

---

# L'intelligence artificielle en études parlementaires : vers un tournant méthodologique ?

**Julien Robin et Jean-François Godbout**

---

🔗 <http://publications-prairial.fr/recp/index.php?id=341>

**DOI : 10.35562/recp.341**

## **Référence électronique**

Julien Robin et Jean-François Godbout, « L'intelligence artificielle en études parlementaires : vers un tournant méthodologique ? », *Revue d'étude et de culture parlementaires* [En ligne], 1 | 2025, mis en ligne le 05 septembre 2025, consulté le 05 février 2026. URL : <http://publications-prairial.fr/recp/index.php?id=341>



# L'intelligence artificielle en études parlementaires : vers un tournant méthodologique ?

Julien Robin et Jean-François Godbout

## PLAN

---

- I. Panorama des usages existants et potentiels des modèles d'IA
    - A. L'intégration concrète des outils d'IA au sein des assemblées parlementaires
    - B. Les usages potentiels des outils d'IA en études parlementaires
      - 1. Le principal potentiel des LLMs : l'analyse de contenu
      - 2. Améliorer les entretiens et questionnaires aux députés
      - 3. Élargir l'exploration de données
      - 4. Modéliser les comportements parlementaires
  - II. Illustration et limites éthiques de l'intégration des outils d'IA
    - A. Exemple d'utilisation d'un LLM pour estimer l'idéologie des députés français
      - 1. Résultats pour la réforme des retraites
      - 2. Résultats pour la loi immigration
    - B. Limite et réflexion
- Conclusion

## TEXTE

---

- <sup>1</sup> Le développement de grands modèles de langage<sup>1</sup> tels que ChatGPT (*Generative Pre-trained Transformer*) d'OpenAI et Mistral AI, suscite l'intérêt de la communauté scientifique, notamment pour leur usage et leur intégration dans la recherche. Le champ des études parlementaires constitue, lui, un terrain particulièrement fertile pour évaluer le potentiel des grands modèles de langage pour l'analyse de textes. En effet, son objet d'étude – le Parlement – produit une vaste quantité de données (débats, scrutins, amendements, lois, procédures, etc.) explorée par les spécialistes<sup>2</sup>. Les LLMs offrent des possibilités en matière d'automatisation de la collecte de données parlementaires massives, qu'elles proviennent de vidéos, d'enregistrements audio ou de discours retranscrits, tout en simplifiant leur catégorisation et en permettant des analyses

textuelles approfondies. Des modèles entraînés pourraient également combler les données partielles ou manquantes, en reconstituant par exemple des débats historiques. Enfin, certains modèles pourraient renforcer les expériences en études parlementaires, notamment dans la modélisation de comportements, puisqu'ils permettent de simuler des réponses similaires à celles de l'humain<sup>3</sup>.

- 2 Si les LLMs ont le potentiel d'élargir les travaux et la profondeur des analyses en sciences sociales<sup>4</sup>, leur intégration en études parlementaires doit être envisagée avec prudence. En effet, le champ s'appuie sur une gamme de méthodes quantitatives (analyse de contenu, expériences) et qualitatives (entretiens et questionnaires) pouvant tirer parti des capacités offertes par les LLMs, mais exige une réflexion sur leurs implications méthodologiques et éthiques. Par ailleurs, les systèmes d'IA rejoignent déjà les capacités humaines dans plusieurs domaines, tels que la reconnaissance vocale, d'image et d'écriture manuscrite, ou encore la compréhension du langage, et tendent à y parvenir dans d'autres (raisonnement prédictif et tests de connaissances générales<sup>5</sup>). Dans ce contexte, le développement des grands modèles de langage interroge le devenir du champ des études parlementaires et de leurs spécialistes : assiste-t-on à un tournant méthodologique ? Autrement dit, la puissance des modèles d'IA questionne le chercheur sur son propre rôle, sur la nature de son travail ainsi que sur la fiabilité des résultats face aux potentiels biais et hallucinations de ce type de modèle.
- 3 Cet article explore l'intégration des grands modèles de langage comme nouvel outil méthodologique en études parlementaires. Nous présenterons d'abord un panorama des fonctions existantes et potentielles de l'IA à usage général dans l'étude des parlements (I). Nous constaterons l'intégration concrète des outils d'IA au sein des assemblées parlementaires (A), nous ferons également un état des lieux des usages potentiels de l'IA dans l'étude des parlements (B). À ce titre, nous notons que notre revue de littérature est marquée par une prédominance des publications en langue anglaise. Nous l'expliquons en partie par l'influence des *legislative studies*<sup>6</sup>, où l'analyse de contenu – une méthode pour laquelle les LLMs présentent le plus grand potentiel – est particulièrement présente, ainsi que par la prépondérance des publications en anglais concernant les LLMs. Ce phénomène pourrait être attribué aux

pratiques et incitations universitaires qui encouragent la publication en anglais pour atteindre un public plus large<sup>7</sup>.

- 4 Dans la section suivante (II), nous illustrerons les capacités des LLMs avec la prédiction de l'orientation idéologique des députés français lors des débats à l'Assemblée nationale sur la réforme des retraites et de la loi immigration en 2023 à l'aide de ChatGPT-4 d'OpenAI. Cet exemple permettra de mettre en lumière le potentiel analytique (A) et les limites des modèles de langage dans l'étude des comportements législatifs. Enfin, nous évoquerons plus largement les implications méthodologiques et éthiques de l'usage des outils d'IA en études parlementaires (B).

## **I. Panorama des usages existants et potentiels des modèles d'IA**

- 5 Après avoir présenté les enjeux de l'intégration des modèles d'IA en études parlementaires, il convient maintenant d'explorer les usages existants et potentiels dans le champ, aussi bien en méthodes quantitatives qu'en méthodes qualitatives.

### **A. L'intégration concrète des outils d'IA au sein des assemblées parlementaires**

- 6 Les outils d'intelligence artificielle sont déjà intégrés dans la réalité parlementaire, que ce soit pour des usages électoraux ou législatifs. Par exemple, l'IA générative permet de réduire les coûts de mobilisation militante<sup>8</sup>, et présente des résultats prometteurs en matière de persuasion des électeurs<sup>9</sup>. Dans les parlements, plus de 200 utilisations potentielles de l'IA sont recensées<sup>10</sup> et certains services parlementaires développent déjà des outils d'IA<sup>11</sup>. Ceux-ci peuvent faciliter le classement de grandes quantités de données, ainsi que contribuer au travail de légistique, comme en France avec LLaMandement, développé par le ministère de l'Économie et des Finances<sup>12</sup>. D'autres outils aident à la transcription et à la traduction des débats, comme le système STAAR<sup>13</sup> au Portugal, ou à l'accompagnement dans le cadre des consultations publiques, comme au Brésil avec un modèle d'analyse de l'opinion publique sur les lois<sup>14</sup>.

## B. Les usages potentiels des outils d'IA en études parlementaires

- 7 Cette section propose un panorama des applications des modèles de langage comme outil méthodologique en études parlementaires. Il s'agit d'explorer le potentiel dans l'analyse de contenu, notamment pour étudier le contenu des débats, dans la réalisation de grilles d'entretien et de questionnaires, ainsi que pour la transcription d'archives historiques et la conception d'expériences en études parlementaires.

