

L'intelligence artificielle générative à la Bibliothèque de l'Assemblée nationale

Une étude de cas sur les usages potentiels de ChatGPT
en contexte professionnel

Rapport du Groupe de travail sur l'intelligence artificielle
Direction de la Bibliothèque

Décembre 2023



Analyse et rédaction

Audrey Houle
Catherine Lanouette
Véronique Boucher-Lafleur
Service de la recherche

En collaboration avec

Andréa Pelletier, Équipe du développement des collections et du traitement documentaire
Nadia Couture, Équipe des archives et de la gestion documentaire
Nathalie Rheault, Service de l'information

Révision linguistique

Danielle Simard
Service de la recherche

Ce document a été réalisé par la
Bibliothèque de l'Assemblée nationale du Québec
Édifice Pamphile-Le May

1035, rue des Parlementaires

Québec (Québec) G1A 1A3

Pour tout renseignement complémentaire sur les travaux de la Bibliothèque ou pour une demande,
n'hésitez pas à nous joindre.

Bibliothèque
418 643-4408

bibliotheque@assnat.qc.ca

ISBN 978-2-550-96700-2

Dépôt légal, Bibliothèque et Archives nationales du Québec, 2024.

Table des matières

Résumé	1
Introduction	2
Mandat confié au Groupe de travail	2
Méthodologie de recherche	3
Résumé des groupes de discussions	5
Service des collections	5
Service de l'information	8
Service de la recherche	9
Phase expérimentale	11
Méthode d'évaluation	12
Organisation des tests	14
Service des collections	15
Service de l'information	15
Service de la recherche	16
Résultats	18
Service des collections	19
Service de l'information	22
Service de la recherche	25



Conclusion.....	30
Recommandations du Groupe de travail sur l'intelligence artificielle	31
Annexe I – Questionnaires des groupes des discussion semi-dirigée	35
Annexe II – Recension des tâches des services de la Bibliothèque	37
Administration de la Bibliothèque de l'Assemblée.....	37
Service des collections	38
Service de l'information.....	39
Service de la recherche.....	41
Annexe III – Grilles d'évaluation.....	43
Grille d'évaluation – Service des collections	43
Grille d'évaluation – Service de l'information	45
Grille d'évaluation – Service de la recherche	46
Annexe IV – Répertoire des essais.....	47



Résumé

Inspiré par la vague de changements technologiques qui transforment l'ensemble des industries à travers le monde, le présent rapport s'intéresse aux usages potentiels de ChatGPT à la Bibliothèque de l'Assemblée nationale du Québec. À cet égard, un groupe de travail composé de personnes issues de tous les services de la Bibliothèque a été mis sur pied pour étudier la question. D'abord, une recension des tâches réalisées par son personnel a été menée. Ensuite, l'équipe de travail a formé des groupes de discussion de chaque service de la Bibliothèque afin d'échanger sur ChatGPT et de recueillir les commentaires et les critiques des panélistes sur cette technologie. Les rencontres ont aussi aidé à déterminer de quelles façons ChatGPT pourrait être intégré au travail de la Bibliothèque.

Suivant les groupes de discussion, le Groupe de travail a conduit une phase expérimentale en collaboration avec les équipes et les services qui composent la Bibliothèque. Ces tests ont pour but d'évaluer le potentiel réel de ChatGPT ainsi que ses forces et ses faiblesses. Pour ce faire, les évaluatrices et les évaluateurs ont rempli des grilles d'évaluation comprenant une série de critères pour estimer le travail de ChatGPT.

Les résultats démontrent que ChatGPT est un outil généraliste, c'est-à-dire qu'il détient des connaissances sur une très grande gamme de sujets. Il produit des réponses très rapidement et jouit d'une bonne capacité d'analyse du langage naturel. Cependant, l'information générée présente des problèmes de validité. ChatGPT fournit parfois de l'information erronée ou inventée. Lorsque questionné sur des sujets plus précis ou techniques, il n'est pas en mesure de produire une réponse satisfaisante. De plus, elles sont exemptes de sources, ce qui constitue un écueil majeur sur le plan méthodologique. Pour ces raisons, ChatGPT ne semble pas être un outil approprié à la plupart des services de la Bibliothèque.

La dernière section du rapport présente les recommandations du groupe de travail sur ChatGPT et plus largement, ses réflexions entourant l'intégration de l'intelligence artificielle à la Bibliothèque.

