, d'après ce modèle, que la rationalité individuelle se caractérise par trois limites. Premièrement, face à un problème, un individu n'est capable ni d'embrasser simultanément toutes les solutions possibles, ni de les hiérarchiser clairement. On peut dire que ses capacités cognitives sont limitées. Deuxièmement, il ne dispose pas d'une vision claire de ses préférences, lesquelles ne sont ni hiérarchisées, ni stables, ni cohérentes. Troisièmement, l'information traitée est toujours incomplète et rarement de première main. Elle peut contenir des biais et se caractérise souvent par une certaine ambiguïté, si bien qu'elle doit nécessairement être interprétée. En conséquence, si les décisions individuelles sont rationnelles, la rationalité des individus impliqués dans ces décisions est limitée et non pas absolue. Si bien que ces limites sont démultipliées au sein d'un réseau composé de concepteurs, de développeurs, d'utilisateurs, d'éditeurs, de documents datés et structurés, de mots clés, d'algorithmes, de critères et de paramètres. Quant aux décisions prises dans un tel réseau, elles sont satisfaisantes et non optimales. Assumer la rationalité limitée d'un individu et – *a fortiori* – d'une organisation, d'un réseau ou d'une solution d'IA générative consiste à abandonner le mythe de la rationalité absolue.

Nous observons que les éditeurs juridiques assument ce postulat de la rationalité limitée dans leurs discours. Leurs discours et leurs comportements sont en effet teintés d'une certaine forme de prudence.

D'abord, en insistant sur la caractère « suffisant » des solutions proposées aux utilisateurs.

16 « extrait. » [A, employé d'une maison d'édition juridique]

Ensuite, en insistant sur la responsabilité des utilisateurs dans l'usage qu'ils font de l'outil.

17 « extrait. » [C, employé d'une maison d'édition juridique]

Enfin, le caractère gratuit de la plupart des solutions proposées jusqu'ici par Larcier et Legitech (malgré des investissements réalisés en interne et via des partenariats avec Open AI pour Larcier et Gemini pour Legitech) reflète probablement leur plus-value encore limitée.

« Notre solution, que nous mettons à disposition de nos clients sans coût supplémentaire, s'appuie sur Gemini, l'IA générative développée par Google. » (Marc-Olivier Lifrange, CEO de Legitech, <https://www.lexnow.lu/blog/articles-4/ia-au-service-de-linformation-juridique-3329>)

Jusqu'ici, les innovations IA ne semblent pas encore se répercuter sur les tarifs pratiqués. Toutefois, Si un utilisateur peut contracter un abonnement individuel avec Strada enrichi de GenIA-L dès 91 EUR, il peut contracter un abonnement individuel avec GenIA-L Assistant dès 279 EUR, soit pour un tarif trois fois plus élevé. Il convient dès lors de considérer avec nuance l'attitude prudente des éditeurs juridiques vis-à-vis de l'IA générative. Leurs investissements dans ce domaine, au travers de partenariats noués avec des géants (Gemini de Google ou ChatGPT d'OpenAI) ne risquent-ils pas de les rendre, à leur tour, captifs vis-à-vis d'opérateurs planétaires eux-mêmes engagés dans une course effrénée à l'innovation et au développement de technologies toujours plus prometteuses ?

Quoiqu'il en soit, les utilisateurs de ces nouveaux outils d'IA générative, quelle que soit leur hétérogénéité ou celle de leurs pratiques, travaillent gratuitement, non plus uniquement en tant qu'auteurs ou évaluateurs de textes pour le compte des éditeurs, mais en tant qu'auteurs de *digital labor* (Casilli, 2015), leur pratiques de recherche contribuant à enrichir les contextes génératifs pour le compte des géants de la tech. Comme si ces innovations en marge du droit inscrivait le droit à la marge de développements civilisationnels bien plus opaques et centralisés.

Références :

- Casilli, A. (2015). *Digital Labor: travail, technologies et conflictualités*. In *Qu'est-ce que le Digital Labor?* Paris : Institut National de l'Audiovisuel (INA).
- Dubois, C., & Delgoffe, B. (2025). Une économie de la captation sous tension: le cas de l'édition juridique en Belgique. *Droit et société*, 120(2), 319-342.
- Simon, H. A. (1972). Theories of bounded rationality. In C. B. McGuire & R. Radner (Eds.), *Decision and organization* (pp. 161–176). Amsterdam: Elsevier.