### 1. Le principal potentiel des LLMs : l'analyse de contenu

- 8 La vaste quantité de données produites par les assemblées législatives (scrutins, amendements, comptes rendus de séances, rapports, etc.) a longtemps été peu exploitée par les chercheurs en raison de la difficulté de les compiler dans des formats utilisables et de l'absence de méthode d'analyse de contenu approprié<sup>15</sup>. Ainsi, les principales études des comportements parlementaires étaient limitées à l'analyse des scrutins nominaux. Cette approche a pourtant été utile pour comprendre notamment la cohésion<sup>16</sup> et l'unité partisane<sup>17</sup>, ainsi que la discipline des députés<sup>18</sup>, même si elle est limitée par la disponibilité des données de vote<sup>19</sup>. Bien que le cas français présente le principe de publicité des votes en séance publique<sup>20</sup>, le vote à main levée est la procédure la plus courante à l'Assemblée nationale, rendant la capture de la plupart des positions plus difficile, car elles ne sont ni enregistrées ni publiées au *Journal officiel*.
- 9 Le développement rapide de la science des données a fait évoluer la discipline vers l'analyse de contenu, fondée sur l'approche *text as data*<sup>21</sup>, notamment afin de combler les lacunes relatives aux données sur les votes individuels. Les débats parlementaires sont donc devenus une source importante dans la recherche en science sociale, particulièrement pour cartographier les arguments et les controverses ainsi que pour comprendre les rôles au sein des assemblées<sup>22</sup>. Si l'approche lexicométrique a été davantage employée pour observer les rôles dans le cas hexagonal<sup>23</sup>, la

littérature anglophone a davantage exploré les préférences et l'idéologie des députés<sup>24</sup>, les enjeux et les saillances des sujets des agendas parlementaires<sup>25</sup>, ainsi que la qualité de la représentation des intérêts des électeurs – également appelée représentation substantive, par exemple en ce qui concerne le genre<sup>26</sup> – à travers l'analyse de contenu.

- 10 Les travaux ont ainsi rapidement intégré les techniques computationnelles et de l'apprentissage automatique pour analyser les positions idéologiques<sup>27</sup>. L'une des méthodes utilisées est le « plongement lexical » (*word embeddings*), qui transforme les mots en ensembles de nombres (vecteurs) pour représenter les discours législatifs. Cette technique permet de regrouper des mots ayant des significations proches, facilitant notamment l'analyse des discours des députés au fil du temps, comme c'est le cas avec l'algorithme Word2vec<sup>28</sup>. Récemment, des travaux ont employé des modèles plus complexes, de type transformer (tels que BERT et plus récemment RoBERTa<sup>29</sup> ou ChatGPT)<sup>30</sup>, pour déterminer les positions idéologiques des députés dans le contexte des débats. Ces modèles permettent une compréhension plus fine du langage en regroupant les mots en séquences, offrant ainsi une représentation plus dynamique que les vecteurs de mots. Une étude récente analysant la rhétorique émotionnelle lors des débats à l'Assemblée nationale entre 2007 et 2024 illustre les capacités de ChatGPT-4 à identifier les émotions et la polarisation dans les discours durant cette période<sup>31</sup>.
- 11 Ces modèles d'apprentissage profond se distinguent par leur grande adaptabilité, leur capacité à contextualiser et à raisonner tout en traitant rapidement de vastes quantités de données, particulièrement dans leurs versions les plus performantes. L'utilisation de ces systèmes pour analyser le comportement législatif est très prometteuse, notamment parce qu'ils n'ont pas nécessairement besoin de préalablement « connaître » les députés pour comprendre l'orientation idéologique de leurs discours<sup>32</sup>. Ce positionnement peut en effet être calculé en fonction du vocabulaire utilisé, certains mots étant plus susceptibles d'être associés sur le plan sémantique aux idéologies de gauche ou de droite. Pour résumer, les LLMs ouvrent la voie à des analyses de plus grande ampleur, tant dans le temps que

dans les quantités de données exploitées<sup>33</sup>, et présentant des estimations comparables à celles de spécialistes<sup>34</sup>.

## **2. Améliorer les entretiens et questionnaires aux députés**

- 12 Les préférences, les attitudes ainsi que le travail des parlementaires sont aussi explorés en profondeur à l'aide d'entretiens et de questionnaires<sup>35</sup>.
- 13 Les LLMs peuvent aider à la conception et à l'élaboration des guides d'entretien et des questionnaires (par exemple, formulation plus claire et neutre). Certains modèles sont capables de produire des messages de type humain et il est envisageable de simuler des entretiens grâce aux modèles préentraînés, à la lumière des travaux de Ken Kato *et al.*<sup>36</sup>, qui ont réussi à créer des discours de députés japonais. D'autres modèles *open source* comme Whisper AI facilitent la transcription et pourraient aider dans le codage qualitatif des entretiens, par exemple pour la catégorisation et l'interprétation de ce type de données<sup>37</sup>.
- 14 Néanmoins, le recours à ces outils présente des limites. Par exemple, la confidentialité des entretiens peut être compromise si les LLMs dépendent de serveurs externes. Par ailleurs, certains modèles peuvent reproduire les biais socioculturels issus de leurs données d'entraînement<sup>38</sup>, que l'on pourrait retrouver en cas de production de discours parlementaires. Enfin, sans précautions de la part des chercheurs, il y a un risque de distorsions dans l'interprétation des propos des parlementaires ou dans la formulation non neutre dans un questionnaire administré.

## **3. Élargir l'exploration de données**

- 15 Les LLMs offrent également des possibilités pour accéder à de nouvelles sources de données, notamment les archives. La numérisation des sources gouvernementales a permis aux historiens du parlementarisme de capturer numériquement une quantité non négligeable d'archives<sup>39</sup>, consolidant les données primaires. Par ailleurs, de nouveaux usages émergent avec la reconnaissance de l'écriture manuscrite (*handwritten text recognition* – HTR). Cette

technique, capable de reproduire un texte traitable par machine à partir d'images de manuscrits historiques, assure des résultats intéressants et certains logiciels comme Transkribus sont devenus plébiscités par la communauté des chercheurs<sup>40</sup>. Les plus récents grands modèles de langage sont aussi très prometteurs dans ce domaine, en réduisant considérablement les taux d'erreurs de retranscription d'archives manuscrites comparativement aux outils de HTR conventionnels<sup>41</sup>. D'autres modèles, entraînés sur des corpus historiques<sup>42</sup>, offrent des perspectives intéressantes afin de contextualiser et de mieux comprendre certains documents légaux. Enfin, leur capacité à analyser des contenus multimodaux, incluant images, vidéos et audios, renforce leur potentiel pour étudier des données qui, jusqu'à présent, étaient difficiles à traiter. L'étude récente de Marie Puren *et al.*<sup>43</sup> illustre bien comment l'utilisation du traitement naturel du langage sur les débats à la Chambre des députés en France, au <sup>xix</sup><sup>e</sup> siècle, peut enrichir l'analyse de ces archives, en facilitant l'accès et l'exploration de ces données complexes. En résumé, l'utilisation de ces modèles constitue un terrain fertile pour mener des analyses sur le temps long, pour comparer tant les époques que les cas et différents types de contenu.

- 16 Comme souvent, l'absence de données constitue l'une des principales limites pour l'analyse des parlements. Par exemple, certaines assemblées législatives provinciales canadiennes ne disposent de retranscriptions de leurs débats que depuis la seconde moitié du <sup>xx</sup><sup>e</sup> siècle<sup>44</sup>. Ainsi, grâce aux techniques permettant de retracer la structure discursive des débats parlementaires existants<sup>45</sup> (émotions, arguments) et à des modèles entraînés avec des données existantes sur les parlementaires (genre, circonscription, parti politique, etc.), il serait possible non seulement de retracer la position des parlementaires, mais aussi de générer le discours type selon le contexte (période parlementaire, texte travaillé, position vis-à-vis du gouvernement, scrutins). La reconstitution de ce type de débats permettrait un suivi dynamique de l'évolution des idées, aidant à identifier des tendances dans les discours parlementaires dans le temps. Elle ouvrirait également la voie à l'exploration expérimentale de différents scénarios parlementaires alternatifs plausibles par le passé.