Introduction

Créé par la compagnie OpenAI, ChatGPT est un agent conversationnel qui intègre les modèles linguistiques d'intelligence artificielle¹. ChatGPT peut effectuer des tâches très variées comme la rédaction de textes, des réponses à des questions factuelles ou la traduction de textes. Il s'alimente dans un très large corpus de données textuelles et il produit l'information demandée rapidement et dans plusieurs langues. Sa capacité de reconnaissance, de traitement et de génération du langage naturel lui permet d'interagir avec l'utilisateur à la manière d'un être humain et même de faire des liens avec les interactions passées.

ChatGPT constitue une application d'intelligence artificielle au même titre que les logiciels Bard ou Bing développés par d'autres compagnies². En pratique, l'intelligence artificielle est un ensemble de techniques fondé sur l'utilisation d'algorithmes propres à traiter l'information, à prendre des décisions automatisées et à formuler des prédictions. Forme avancée et accessible d'intelligence artificielle, ChatGPT est un outil polyvalent qui offre différentes avenues aux professionnelles et professionnels de l'Assemblée comme la rédaction de textes et la génération d'information sur un sujet donné.


La Bibliothèque de l'Assemblée nationale est d'ailleurs familière avec l'usage de l'intelligence artificielle. Depuis 2019, elle utilise le logiciel *Transkribus* pour le traitement de certains fonds d'archives. Il s'agit d'un logiciel accessible en ligne capable de reconnaître, de numériser, de retranscrire et d'effectuer des recherches à partir de documents historiques. À l'aide d'un modèle créé par l'utilisateur, le logiciel retranscrit automatiquement les caractères manuscrits en texte numérique à l'aide de l'intelligence artificielle. Son usage au cours des dernières années a permis aux historiens de l'Assemblée nationale d'accélérer nettement leur traitement documentaire, témoignant ainsi de l'intérêt pour l'intégration des technologies au sein du flux de travail de la Bibliothèque.

Mandat confié au Groupe de travail

L'arrivée de ChatGPT a soulevé l'intérêt de la Direction de la Bibliothèque en raison de son accessibilité. Soucieuse d'optimiser les processus de travail et de demeurer en phase avec les dernières innovations technologiques, la Direction a lancé un projet pilote pour mesurer les capacités de ChatGPT dans son contexte organisationnel. La disponibilité de l'outil, sa rapidité et les avancées récentes en matière de traitement du langage naturel par OpenAI dans les derniers mois ont milité en faveur de ce choix. D'autres logiciels spécialisés pourraient éventuellement être envisagés. Cependant, la performance avérée de ChatGPT dans des tâches afférentes aux activités de la Bibliothèque a encouragé la Direction à commencer par ce logiciel pour évaluer les possibilités offertes par l'intelligence artificielle générative.

¹ Pour plus d'informations sur ChatGPT, voir : Audrey Houle, « [ChatGPT, qu'es-tu?](#) » : portrait de l'intelligence artificielle au Québec, *Première Lecture*, 26 avril 2023.

² Imad Khan, [ChatGPT vs. Bing vs. Google Bard: Which AI Is the Most Helpful?](#), *CNET*, 10 avril 2023.



ChatGPT a été la cible de critiques, notamment en raison des informations inexactes qu'il génère et des problèmes méthodologiques engendrés par l'absence de références dans ses réponses, la démarche du Groupe de travail permettrait de tester l'outil et de vérifier les fondements de ces critiques.

Cette réflexion sur ChatGPT s'inscrit dans la poursuite de l'objectif 3.1 du Plan stratégique 2023-2027 de la Bibliothèque ainsi que dans les priorités du Plan d'action 2022-2023 du Service de la recherche.

Plan stratégique 2023-2027 de la Bibliothèque

- ▶ Objectif 3.1 : Optimiser les processus de travail de la Bibliothèque

Priorités du plan d'action 2022-2023 du Service de la recherche


- ▶ Évaluer l'apport potentiel de ChatGPT aux processus de travail de la Bibliothèque
- ▶ Examiner l'apport d'autres outils technologiques capables de soutenir le travail de la Bibliothèque en matière de : reconnaissance d'images, moissonnage de données, traitement de données sérielles, visualisation de données, etc.

Méthodologie de recherche

La méthodologie de recherche employée par le Groupe de travail se décline en deux étapes. D'abord, des groupes de discussion semi-dirigés ont été formés par le Groupe de travail avec chaque équipe et service de la Bibliothèque³. Au total, quatre rencontres d'une durée de soixante minutes se sont tenues en format hybride entre le 10 et le 17 mai 2023. Les discussions étaient semi-dirigées par une animatrice. Les questions étaient de type ouvert ou semi-ouvert. Elles ont été légèrement adaptées aux réalités de chaque service pour assurer la fluidité de la discussion tout en conservant l'homogénéité entre les questionnaires. Ces rencontres ont été l'occasion d'échanger avec les membres de chaque service sur leurs connaissances de ChatGPT et de l'intelligence artificielle, les façons dont ces technologies pourraient être intégrées à leur travail et leurs critiques et préoccupations face à l'outil. À partir des conclusions des groupes de discussion, le Groupe de travail a déterminé des usages potentiels de ChatGPT pour chaque service de la Bibliothèque.

Durant l'étape expérimentale, des tests ont été réalisés sur ChatGPT Plus pour mesurer sa capacité à accomplir certaines tâches typiques du personnel de la Bibliothèque. L'élaboration des tests s'est fondée sur la littérature recensée et les aspects soulevés par les groupes de discussion semi-dirigés. Les résultats

³ Les questionnaires utilisés pour les entrevues sont publiés à l'Annexe III.



produits par ChatGPT ont été comparés à des travaux réalisés par le personnel de la Bibliothèque. Les tests ont été effectués sur un abonnement payant afin d'obtenir les meilleurs résultats possibles. Contrairement à la version gratuite, ChatGPT Plus utilise la plus récente technologie développée par OpenAI, il offre un temps de réponse plus rapide et un accès prioritaire aux mises à jour et aux nouvelles fonctionnalités⁴. Il peut également générer des réponses plus longues que la version gratuite.

Certaines tâches ont été exclues de la phase test pour différentes raisons : impossibilité de conduire la tâche de façon numérique, limitations inhérentes à ChatGPT ainsi que le temps et les ressources allouées au Groupe de travail. Il est à noter qu'aucun test n'a été mené par l'Équipe des archives et de la gestion documentaire. Plusieurs considérations ont mené à cette décision telles que la nature physique du traitement des pièces d'archives et les impératifs liés à la protection des renseignements personnels et à la cybersécurité des documents manipulés. Enfin, les tâches administratives ont été exclues de la phase expérimentale. Bien qu'elles comportent un intérêt pour l'usage de l'intelligence artificielle, le Groupe de travail a donné la priorité aux tâches plus spécialisées réalisées par les membres de la Bibliothèque de l'Assemblée nationale en raison des délais impartis et du mandat alloué.

⁴ OpenAI, [Introducing ChatGPT Plus](#), 1^{er} février 2023.

Résumé des groupes de discussion

La Bibliothèque de l'Assemblée nationale offre des services personnalisés aux clientèles politiques et administratives qu'elle dessert. Elle a pour mission de favoriser la vitalité démocratique : elle produit, conserve, valorise et diffuse de l'information et des connaissances faisant autorité auprès des parlementaires, de l'administration et de la société civile⁵. Pour ce faire, elle est composée de près de 70 employées et employés répartis dans trois services : le Service de la recherche, le Service de l'information et le Service des collections.

La création de groupes de discussion dans le cadre des travaux du Groupe de travail sur l'intelligence artificielle avait des visées multiples. D'une part, le Groupe de travail souhaitait bénéficier de l'expérience de l'ensemble du personnel de la Bibliothèque afin d'alimenter ses travaux. Les membres de la Bibliothèque de l'Assemblée réalisent un ensemble de tâches variées qui diffèrent grandement d'une équipe à l'autre. Le Groupe de travail voulait donc recueillir le point de vue et l'expérience de tous les employées et employés intéressés afin que l'exercice reflète adéquatement leur travail. D'autre part, les possibilités avec un outil comme ChatGPT sont multiples et leur exploration requiert une fine connaissance des tâches réalisées par chacun en plus d'une part d'imagination. Les groupes de discussion nourrissent la réflexion et profitaient de la créativité de l'ensemble des membres de la Bibliothèque afin de déterminer les usages potentiels. Enfin, l'idée d'implanter une technologie telle que ChatGPT est susceptible de soulever plusieurs questions et inquiétudes. Les groupes de discussion avaient finalement pour objectif d'ouvrir la discussion avec le personnel de la Bibliothèque sur les zones de sensibilités tout en créant un canal de communication ouvert à toutes et à tous sur le sujet.


Les résultats de ces travaux sont concluants. Les employées et employés de la Bibliothèque ont participé en grand nombre et les échanges ont permis au Groupe de travail d'établir des bases solides pour les étapes subséquentes du mandat. Au total, près de la moitié des effectifs de la Bibliothèque ont participé à ces rencontres. Certains services ont également poursuivi leur réflexion à la suite de la rencontre grâce au canal Teams créé à cet effet. Plusieurs recommandations de ce rapport tirent leur source de ces échanges.

La section ci-après présente un résumé des échanges réalisés lors de chaque groupe de discussion.