## 4. Modéliser les comportements parlementaires

- 17 Dans cette perspective, les LLMs peuvent également être utiles dans les méthodes expérimentales en études parlementaires. Les expériences sont devenues une méthode d'enquête importante en science politique à la fin du <sup>xx</sup><sup>e</sup> siècle <sup>46</sup>, et l'étude des parlements n'y fait pas exception <sup>47</sup>. Ainsi, les modèles de langage préentraînés pourraient modéliser des scénarios complexes en simulant des interactions entre les parlementaires, en matière de comportement de vote, selon les coalitions parlementaires ou encore selon les changements formels (par exemple, changement de procédures) ou politiques au sein des assemblées législatives. Bien que des cadres analytiques fondés sur des modèles informatiques existent déjà <sup>48</sup>, ils sont souvent limités par leur simplicité et leur manque de réalisme. Les LLMs ouvrent ainsi la voie à des comportements beaucoup plus sophistiqués, permettant de tester diverses théories, hypothèses et scénarios contre-factuel, dans un environnement simulé qui ne nécessite pas l'intervention des législateurs. En effet, les LLMs pourraient mener à la création de simulations complexes et réalistes, dans lesquelles des agents conversationnels d'IA seraient amenés à évoluer dans une interface de type bac à sable <sup>49</sup>. Dans de tels scénarios, il serait possible de peupler une assemblée de députés virtuels en leur attribuant une biographie, des objectifs législatifs, une identité politique et un ensemble de règles de fonctionnement. On pourrait alors observer, par exemple, leur comportement législatif en faisant varier certaines caractéristiques structurelles ou partisans, telles que les règles de procédure, la taille de leur groupe parlementaire ou la cohésion partisane. Ces simulations ne remplaceraient évidemment pas la réalité, mais elles pourraient offrir un environnement expérimental synthétique permettant de tester des prédictions théoriques à un coût relativement faible pour les chercheurs <sup>50</sup>.

## II. Illustration et limites éthiques de l'intégration des outils d'IA

- 18 Après avoir exploré le potentiel des grands modèles de langage, nous faisons une démonstration de leur usage pour la prédiction idéologique et du vote des députés français, au sujet de la réforme des retraites et de la loi immigration en 2023. Malgré des résultats prometteurs, nous aborderons les limites de l'utilisation de l'IA en études parlementaires.

### A. Exemple d'utilisation d'un LLM pour estimer l'idéologie des députés français

- 19 L'utilisation de grands modèles de langage, comme ChatGPT, peut s'avérer simple et rapide pour effectuer des analyses. Nous en faisons l'illustration en mobilisant ChatGPT-4 pour estimer l'idéologie des présidents des groupes parlementaires<sup>51</sup> lors des débats sur la réforme des retraites de 2023<sup>52</sup> et sur la loi immigration de décembre 2023<sup>53</sup>. Ces exemples constituent deux moments cristallisants lors des débats à l'Assemblée nationale et ont montré les fractures idéologiques tant entre les blocs qu'au sein de ceux-ci. Après avoir collecté<sup>54</sup> et anonymisé<sup>55</sup> les interventions des députés, nous les avons soumises à ChatGPT-4 en lui demandant de donner un score idéologique sur une échelle gauche-droite unidimensionnelle de 100 points<sup>56</sup>, ainsi que de prédire son vote pour ou contre le projet de loi. Nous avons conçu une requête spécifique pour chaque texte de loi destinée à guider ChatGPT-4 dans l'analyse des interventions<sup>57</sup>. La requête structure l'interaction avec le modèle en plusieurs étapes : le modèle reçoit un document au format JSON<sup>58</sup> compilant toutes les interventions anonymisées d'un député (1) ; il prend connaissance des arguments présentés (2) ; et évalue si le député est « pour », ou « contre » le projet de loi ou si sa position est « inconnue » (3) ; le modèle attribue ensuite un score sur l'échelle gauche-droite (4).
- 20 Les résultats montrent que le modèle est capable de prédire les positions idéologiques des députés ainsi que leur vote, à l'exception de trois cas pour la prédiction de vote sur la loi immigration. Alors

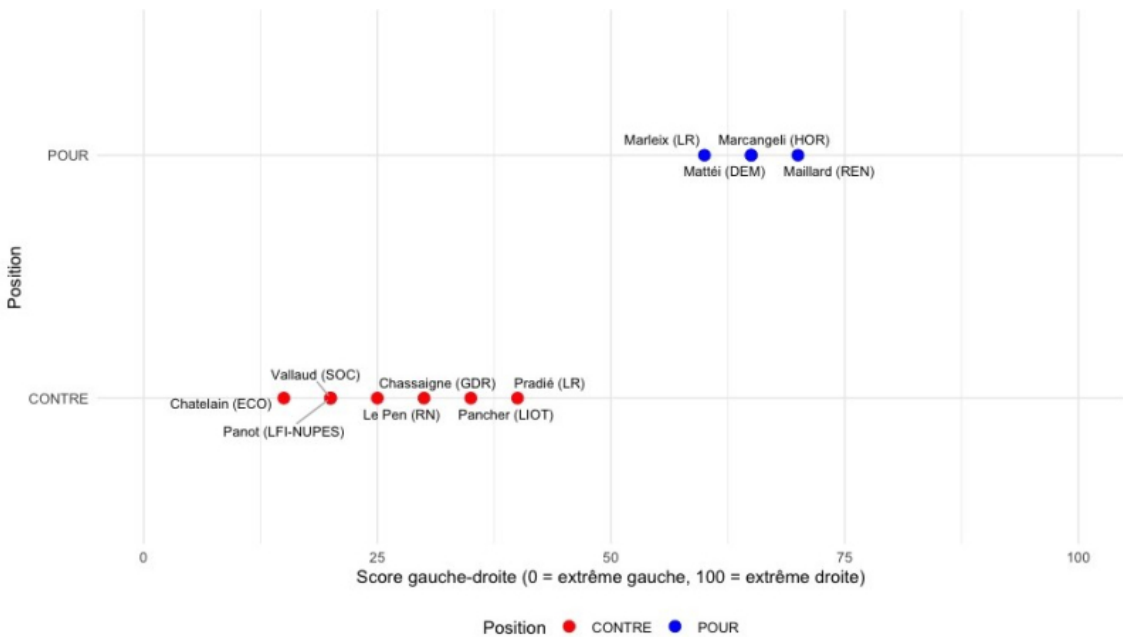
que les données de vote sont vérifiables, car ce sont des votes enregistrés, nous discutons des résultats sur la prédiction du positionnement idéologique des députés avec les dernières données du Chapel Hill Expert Survey (CHES) de 2024<sup>59</sup>.

## 1. Résultats pour la réforme des retraites

- 21 Les résultats de ChatGPT-4 (figure 1 et tableau 1) sont similaires aux votes de la motion de censure transpartisane déposée par le groupe LIOT<sup>60</sup> le 20 mars 2023. Rappelons que le gouvernement Borne a entrepris la réforme des retraites, à travers le Projet de loi de financement rectificative de la sécurité sociale (PLFRSS) pour 2023. L'utilisation de cette procédure budgétaire par le gouvernement était alors considérée comme stratégique, car elle permettait de contourner la procédure législative ordinaire et autorisait le gouvernement à utiliser la procédure de l'article 49, alinéa 3 de la Constitution<sup>61</sup>. Cette procédure, utilisable sans limites dans le cadre de la procédure budgétaire<sup>62</sup>, engage la responsabilité du gouvernement et permet de faire adopter un projet de loi sans vote de l'Assemblée nationale, sauf si une motion de censure est adoptée. En réaction au recours à l'art. 49, al. 3, deux motions de censure sont déposées, l'une par le groupe RN<sup>63</sup>, l'autre par le groupe LIOT. Pour notre démonstration, nous considérons la motion déposée par le groupe LIOT comme un vote exprimant les positions des députés sur la réforme des retraites plutôt que sur la responsabilité du gouvernement. En effet, cette motion, plus consensuelle que celle qui a été déposée par le groupe RN et de par sa nature transpartisane, a fait office de vote sur l'ensemble du texte. Au niveau de notre analyse, nous remarquons, par exemple, que Marine Le Pen, présidente du groupe RN, obtient le score de 30 (centre gauche), ce qui peut sembler surprenant étant donné que le RN est traditionnellement classé à l'extrême droite de l'échelle, comme le montrent les diverses vagues du Chapel Hill Expert Survey<sup>64</sup> (CHES), avec un score de 8,81 sur 10 en 2024. Cependant, ChatGPT ne se trompe pas puisque les positions économiques du RN se sont rapprochées du centre de l'échelle (8,7 en 1999 contre 6 en 2024<sup>65</sup>), et la question des retraites est devenue l'un des principaux thèmes de Marine Le Pen, à l'image des enjeux socio-économiques<sup>66</sup>. À l'inverse, le président du groupe LR<sup>67</sup>, Olivier Marleix, obtient le score de 60, en raison de son

opposition aux taxations excessives des retraités, notamment les plus modestes, soulignant que les politiques fiscales actuelles ont réduit leur pouvoir d'achat. Ce score rapproche davantage le groupe LR du centre de l'échelle de la variable économique, comparativement aux données du CHES en 2024 (7,81 sur 10). Nous avons aussi testé un député considéré comme rebelle sur ce vote au sein des LR, Aurélien Pradié, ayant voté la motion de censure, contrairement au groupe<sup>68</sup> et malgré la discipline de parti<sup>69</sup>. Ce dernier obtient le score de 40, que ChatGPT justifie par le manque de clarté du projet de réforme, notamment concernant les effets négatifs sur les carrières longues.

Figure 1 – Position des députés sur la réforme des retraites (2023)



Crédits : Julien Robin et Jean-François Godbout

Source : Estimation des positions par ChatGPT-4, voir <https://chatgpt.com/share/677b2088-73ac-8005-8465-9432c3ba5fd6>.

Tableau 1 – Position des députés sur la réforme des retraites (2023)

Nom	Prénom	Groupe	Code	Rôle	Score idéologique	Estimation vote	Vote réel
CHATELAIN	Cyrielle	ECO	PA794008	Présidente de groupe	15	CONTRE	POUR
CHASSAIGNE	André	GDR	PA267306	Président de groupe	25	CONTRE	POUR
LE PEN	Marine	RN	PA720614	Président de groupe	30	CONTRE	POUR
MAILLARD	Sylvain	REN	PA717379	Président de groupe	70	POUR	Ne participe pas au vote
MARCANGELI	Laurent	HOR	PA605782	Président de groupe	65	POUR	Ne participe pas au vote
MARLEIX	Olivier	LR	PA606098	Président de groupe	60	POUR	Ne participe pas au vote
MATTÉI	Jean-Paul	DEM	PA720728	Président de groupe	65	POUR	Ne participe pas au vote
PANCHER	Bertrand	LIOT	PA333421	Président de groupe	35	CONTRE	POUR
PANOT	Mathilde	LFI	PA720892	Présidente de groupe	20	CONTRE	POUR
PRADIÉ	Aurélien	LR	PA720100	Député rebelle au groupe LR	40	CONTRE	POUR
VALLAUD	Boris	SOC	PA719930	Président de groupe	20	CONTRE	POUR

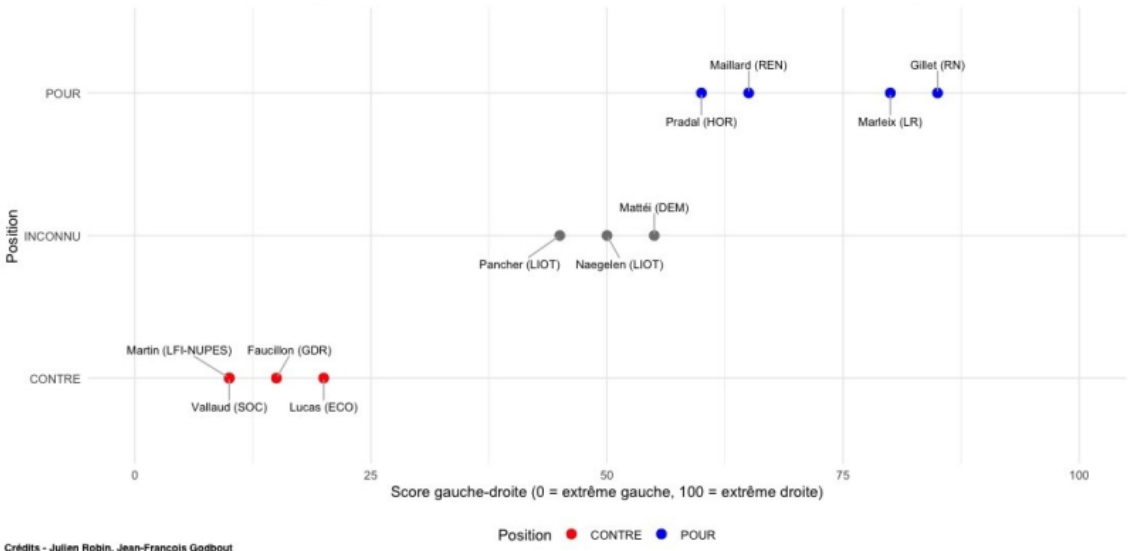
Crédits : Julien Robin et Jean-François Godbout

Source : Estimation des positions par ChatGPT-4, voir <https://chatgpt.com/share/677b2088-73ac-8005-8465-9432c3ba5fd6>.

## 2. Résultats pour la loi immigration

- 22 Les prédictions sur la loi immigration (figure 2 et tableau 2) révèlent quelques difficultés. Par exemple, le modèle n'a pas su déterminer si les députés du groupe LIOT voteraient pour ou contre le texte. Son président, Bertrand Pancher, qui avait voté contre, obtient la note de 45 sans prédiction de son vote. ChatGPT le justifie par une position qu'il estime équilibrée. En revanche, le député Christophe Naegelen, également membre de LIOT, qui avait voté en faveur de la loi, obtient le score de 50, mais là encore, le modèle ne parvient pas à prédire correctement son vote. Cela peut s'expliquer par ses interventions nuancées, entre la nécessité de contrôles plus stricts de l'immigration et la régularisation des travailleurs.
- 23 Le résultat pour le président du groupe démocrate, Jean-Paul Matteï, est intéressant puisqu'il obtient la note de 55, justifiée selon le modèle par un discours équilibré entre promotion de l'intégration et reconnaissance de la migration comme une constante historique. Bien que le député se soit abstenu, le modèle prédit un vote en faveur du texte, probablement en raison du rappel qu'il a fait, lors de son intervention, du soutien à la version initiale du projet de loi<sup>70</sup>. Nous notons que ce score est proche des données du CHES pour le parti Mouvement démocrate (MoDem), puisque le score est de 5,36 sur 10 en 2024<sup>71</sup>.

Figure 2 – Position des députés sur la loi immigration (2023)



Crédits : Julien Robin et Jean-François Godbout

Source : Estimation des positions par ChatGPT-4, voir <https://chatgpt.com/share/677b20fd-f624-8005-b733-c1a135d70b8b>.

Tableau 2 – Position des députés sur la loi immigration (2023)

Nom	Prénom	Groupe	Code	Rôle	Score idéologique	Estimation vote	Vote réel
FAUCILLON	Elsa	GDR	PA721896	Oratrice pour le groupe GDR	15	CONTRE	CONTRE
GILLET	Yohann	RN	PA793832	Orateur pour le groupe RN	85	POUR	POUR
LUCAS	Benjamin	ECO	PA795636	Orateur pour le groupe ECO	20	CONTRE	CONTRE
MAILLARD	Sylvain	REN	PA717379	Président et orateur pour le groupe	65	POUR	POUR
MARLEIX	Olivier	LR	PA606098	Président et orateur pour le groupe	80	POUR	POUR
MARTIN	Élisa	LFI	PA342384	Oratrice pour le groupe LFI- NUPES	10	CONTRE	CONTRE
MATTÉI	Jean-Paul	DEM	PA720728	Président et orateur pour le groupe	55	INCONNU	ABSTENTION
NAEGELEN	Christophe	LIOT	PA721486	Député LIOT, orateur et rebelle	50	INCONNU	POUR
PANCHER	Bertrand	LIOT	PA333421	Président de groupe	45	INCONNU	CONTRE
PRADAL	Philippe	HOR	PA793306	Orateur pour le groupe HOR	60	POUR	POUR
VALLAUD	Boris	SOC	PA719930	Président et orateur pour le groupe	10	CONTRE	CONTRE

Crédits : Julien Robin et Jean-François Godbout

Source : Estimation des positions par ChatGPT-4, voir <https://chatgpt.com/share/677b20fd-f624-8005-b733-c1a135d70b8b>.