Service des collections

Le Service des collections comprend deux équipes : l'Équipe du développement des collections et du traitement documentaire et l'Équipe des archives et de la gestion documentaire. La première équipe veille au développement et à la conservation des différentes collections de la Bibliothèque. Elle est aussi responsable du catalogage des collections et de l'indexation du *Journal des débats*. La seconde équipe est

⁵ Bibliothèque de l'Assemblée nationale, [À propos – mission](#).



responsable de la gestion des documents parlementaires, institutionnels et administratifs de l'Assemblée nationale. Elle réalise l'acquisition, le traitement et la conservation des archives et des objets patrimoniaux détenus par la Bibliothèque. Cette équipe dispose aussi d'une expertise en numérisation. Ce service comprend des techniciennes et des techniciens en documentation, des bibliothécaires, des analystes en gestion intégrée des documents, des archivistes, des agentes et des agents de bureau et d'une technicienne en arts appliqués et graphisme.

Les membres de ces deux équipes ont été consultés distinctement pour respecter leurs particularités.

Équipe des archives et de la gestion documentaire

La rencontre avec l'Équipe des archives et de la gestion documentaire a porté davantage sur l'intelligence artificielle en raison des limites inhérentes à ChatGPT dans le contexte archivistique et de la gestion documentaire. Certains membres de l'équipe sont sceptiques quant à l'efficacité et au potentiel réel d'un logiciel comme ChatGPT dans leur travail, compte tenu des tâches liées à la manipulation d'objets physiques. Il pourrait être envisagé d'utiliser ChatGPT pour effectuer des recherches préliminaires sur des événements et des personnes liées aux fonds d'archives. L'équipe pense que cela pourrait leur faire comprendre plus rapidement et plus globalement la portée d'un enjeu dans les activités d'un parlementaire. Certaines inquiétudes ont néanmoins été soulevées sur la fiabilité des données et à la sécurité de l'information.

En ce qui a trait à la numérisation et au graphisme, ChatGPT ne prend pas en charge directement la reconnaissance optique de caractères et le traitement des images. Dans ce contexte, d'autres outils d'intelligence artificielle, comme les programmes de création d'images, semblent plus pertinents que ChatGPT. Les exemples de Dall-E, de Midjourney et de Stockimg.ai ont été évoqués. La prochaine version de Photoshop pourrait aussi répondre à certains besoins de l'équipe, considérant qu'elle intègre une intelligence artificielle qui permet d'automatiser davantage le traitement d'images. Il a aussi été question du logiciel Skylum, utilisé par Bibliothèque et Archives nationales du Québec (BAnQ), qui aide à reconstruire une image ou un texte. Il peut améliorer la plume et la clarté, en plus d'optimiser la qualité de la numérisation.

Dans la gestion intégrée des documents (GID) et les archives, l'utilisation de ChatGPT nécessite un accès aux bases de données de l'Assemblée. L'équipe de la GID et des archives traitent des informations internes qui sont gérées et conservées dans des bases de données uniquement accessibles dans les réseaux sécurisés de l'Assemblée.

Avant d'autoriser toute connexion entre ChatGPT et les bases de données de l'Assemblée, les membres de l'équipe ont souligné l'importance d'évaluer les risques liés à la protection des renseignements qui y sont stockés. Il convient de prendre en considération que ChatGPT utilise un processus d'apprentissage automatisé et adaptatif en incorporant les informations fournies pour répondre aux questions de ses utilisateurs. Cette caractéristique pourrait constituer une faille de sécurité si elle n'est pas rigoureusement encadrée, d'où la nécessité d'une analyse approfondie du logiciel par les services informatiques et technologiques de l'Assemblée.

Équipe du développement des collections et du traitement documentaire

Plusieurs des tâches effectuées au sein de l'équipe du développement des collections et du traitement documentaire pourraient tirer avantage d'une éventuelle collaboration avec des outils d'intelligence artificielle.

En ce qui concerne le catalogage, la méthode d'importation des notices bibliographiques existe depuis longtemps dans les milieux documentaires, notamment dans les bibliothèques publiques ou universitaires. Or, l'équipe souligne qu'il est moins pertinent dans un contexte de bibliothèque spécialisée comme la Bibliothèque de l'Assemblée nationale. L'institution étant un lieu unique au Québec, l'utilisation d'une méthode générique est difficilement applicable en raison des spécificités de la collection, du catalogue et les normes en vigueur. De plus, le traitement documentaire est susceptible d'être adapté en fonction des usagères et des usagers. Toute forme d'automatisation envisagée devra tenir compte de cette réalité selon les membres de l'équipe.

À cet effet, l'exemple de la Bibliothèque royale de Belgique a fait l'objet de discussions. L'institution belge utilise un outil sur Microsoft Power Platform capable d'effectuer un premier jet de catalogage à partir d'une page titre⁶. Aux yeux de l'Équipe du développement des collections et du traitement documentaire, l'utilisation d'un tel logiciel pourrait être très pertinente.


Le travail d'indexation pourrait aussi profiter de l'intelligence artificielle. L'outil choisi doit toutefois employer le langage contrôlé associé, comme un thésaurus, un répertoire de descripteurs ou de vedettes-matière.

Une partie du travail de cette équipe est déjà automatisée. En effet, pour réaliser la veille des publications gouvernementales, l'équipe utilise le logiciel WebSite-Watcher qui fonctionne par requêtes préenregistrées. Cependant, l'outil demeure imparfait pour les besoins de l'équipe. Les notifications sont émises dès qu'une page est modifiée, et ce, peu importe la modification. L'intelligence artificielle est susceptible d'améliorer ce processus.

Les membres de l'équipe ont aussi indiqué certaines tâches moins propices à l'utilisation de l'intelligence artificielle. Les tâches manuelles liées au traitement documentaire ont été évoquées : ouverture de colis, vérification des bons de commande ou des factures, traitement physique des livres, etc. Les tâches qui nécessitent un travail de réflexion et l'application d'un jugement critique se prêtent moins à l'intelligence artificielle. Il peut notamment être question du choix de termes à ajouter au thésaurus ou à modifier, à la sélection du nombre de points d'accès à donner lors du catalogage et aux décisions relatives à la mise en valeur d'une collection spécifique.

En somme, le jugement de la personne au catalogage ne peut être remplacé par l'intelligence artificielle. Plusieurs choix doivent être faits en fonction du document afin de maintenir l'uniformité et la constance du

⁶ Inetum realdolmen, [L'intelligence artificielle accélère le catalogage des livres à la Bibliothèque royale de Belgique](#).



catalogue. Dans le même ordre d'idées, il importe de veiller à préserver la spécificité du catalogue et des collections de la Bibliothèque.

En général, il ressort des discussions que l'intelligence artificielle est mieux indiquée aux documents déjà en format électronique puisqu'ils ne nécessitent pas de numérisation.

Service de l'information

Le Service de l'information offre une assistance aux usagères et aux usagers de la Bibliothèque en matière de recherche documentaire, de veille et de formation. En plus des services documentaires, son personnel s'occupe de l'accueil des visiteuses et des visiteurs à la Bibliothèque, des prêts ainsi que du classement des documents et du développement des collections. Il est composé de bibliothécaires, de techniciennes et de techniciens en documentation et d'agentes et d'agents de bureau.

Plusieurs usages potentiels de ChatGPT ont été identifiés lors de la rencontre avec les membres du Service de l'information. L'outil pourrait notamment être utile au développement des collections, à la recherche d'information, à la rédaction de résumés, à la saisie de références dans une base de données ou à la production d'une bibliographie. D'autres avenues de nature administrative ont été envisagées comme la rédaction mot pour mot de courriels ou la retranscription des rencontres d'équipe.

Les membres du Service de l'information ont toutefois soulevé certaines limites de ChatGPT. La principale étant que ChatGPT n'utilise pas le [Journal des débats](#) pour répondre à une requête. De plus, le logiciel n'a pas accès au contenu d'un document PDF disponible sur un site Web ni au contenu de bases de données internes ou spécialisées. Il s'agit de ressources utiles au Service de l'information, ce qui limite la pertinence de ChatGPT. Enfin, dans sa forme actuelle, ChatGPT ne se prête pas actuellement à la production de revues de presse, car le corpus de données avec lequel il a été entraîné se termine en septembre 2021.

Il a aussi été souligné que ChatGPT a le potentiel de transformer la nature des demandes que le Service reçoit. En effet, l'outil pourrait améliorer de façon substantielle la démocratisation de l'information et son accès, comme ce fut le cas avec l'arrivée de Google. Cette avancée a engendré une complexification des demandes reçues par le Service de l'information. Il est noté qu'à l'heure actuelle, les usagères et les usagers font souvent appel au Service quand ils ne trouvent pas l'information sur des moteurs de recherche en ligne. Il est aussi soulevé qu'une demande à ChatGPT ne requiert pas l'envoi d'un courriel ou d'un formulaire et que la réponse est obtenue pratiquement sans délai.

Des questions ont été soulevées à savoir si l'utilisation d'un tel outil se traduirait par une diminution des demandes au Service ou, à l'inverse, par une augmentation de leur niveau de complexité. Devant ces transformations, le Service de l'information considère qu'il est essentiel que la Bibliothèque se positionne comme une référence dans l'utilisation de ChatGPT pour les usagères et les usagers.