24 En résumé, notre exemple de l'utilisation de ChatGPT-4 montre qu'un grand modèle de langage préentraîné semble capable de fournir une estimation idéologique des députés, en quelques secondes et sur l'ensemble d'un débat. Notre exemple ayant été fait à petite échelle, un travail de recherche futur permettrait d'affiner la fiabilité d'un tel modèle, notamment pour prédire la dissension au sein des groupes ou les changements de groupe, par exemple.

## B. Limite et réflexion

- 25 Si notre exemple montre les facilités d'usage d'un LLM comme ChatGPT, ce dernier présente tout de même une barrière tarifaire pour une utilisation plus large. Ces modèles reposent sur des *tokens* – unités de texte pour l'analyse et la compréhension du contenu. Ainsi, un texte de 1 000 mots correspond à environ 1 500 *tokens*. Selon la version du modèle, la capacité de traitement des *tokens* est variable et souvent limitée<sup>72</sup>, nécessitant, pour des volumes de données importants, des versions plus performantes et donc coûteuses<sup>73</sup>. Le recours aux modèles *open source* tels que MiXtral (Mistral AI) ou Llama (Meta) est possible, mais nécessite une configuration locale et des processeurs graphiques (*graphics processing unit* – GPU) pouvant être onéreux. Par ailleurs, l'efficacité algorithmique des LLMs repose principalement sur l'utilisation de GPU et du cloud, dont les infrastructures matérielles et le déploiement soulèvent des enjeux environnementaux, notamment en matière d'extraction de terres rares et de consommation d'énergie<sup>74</sup>.
- 26 Au-delà des conséquences environnementales, les biais algorithmiques ajoutent de la complexité à une perspective éthique de l'usage des LLMs. Comme nous l'avons abordé précédemment, les modèles préentraînés ne sont pas exempts de tout biais socioculturel, dont les hallucinations sont une forme d'expression.
- 27 Enfin, l'utilisation des LLMs interroge le rôle du chercheur vis-à-vis de sa propre recherche. Leurs usages peuvent être utiles dans le processus de recherche : revue de littérature<sup>75</sup>, formulation d'hypothèses<sup>76</sup> ou aide à la rédaction<sup>77</sup> – dont on observe une utilisation croissante<sup>78</sup>. Cependant, le risque est une possible dépossession du chercheur sur son travail. Outre les enjeux de collecte de données ou de propriété intellectuelle, une trop grande imprudence dans l'usage des modèles d'IA affecterait à long terme l'autonomie intellectuelle et la capacité créatrice des chercheurs. À l'instar des appels à son alignement<sup>79</sup> – c'est-à-dire l'harmonisation entre les valeurs et les objectifs établis par les humains et ceux qui sont poursuivis par les modèles génératifs<sup>80</sup> –, une utilisation responsable de l'IA, inspirée de la Déclaration

de Montréal<sup>81</sup>, apparaît essentielle pour préserver l'intégrité de la communauté de chercheurs en études parlementaires.

## Conclusion

- 28 L'intégration des grands modèles de langage en études parlementaires offre un réel potentiel pour les chercheurs. Ceux-ci facilitent la collecte et l'analyse d'une grande quantité de données, tout en permettant l'étude, la modélisation et la simulation des comportements parlementaires. Les LLMs peuvent enrichir l'étude des parlements en favorisant la comparaison et en comblant les données historiques partielles, voire manquantes. Malgré des résultats prometteurs comme nous l'avons montré avec la prédiction idéologique des députés français par ChatGPT, leur utilisation soulève des enjeux méthodologiques et éthiques. Les biais, les hallucinations de l'IA ainsi que les limites techniques et financières pourraient néanmoins affecter l'intégration de l'IA dans le champ. Une utilisation prudente et réflexive de ces modèles est nécessaire pour garantir la qualité et l'intégrité à l'avenir de la recherche en études parlementaires.

## NOTES

---

1 En anglais, *large language model* (LLM). Les grands modèles de langage sont constitués par apprentissage profond (*deep learning*), à partir de mégadonnées textuelles. Ceux-ci constituent la base du contenu généré par les agents conversationnels pour produire des réponses en langage naturel de façon syntaxiquement et sémantiquement correcte. Voir « Grand modèle de langage », *Office québécois de la langue française, Grand dictionnaire terminologique*, URL : <https://vitrinelinguistique.oqlf.gouv.qc.ca/fiche-gdt/fiche/26570880/grand-modele-de-langage> [consulté le 7 mars 2025].

2 J. B. Slapin et S.-O. Proksch, « Words as Data : Content Analysis in Legislative Studies », dans S. Martin, T. Saalfeld et K. W. Strøm (dir.), *The Oxford Handbook of Legislative Studies*, Oxford, Oxford University Press, 2014, p. 126-144, DOI : [10.1093/oxfordhb/9780199653010.013.0033](https://doi.org/10.1093/oxfordhb/9780199653010.013.0033).



- 3 S. Bubeck, V. Chandrasekaran, R. Eldan, J. Gehrke, E. Horvitz, E. Kamar, p. Lee, Y. T. Lee, Y. Li et S. Lundberg, « Sparks of Artificial General Intelligence: Early Experiments with Gpt-4 », *arXiv preprint*, 2023, DOI : [10.48550/arXiv.2303.12712](https://doi.org/10.48550/arXiv.2303.12712).
- 4 C. A. Bail, « Can Generative AI Improve Social science? », *Proceedings of the National Academy of Sciences*, vol. 121, n° 21, 2024, DOI : [10.1073/pnas.2314021121](https://doi.org/10.1073/pnas.2314021121).
- 5 D. Kiela et al., « Test Scores of AI Systems on Various Capabilities Relative to Human Performance », 2023, URL : <https://ourworldindata.org/grapher/test-scores-ai-capabilities-relative-human-performance> [consulté le 3 janvier 2025].
- 6 C. Benoît et O. Rozenberg, « Political Science Approaches to Legislatures », dans *Handbook of Parliamentary Studies*, Cheltenham, Edward Elgar Publishing, 2020, p. 200-222.
- 7 J.-F. Godbout, « Les cinquante ans de la *Revue canadienne de science politique* : Le bilinguisme en déclin ? », *Canadian Journal of Political Science / Revue canadienne de science politique*, vol. 50, n° 1, 2017, p. 3-11, DOI : [10.1017/S0008423917000312](https://doi.org/10.1017/S0008423917000312).
- 8 F. Foos, « The Use of AI by Election Campaigns », *OSF Preprints*, 2024, <https://osf.io/preprints/osf/zm2r6> [consulté le 3 janvier 2025].
- 9 K. Hackenburg, L. Ibrahim, B. M. Tappin et M. Tsakiris, « Comparing the Persuasiveness of Role-Playing Large Language Models and Human Experts on Polarized US Political Issues », *OSF Preprints*, 2023, DOI : [10.31219/osf.io/zm2r6](https://doi.org/10.31219/osf.io/zm2r6).
- 10 J. von Lucke, F. Fitsilis et J. Etscheid, « Using Artificial Intelligence for Legislation--Thinking About and Selecting Realistic Topics », *Working paper Conference EGOV-CeDEM-ePart 2022*, 2022, URL : <https://asgp.co/wp-content/uploads/2023/10/Academic-article-on-using-AI-for-legislation.pdf> [consulté le 3 janvier 2025].
- 11 Union interparlementaire, « Scénarios d'utilisation de l'IA dans les parlements », *Bulletin de l'innovation*, n° 18, 2024, URL : <https://www.ipu.org/fr/bulletin-de-linnovation/histoires/scenarios-dutilisation-de-lia-dans-les-parlements> [consulté le 4 janvier 2025]. Voir également le détail des scénarios sur la page de l'Union interparlementaire qui leur est consacrée : URL : <https://www.ipu.org/fr/ai-use-cases> [consulté le 29 juin 2025].

- 12 J. Gesnouin, Y. Tannier, C. Gomes Da Silva, H. Tapory, C. Brier, H. Simon, R. Rozenberg, H. Woehrel, M. El Yakaabi et T. Binder, « LLaMandement: Large Language Models for Summarization of French Legislative Proposals », *arXiv preprint*, 2024, DOI : [10.48550/arXiv.2401.16182](https://doi.org/10.48550/arXiv.2401.16182).
- 13 Système basé sur la reconnaissance vocale automatique en utilisant Whisper d'OpenAI, réduisant notamment les erreurs de transcription et documentant plus rapidement le contenu des séances plénières et en commission. Voir p. Nascimento, J. C. Ferreira et F. Batista, « Automatic Transcription System for Parliamentary Debates in the Context of Assembly of the Republic of Portugal », *International Journal of Speech Technology*, vol. 27, n° 3, 2024, p. 613-635, DOI : [10.1007/s10772-024-10126-4](https://doi.org/10.1007/s10772-024-10126-4).
- 14 N. F. F. Da Silva et al., « Evaluating Topic Models in Portuguese Political Comments About Bills from Brazil's Chamber of Deputies », dans A. Britto et K. Valdivia Delgado (dir.), *Intelligent Systems. BRACIS 2021*, São Paolo, Springer, 2021, p. 104-120, DOI : [10.1007/978-3-030-91699-2\\_8](https://doi.org/10.1007/978-3-030-91699-2_8).
- 15 J. B. Slapin et S.-O. Proksch, « Words as Data... », *op. cit.*
- 16 H. Rosenthal et E. Voeten, « Analyzing Roll Calls with Perfect Spatial Voting: France 1946–1958 », *American Journal of Political Science*, vol. 48, n° 3, 2004, p. 620-632, DOI : [10.2307/1519920](https://doi.org/10.2307/1519920).
- 17 J.-F. Godbout et M. Foucault, « French Legislative Voting in the Fifth Republic », *French Politics*, vol. 11, n° 4, 2013, p. 307-331, DOI : [10.1057/fp.2013.17](https://doi.org/10.1057/fp.2013.17).
- 18 N. Sauger, « Party Discipline and Coalition Management in the French Parliament », *West European Politics*, vol. 32, n° 2, 2009, p. 310-326, DOI : [10.1080/01402380802670602](https://doi.org/10.1080/01402380802670602).
- 19 C'est par exemple le cas à la Diète japonaise, où seul un nombre limité de lois fait l'objet de votes enregistrés. Voir K. Kato, A. Purnomo, C. Cochrane et R. Saqur, « L(u)PIN: LLM-Based Political Ideology Nowcasting », *arXiv preprint*, 2024, DOI : [10.48550/arXiv.2405.07320](https://doi.org/10.48550/arXiv.2405.07320).
- 20 L'article 63, alinéa 1 du Règlement de l'Assemblée nationale dispose que les votes « s'expriment, soit à main levée, soit par assis et levé, soit au scrutin public ordinaire, soit au scrutin public à la tribune ».
- 21 M. M. Taylor-Robinson, E. Crewe et S. Martin, « The Present of Parliamentary and Legislative Studies », *Parliamentary Affairs*, vol. 75, n° 4, 2022, p. 740-753, DOI : [10.1093/pa/gsac010](https://doi.org/10.1093/pa/gsac010).

- 22 C. de Galembert, O. Rozenberg et C. Vigour, *Faire parler le Parlement. Méthodes et enjeux de l'analyse des débats parlementaires pour les sciences sociales*, Paris, LGDJ, 2013.
- 23 C. Vigour, « Réflexivité, rôles parlementaires et discours de légitimation. Les apports du logiciel Alceste à l'analyse lexicale d'un débat parlementaire relatif aux rapports entre justice et politique », dans C. de Galembert, O. Rozenberg et C. Vigour (dir.), : 'op. cit.', p. 127-47.
- 24 M. Laver, K. Benoit et J. Garry, « Extracting Policy Positions from Political Texts Using Words as Data », *American Political Science Review*, vol. 97, n° 2, 2003, p. 311-331, DOI : [10.1017/S0003055403000698](https://doi.org/10.1017/S0003055403000698) ; J. B. Slapin et S.-O. Proksch, « A Scaling Model for Estimating Time-Series Party Positions from Texts », *American Journal of Political Science*, vol. 52, n° 3, 2008, p. 705-722, DOI : [10.1111/j.1540-5907.2008.00338.x](https://doi.org/10.1111/j.1540-5907.2008.00338.x).
- 25 J. Grimmer, « A Bayesian Hierarchical Topic Model for Political Texts: Measuring Expressed Agendas in Senate Press Releases », *Political Analysis*, vol. 18, n° 1, 2010, p. 1-35, DOI : [10.1093/pan/mpp034](https://doi.org/10.1093/pan/mpp034).
- 26 M. Tremblay, « Do Female MPs Substantively Represent Women? A Study of Legislative Behaviour in Canada's 35th Parliament », *Canadian Journal of Political Science / Revue canadienne de science politique*, vol. 31, n° 3, 1998, p. 435-465, DOI : [10.1017/S0008423900009082](https://doi.org/10.1017/S0008423900009082).
- 27 K. M. Quinn, B. L. Monroe, M. Colaresi, M. H. Crespin et D. R. Radev, « How to Analyze Political Attention with Minimal Assumptions and Costs », *American Journal of Political Science*, vol. 54, n° 1, 2010, p. 209-228, DOI : [10.1111/j.1540-5907.2009.00427.x](https://doi.org/10.1111/j.1540-5907.2009.00427.x).
- 28 M. Van Lange et R. Futselaar, « Debating Evil: Using Word Embeddings to Analyze Parliamentary Debates on War Criminals in The Netherlands », *Contributions to Contemporary History*, vol. 59, n° 1, 2019, DOI : [10.51663/pnz.59.1.07](https://doi.org/10.51663/pnz.59.1.07).
- 29 K. Kato, A. Purnomo, C. Cochrane et R. Saqur, *op. cit.*
- 30 P. Y. Wu, J. A. Tucker, J. Nagler et S. Messing, « Large Language Models Can Be Used to Estimate the Ideologies of Politicians in a Zero-Shot Learning Setting », *arXiv preprint*, 2023, DOI : [10.48550/arXiv.2303.12057](https://doi.org/10.48550/arXiv.2303.12057).
- 31 Y. Algan, T. Renault et H. Subtil, « La Fièvre parlementaire : ce monde où l'on catche ! », *Observatoire du bien-être du Cepremap*, 2025, URL : <https://www.ceprenmap.fr/depot/2025/01/2025-01-Fievre-Parlementaire.pdf> [consulté le 7 mars 2025].