Service de la recherche


Le Service de la recherche détient une expertise en analyse et en rédaction. Il produit des travaux de recherche aux parlementaires à titre individuel, aux commissions parlementaires, aux délégations internationales et interparlementaires ainsi qu'à l'administration de l'Assemblée nationale. Il a notamment pour vocation de favoriser le développement et la diffusion des connaissances sur l'histoire, le parlementarisme et les institutions parlementaires. Ce service est composé d'analystes aux formations universitaires variées, d'historiens, d'une technicienne en administration et d'une réviseuse linguistique.

Étant un outil spécialisé en langage naturel, ChatGPT revêt un intérêt pour le Service de la recherche à différentes étapes des travaux. En amont, il a été évoqué que ChatGPT serait utile pour expliquer des concepts complexes et pour donner une vue d'ensemble sur un sujet. Des préoccupations ont toutefois été soulevées en ce qui a trait à la fiabilité des données. Des membres de l'équipe ayant déjà utilisé ChatGPT ont mentionné avoir vu l'outil transmettre des informations erronées⁷. Considérant cela, l'équipe met en doute la fiabilité de l'outil. ChatGPT est plutôt considéré comme un bon collaborateur à l'étape du remue-méninge, à la manière de ressources telles que Wikipédia qui peuvent être consultées dans les premiers jalons d'une recherche.

De plus, les membres du Service ont souligné les limites de ChatGPT dans sa version actuelle. Par exemple, ses connaissances sont assez limitées sur des sujets locaux ou régionaux. Il s'agit d'une limite importante considérant que les analystes du Service sont souvent appelés à rédiger des notes d'information sur des enjeux de circonscriptions précises. Le biais de la langue a aussi été évoqué, car beaucoup moins de données en français qu'en anglais se trouvent sur le Web. La qualité des résultats est susceptible d'être nettement inférieure pour un usage en français. En outre, des membres de l'équipe ont souligné que ChatGPT ne cite pas les sources qu'il utilise pour répondre à une question. Cela pose un problème de droit d'auteur en plus de complexifier la vérification de l'exactitude de l'information. Enfin, des analystes ont des doutes sur la sensibilité de l'outil pour traiter de questions délicates. Pour toutes ces raisons, la rédaction d'une note par ChatGPT ne semble pas envisageable à courts et moyens termes. Même dans un futur plus lointain, des membres du Service soulignent qu'un travail de vérification devra être réalisé par un humain.

La production de résumés de mémoires apparaît plus pertinente pour les membres du Service de la recherche. Considérant que dans sa version actuelle, ChatGPT n'accepte pas le téléversement de documents, des questions sont soulevées quant à la démarche à suivre. De plus, ses réponses ont une longueur de jetons maximale et l'on ignore s'il en est de même pour les requêtes.

⁷ Lors de la rencontre avec le Service de la recherche, un membre de l'équipe présente un test qu'il a effectué avec ChatGPT pour la génération de nouvelles données et le calcul de statistiques de base (par exemple, un écart-type). Les résultats montrent que l'outil n'est pas toujours fiable et que les résultats des calculs sont parfois erronés. Or, il y a un potentiel créatif à explorer notamment pour la génération de données ou de contenu fictif.




L'apport de ChatGPT à l'étape postrédaction a aussi été discuté. L'idée de soumettre à l'outil un travail et lui demander de réécrire la note dans le but de changer le ton, d'ajouter des figures de style ou de rendre le texte plus fluide est évoquée.

Les groupes de discussion étaient aussi une occasion pour le personnel de la Bibliothèque d'exprimer leurs critiques par rapport à l'intégration de l'intelligence artificielle dans leur quotidien. Plusieurs personnes ont fait part de leur inquiétude en ce qui a trait à la fiabilité des données et à l'impossibilité d'identifier les sources utilisées par ChatGPT. Ainsi, une validation des informations fournies par l'outil semble essentielle. Dans cette perspective, des personnes se demandent si l'usage de ChatGPT est une véritable économie de temps.

Il a aussi été question du sens à donner à son travail. La monotonie que peut entraîner le travail de collaboration avec l'IA a été soulevée, de même que la préoccupation que le rôle de l'humain se résume à vérifier le travail de la machine. Un point de vue inverse a toutefois été amené : faire le choix d'effectuer une tâche qui peut être réalisée par une machine ne semble pas valorisant pour tous. À cet égard, dans toute éventuelle collaboration avec l'intelligence artificielle la personne conservera ses facultés humaines et son libre arbitre.