- 32 G. Le Mens et A. Gallego, « Positioning Political Texts With Large Language Models by Asking and Averaging », *arXiv preprint*, 2023, DOI : [10.48550/arXiv.2311.16639](https://doi.org/10.48550/arXiv.2311.16639).
- 33 R. Pujari et D. Goldwasser, « Understanding Politics via Contextualized Discourse Processing », *arXiv preprint*, 2020, DOI : [10.48550/arXiv.2012.15784](https://doi.org/10.48550/arXiv.2012.15784).
- 34 D. Bol et P.-H. Bono, « Est-ce que ChatGPT est une bonne politologue ? », *Note de recherche, Élections européennes 2024, vague 3 de l'enquête électorale*, note 5, 2024, URL : [https://www.sciencespo.fr/cevipof/sites/sciencespo.fr/cevipof/files/Noteelectionseuropeennes\\_DB%26PHB\\_chatGPT\\_mai2024\\_VF.pdf](https://www.sciencespo.fr/cevipof/sites/sciencespo.fr/cevipof/files/Noteelectionseuropeennes_DB%26PHB_chatGPT_mai2024_VF.pdf) [consulté le 17 décembre 2024].
- 35 S. Bailer, « Interviews and Surveys in Legislative Studies », dans S. Martin, T. Saalfeld, et K. W. Strøm (dir.), *op. cit.*, p. 167-184, DOI : [10.1093/oxfordhb/9780199653010.013.0011](https://doi.org/10.1093/oxfordhb/9780199653010.013.0011).
- 36 K. Kato, A. Purnomo, C. Cochrane et R. Saqur, *op. cit.*
- 37 S.-C. Dai, A. Xiong et L.-W. Ku, « LLM-In-The-Loop: Leveraging Large Language Model for Thematic Analysis », *arXiv preprint*, 2023, DOI : [10.48550/arXiv.2310.15100](https://doi.org/10.48550/arXiv.2310.15100).
- 38 A. Abid, M. Farooqi et J. Zou, « Large Language Models Associate Muslims With Violence », *Nature Machine Intelligence*, vol. 3, n° 6, 2021, p. 461-463, DOI : [10.1038/s42256-021-00359-2](https://doi.org/10.1038/s42256-021-00359-2).
- 39 A. Keeling et J. Sandlos, « Shooting the Archives: Document Digitization for Historical–Geographical Collaboration », *History Compass*, vol. 9, n° 5, 2011, p. 423-432, DOI : [10.1111/j.1478-0542.2011.00771.x](https://doi.org/10.1111/j.1478-0542.2011.00771.x).
- 40 J. Nockels, p. Gooding, S. Ames et M. Terras, « Understanding the Application of Handwritten Text Recognition Technology in Heritage Contexts: A Systematic Review of Transkribus in Published Research », *Archival Science*, vol. 22, n° 3, 2022, p. 367-392, DOI : [10.1007/s10502-022-09397-0](https://doi.org/10.1007/s10502-022-09397-0).
- 41 M. Humphries, L. C. Leddy, Q. Downton, M. Legace, J. McConnell, I. Murray et E. Spence, « Unlocking the Archives: Large Language Models Achieve State-of-the-Art Performance on the Transcription of Handwritten Historical Documents », *Social Science Research Network*, 2024, DOI : [10.2139/ssrn.5006071](https://doi.org/10.2139/ssrn.5006071).
- 42 S. Sarker, A. T. Hamad, H. Alshammari, V. Grieco et p. Rao, « Seventeenth-Century Spanish American Notary Records for Fine-Tuning

Spanish Large Language Models », *arXiv preprint*, 2024, DOI : [10.48550/arXiv.2406.05812](https://doi.org/10.48550/arXiv.2406.05812).

43 M. Puren *et al.*, « From Parliamentary History to Digital and Computational History: A NLP-Friendly TEI Model for Historical Parliamentary Proceedings », *Digital Scholarship in the Humanities*, vol. 40, suppl. 1, 2025, DOI : [10.1093/llc/fqae071](https://doi.org/10.1093/llc/fqae071).

44 Les débats parlementaires n'existaient pas en Alberta avant 1972, en Colombie-Britannique (après 1970) et en Ontario (après 1974).

45 J. Bara, A. Weale et A. Biquelet, « Analysing Parliamentary Debate with Computer Assistance », *Swiss Political Science Review*, vol. 13, n° 4, 2007, p. 577-605, DOI : [10.1002/j.1662-6370.2007.tb00090.x](https://doi.org/10.1002/j.1662-6370.2007.tb00090.x).

46 J. N. Druckman, D. P. Green, J. H. Kuklinski et A. Lupia, « The Growth and Development of Experimental Research in Political Science », *American Political Science Review*, vol. 100, n° 4, 2006, p. 627-635, DOI : [10.1017/S0003055406062514](https://doi.org/10.1017/S0003055406062514).

47 J. N. Druckman, T. J. Leeper et K. J. Mullinix, « The Experimental Study of Legislative Behaviour », dans S. Martin, T. Saalfeld et K. W. Strøm (dir.), *op. cit.*, DOI : [10.1093/oxfordhb/9780199653010.013.0029](https://doi.org/10.1093/oxfordhb/9780199653010.013.0029).

48 *Ibid.*

49 J. S. Park, J. C. O'Brien, C. J. Cai, M. R. Morris, p. Liang et M. S. Bernstein, « Generative Agents: Interactive Simulacra of Human Behavior », *arXiv preprint*, 2023, DOI : [10.48550/arXiv.2304.03442](https://doi.org/10.48550/arXiv.2304.03442).

50 S. Jordan, H. L. Paul et A. Q. Philips, « How to Cautiously Uncover the “Black Box” of Machine Learning Models for Legislative Scholars », *Legislative Studies Quarterly*, vol. 48, n° 1, 2023, p. 165-202, DOI : [10.1111/lsq.12378](https://doi.org/10.1111/lsq.12378).

51 Ou des orateurs désignés dans le cadre de la discussion générale.

52 Projet de loi de financement rectificative de la sécurité sociale pour 2023.

53 Projet de loi pour contrôler l'immigration, améliorer l'intégration.

54 Comptes rendus des séances des 6, 7, 8, 10, 13, 14, 15, 16 et 17 février 2023 pour la réforme des retraites et le 12 décembre 2023 pour la loi immigration. Les données sont disponibles à l'adresse suivante : <https://data.assemblee-nationale.fr/archives-16e/debats> [consulté le 29 juin 2025].

55 Les précisions sont disponibles en annexe.

- 56 Nous reprenons la même échelle développée dans G. Le Mens et A. Gallego, « Positioning Political Texts With Large Language Models by Asking and Averaging », 2023, *arXiv preprint*, DOI : [10.48550/arXiv.2311.16639](https://doi.org/10.48550/arXiv.2311.16639). Les auteurs utilisent cette échelle pour évaluer la position idéologique de divers textes politiques (tweets de membres du Congrès américain, manifestes des partis politiques britanniques et discours de législateurs européens. L'échelle a été conçue pour capturer les nuances idéologiques des textes, permettant de positionner les textes sur un même continuum idéologique, et ce, avec une granularité plus fine.
- 57 Les requêtes sont disponibles en annexe et les échanges avec ChatGPT sont disponibles à l'adresse suivante : <https://chatgpt.com/share/675f5c06-e590-8005-8e16-765f7d0949a4> [consulté le 29 juin 2025].
- 58 *JavaScript Object Notation* : ce format est préféré en raison de l'extraction et du nettoyage des débats au format XML (*Extensible Markup Language*).
- 59 J. Rovny, R. Bakker, L. Hooghe, S. Jolly, G. Marks, J. Polk, M. Steenbergen et M. Vachudova, « 25 Years of Political Party Positions in Europe: The Chapel Hill Expert Survey, 1999-2024 », *Chapel Hill Expert Survey*, 2024 working paper, URL : [https://www.chesdata.eu/s/CHES\\_2024\\_final\\_v2-c32x.dta](https://www.chesdata.eu/s/CHES_2024_final_v2-c32x.dta) [consulté le 13 mars 2025].
- 60 Libertés, indépendants, outre-mer et territoires.
- 61 A. Granero, « Réformes de retraites : par quels moyens législatifs le gouvernement peut-il la faire adopter ? », *The Conversation*, 17 janvier 2023, URL : <https://theconversation.com/reforme-des-retraites-par-quels-moyens-legislatifs-le-gouvernement-peut-il-la-faire-adopter-197929> [consulté le 2 mars 2025].
- 62 La procédure budgétaire comprend les lois de finances (LF) et les lois de financement de la sécurité sociale (LFSS), que ce soit pour les projets de loi initiaux (PLF et PLFSS), c'est-à-dire en année  $n-1$ , ou pour les versions rectificatives, c'est-à-dire adoptées au cours de l'année budgétaire  $n$ . Pour plus d'explication, nous référons le lecteur à G. Carcassonne et M. Guillaume, *La Constitution*, Paris, Points, 16<sup>e</sup> éd., 2022, § 335.
- 63 Rassemblement national.
- 64 L'échelle est sur 10 points (0 = extrême gauche ; 5 = centre ; 10 = extrême droite) : 8,7 (1999), 9,92 (2002), 10 (2006), 9,88 (2010) et 9,63 (2014). Voir S. Jolly, R. Bakker, L. Hooghe, G. Marks, J. Polk, J. Rovny, M. Steenbergen et

M. Vachudova, « Chapel Hill Expert Survey Trend File, 1999-2019 », *Electoral Studies*, 2022, DOI: [10.1016/j.electstud.2021.102420](https://doi.org/10.1016/j.electstud.2021.102420)

65 Voir la variable LRECON du CHES qui mesure la distance idéologique sur les questions économiques (rôle de l'État dans l'économie, les dépenses publiques et l'État-providence).