De plus, des questions sur le réalisme du projet ont été soulevées par la majorité des services et des équipes. Les impératifs liés à la protection des renseignements que la Bibliothèque détient, la cybersécurité ainsi que la lenteur de l'intégration de technologies dans l'environnement numérique de l'Assemblée pourraient constituer un frein à l'usage de l'intelligence artificielle.

Phase expérimentale



Afin d'évaluer le potentiel réel de ChatGPT pour les services de la Bibliothèque, des tests ont été réalisés par les membres du Groupe de travail sur l'intelligence artificielle et des membres du personnel. Les tests ont été choisis en fonction de leur pertinence pour chaque service, leur potentiel d'automatisation et leur faisabilité. L'ensemble des tests a été conduit en français puisqu'il s'agit de la langue utilisée dans l'environnement de travail de la Bibliothèque. Elle répond à la [directive linguistique](#) de l'Assemblée nationale prônant une utilisation exemplaire de la langue française.

Dans la sélection des tests, les contraintes inhérentes au fonctionnement de ChatGPT ont été prises en compte. En tant que modèle de langage naturel, ChatGPT a été entraîné avec un corpus de données qui s'arrête en septembre 2021. Ses capacités sont donc limitées non seulement sur les questions qui touchent l'actualité, mais aussi sur l'évolution technologique postérieure à cette date. Les tests se basaient donc sur des données antérieures à 2021 pour assurer la comparabilité des résultats.

De plus, en tant qu'agent conversationnel, les résultats produits par ChatGPT se limitent à générer du texte à la suite d'une question ou d'une demande qui lui est soumise. Le texte est limité à 4096 jetons, ce qui inclut les mots, les signes de ponctuation et les espaces. De plus, l'interface de ChatGPT n'est pas en mesure de reconnaître ou de générer des images ou des graphiques au moment d'écrire ces lignes. Il est également impossible de lui donner accès à un fichier de texte afin qu'il se nourrisse de son contenu. Enfin, le système ne cite pas ses sources lorsqu'il produit des réponses.

Pour ces raisons, certaines tâches ont dû être exclues de la phase expérimentale. Le présent rapport vise à présenter des usages potentiels de ChatGPT pour la Bibliothèque de l'Assemblée nationale. Il est donc possible que d'autres tâches soient automatisables à l'aide d'autres logiciels.

Méthode d'évaluation

Les résultats des tests sont mesurés en fonction d'une grille d'évaluation établie à cette fin. La grille a été conçue pour l'ensemble des tests. Elle comporte sept critères d'évaluation qui portent autant sur le fond que sur la forme des résultats obtenus. L'appréciation des différents critères est adaptée en fonction des besoins des services.


La grille comporte les critères d'évaluation suivant :

Validité des éléments de la réponse

Le critère de validité vise à mesurer la véracité et la fiabilité des éléments de réponses.

Pertinence de la réponse

La mesure de la pertinence des éléments se distingue de la validité des éléments. Ce critère d'évaluation mesure le résultat obtenu dans son ensemble en fonction de la réponse attendue. Au-delà du caractère factuel de l'information, le résultat obtenu doit correspondre aux attentes



de l'utilisateur par rapport à la requête initiale. Le critère de pertinence mesure également le degré de connaissances de l'outil dans le contexte évalué.

Cohérence des propos

La mesure de la cohérence des propos vise à évaluer l'enchaînement logique des idées.

La neutralité des termes

Le critère de neutralité mesure la connotation du vocabulaire utilisé par ChatGPT. Il permet également d'évaluer si le propos est teinté d'une opinion.

Performance et rapidité du processus

La mesure de la performance tient compte des périodes d'indisponibilité du logiciel payant ainsi que du nombre de questions de relance nécessaire afin d'obtenir le résultat souhaité.

Structuration de la réponse

Le critère de la structuration de la réponse constitue une appréciation de forme dans la présentation des informations à savoir si les éléments sont présentés dans un ordre cohérent.

Qualité de la langue

La mesure de la qualité de la langue vise à apprécier la qualité du vocabulaire, de la syntaxe et des expressions générés par ChatGPT.


Échelle de notation

Une échelle de notation a été retenue comme mode d'évaluation pour chaque critère plutôt qu'une notation numérique ou une cote. Plusieurs raisons justifient ce choix. D'abord, il s'agit d'une variable catégorique et ordinale. Cela signifie que ses valeurs sont limitées, non quantifiables et qu'elles s'ordonnent entre elles selon un ordre naturel. La méthode permet une meilleure comparabilité entre les critères et les tests puisque les valeurs d'évaluation sont similaires et standardisées. En conséquence, les points forts et les points faibles de ChatGPT seront identifiables facilement et rapidement.