66 G. Ivaldi, « Marine Le Pen, Éric Zemmour : social-populisme contre capitalisme populaire », *Note de recherche. Le Baromètre de la confiance politique*, 2022, Sciences Po / CEVIPOF, vague 13, URL : [https://www.sciencespo.fr/cevipof/sites/sciencespo.fr.cevipof/files/NoteBaroV13\\_GI\\_socialpopulisme\\_mars2022\\_V4.pdf](https://www.sciencespo.fr/cevipof/sites/sciencespo.fr.cevipof/files/NoteBaroV13_GI_socialpopulisme_mars2022_V4.pdf) [consulté le 20 décembre 2024].

67 Les Républicains.

68 JO AN, *Débats*, séance du 20 mars 2023, p. 2711.

69 « Les Républicains : Aurélien Pradié, numéro deux du parti, démis de ses fonctions après “ses prises de position répétées” sur la réforme des retraites », *Le Monde*, 18 février 2023, URL : [https://www.lemonde.fr/politique/article/2023/02/18/aurelien-pradie-demis-de-ses-fonctions-de-numero-deux-de-lr-a-la-suite-de-ses-prises-de-position-sur-la-reforme-des-retraites\\_6162385\\_823448.html](https://www.lemonde.fr/politique/article/2023/02/18/aurelien-pradie-demis-de-ses-fonctions-de-numero-deux-de-lr-a-la-suite-de-ses-prises-de-position-sur-la-reforme-des-retraites_6162385_823448.html) [consulté le 7 mars 2025].

70 JO AN, *Débats*, séance du 11 décembre 2023. Voir notamment la position de l'orateur du groupe, Erwan Balanant, qui votera contre la version révisée et issue de la commission mixte paritaire.

71 La variable Immigration policy mesure la position des partis concernant l'immigration, sur une échelle de 0 à 10, où 0 représente une position soutenant fortement une politique libérale en matière d'immigration ; et 10 représente une position soutenant fortement une politique restrictive.

72 Par exemple, le modèle utilisé pour notre cas, GPT-4o, est limité à 8 192 tokens.

73 Par exemple, la capacité du modèle GPT-4-32k d'OpenAI monte à 32 000 tokens, mais le coût par token est deux fois plus élevé.

74 AI Action Summit, *Rapport international sur la sûreté de l'IA*, présidé par Yoshua Bengio, 2025, URL : [https://assets.publishing.service.gov.uk/media/67af5a7363d7f4ccee9e236/International\\_AI\\_Safety\\_Report\\_2025\\_executive\\_summary\\_french.pdf](https://assets.publishing.service.gov.uk/media/67af5a7363d7f4ccee9e236/International_AI_Safety_Report_2025_executive_summary_french.pdf) [consulté le 7 mars 2025].

75 G. Wagner, R. Lukyanenko et G. Paré, « Artificial Intelligence and the Conduct of Literature Reviews », *Journal of Information Technology*, vol. 37, n° 2, 2022, p. 209-226, DOI : [10.1177/02683962211048201](https://doi.org/10.1177/02683962211048201).

- 76 Y. J. Park, D. Kaplan, Z. Ren, C.-W. Hsu, C. Li, H. Xu, S. Li et J. Li, « Can ChatGPT Be Used to Generate Scientific Hypotheses? », *Journal of Materiomics*, vol. 10, n° 3, 2024, p. 578-584, DOI : [10.1016/j.jmat.2023.08.007](https://doi.org/10.1016/j.jmat.2023.08.007).
- 77 M. Khalifa et M. Albadawy, « Using Artificial Intelligence in Academic Writing and Research: An Essential Productivity Tool », *Computer Methods and Programs in Biomedicine Update*, vol. 5, 2024, DOI : [10.1016/j.cmpbup.2024.100145](https://doi.org/10.1016/j.cmpbup.2024.100145).
- 78 W. Liang, Y. Zhang, Z. Wu, H. Lepp, W. Ji, X. Zhao, H. Cao, S. Liu, S. He et Z. Huang, « Mapping the Increasing Use of LLMs in Scientific Papers », *arXiv preprint*, 2024, DOI : [10.48550/arXiv.2404.01268](https://doi.org/10.48550/arXiv.2404.01268).
- 79 Y. Bengio, « Des IA comme chercheurs idéalisés : sécuritaires et utiles ? », *Yoshua Bengio (blog)*, 12 mai 2023, URL : <https://yoshuabengio.org/fr/2023/05/12/des-ia-comme-chercheurs-idealises-securitaires-et-utiles/> [consulté le 7 mars 2025].
- 80 Y. Bengio, « Comment des 'IA nocives pourraient apparaître », *Yoshua Bengio (blog)*, 30 mai 2023, URL : <https://yoshuabengio.org/fr/2023/05/30/comment-des-ia-nocives-pourraient-apparaître/> [consulté le 7 mars 2025].
- 81 Université de Montréal, *La déclaration de Montréal pour un développement responsable de l'intelligence artificielle*, 2018, URL : [https://declarationmontreal-iaresponsable.com/wp-content/uploads/2023/01/UdeM\\_Decl\\_IA\\_Resp\\_LA\\_Declaration\\_FR\\_web\\_4juin2019.pdf](https://declarationmontreal-iaresponsable.com/wp-content/uploads/2023/01/UdeM_Decl_IA_Resp_LA_Declaration_FR_web_4juin2019.pdf) [consulté le 7 mars 2025].

## RÉSUMÉS

---

### Français

L'article explore l'intégration des grands modèles de langage (LLMs) en études parlementaires. Ces outils, comme ChatGPT, facilitent l'automatisation de la collecte, l'analyse et la catégorisation des données parlementaires (débats, scrutins, etc.), mais permettent aussi de simuler les comportements législatifs. Si des perspectives méthodologiques sont possibles, l'utilisation des LLMs soulève des enjeux éthiques, concernant les biais algorithmiques et la position même des spécialistes du champ. L'article présente aussi les résultats de la prédiction par ChatGPT-4 de l'orientation idéologique de députés français sur la réforme des retraites et la loi immigration en 2023.



## English

This article explores the integration of large language models (LLMs) in parliamentary studies. These tools, such as ChatGPT, make it easier to automate the collection, analysis and categorization of parliamentary data (debates, votes, etc.), and also to simulate legislative behavior. While methodological perspectives are possible, the use of LLMs raises ethical issues concerning algorithmic bias and the very position of specialists in the field. The article also presents the results of ChatGPT-4's prediction of the ideological orientation of French MPs on pension reform and immigration law in 2023.

## INDEX

---

### Mots-clés

intelligence artificielle, IA, grands modèles de langage, études parlementaires, ChatGPT, analyse de contenu, comportements parlementaires, éthique de l'IA

### Keywords

artificial intelligence, AI, large language models, parliamentary studies, ChatGPT, content analysis, parliamentary behaviour, AI ethics

## AUTEURS

---

### Julien Robin

Doctorant en science politique, université de Montréal, Mila – Institut québécois d'intelligence artificielle

IDREF : <https://www.idref.fr/290558204>

### Jean-François Godbout

Professeur titulaire de science politique, université de Montréal, Mila – Institut québécois d'intelligence artificielle

IDREF : <https://www.idref.fr/290558247>