Considérant que ChatGPT génère du texte et de l'information, il est plus facile d'apprécier son travail selon des catégories plutôt qu'en lui attribuant une valeur numérique. Voici une description des catégories utilisées pour l'échelle de notation.

Excellent : Une réponse excellente dépasse les attentes. Elle démontre une compréhension approfondie du sujet, une excellente maîtrise du contenu, une clarté dans l'expression et une capacité à fournir des arguments pertinents.

Très bon : Une réponse très bonne est de haute qualité. Elle démontre une compréhension du sujet, une maîtrise adéquate du contenu, une organisation claire des idées et une capacité à



présenter des arguments. Une réponse très bonne répond aux attentes de manière satisfaisante et se démarque par sa qualité supérieure.

Bon : Une réponse bonne est satisfaisante. Elle démontre une compréhension adéquate du sujet, une connaissance de base du contenu et une organisation cohérente. Bien qu'elle puisse manquer de certains éléments ou de détails approfondis, une réponse bonne répond globalement aux attentes et présente une qualité acceptable. Il n'y a pas d'erreurs factuelles.

Mauvais : Une réponse mauvaise est insatisfaisante. Elle démontre une compréhension limitée du sujet, des connaissances inadéquates du contenu et une organisation confuse ou incohérente. Une réponse mauvaise présente des lacunes significatives, des erreurs fréquentes ou des informations incorrectes qui ne répondent pas aux attentes.

Très mauvais : Une réponse très mauvaise est extrêmement insatisfaisante. Elle démontre une compréhension très limitée ou inexistante du sujet, des connaissances largement inadéquates du contenu et une organisation incohérente ou inintelligible. Une réponse très mauvaise ne répond pas aux attentes et présente des erreurs majeures ou une absence totale de compréhension⁸.

Si certains critères ne sont pas applicables, il est possible d'utiliser la mention **N/D**.

Organisation des tests

La phase expérimentale s'est déroulée du 19 juin au 1^{er} août. Au total, cette phase comprend huit différents types de tests qui se déclinent en 70 essais répartis entre les trois services de la Bibliothèque. Comme mentionné précédemment, l'Équipe des archives et de la gestion documentaire a été exclue de la phase expérimentale.

Afin d'outiller les différents testeurs dans leur démarche, un [Guide de l'utilisateur ChatGPT](#) a été élaboré par les membres du Groupe de travail. Ce guide contient l'ensemble des informations nécessaires pour faciliter l'accès à la plateforme et son utilisation.

Pour chaque test, une [grille d'évaluation](#) est remplie. L'ensemble des tests et des essais effectués sont consignés dans le [répertoire des essais](#) qui contient des informations sur le type de test, les essais réalisés et le nom des personnes ayant réalisé chaque essai.

Au moment d'écrire ces lignes, la version payante de ChatGPT est limitée à 25 questions par période de trois heures. Un [calendrier](#) a donc été créé pour coordonner l'ensemble des personnes impliquées dans la phase test.

⁸ La description de chaque niveau de notation a été effectuée en collaboration avec ChatGPT.

Service des collections

L'Équipe du développement des collections a mené deux types de tests soit le catalogage et l'indexation. Les tâches reliées à l'acquisition de documents ont été exclues de la phase test puisque ce volet est couvert par le Service de l'information.


Le catalogage consiste à décrire des documents afin que ceux-ci soient repérables dans un catalogue. Pour ce faire, des informations sur les documents sont entrées dans des notices bibliographiques comme le titre, l'auteur ou le sujet. Le catalogage à la Bibliothèque de l'Assemblée nationale est effectué conformément au format d'échange de données bibliographiques MARC21 et à la norme RDA. Pour la phase expérimentale, les tests consistent à demander à ChatGPT de générer des notices bibliographiques sur plusieurs types de documents comme des publications gouvernementales, un site Web, un article de périodique et un mémoire déposé en commission parlementaire.

De son côté, l'indexation permet de décrire le contenu intellectuel d'un document à l'aide de descripteurs, c'est-à-dire de mots-clés pertinents en fonction du sujet. Une fois indexé, le document contient un ensemble d'étiquettes qui permettent de l'identifier lors d'une recherche dans un catalogue. Les tests d'indexation consistent à donner un extrait de texte à ChatGPT et à lui demander de fournir entre trois et cinq mots-clés ou expressions pour décrire un extrait de texte en les classant par ordre de pertinence. Les extraits de texte proviennent du [Journal des débats](#) et sont tirés de différentes rubriques des travaux de l'Assemblée nationale : déclarations de députés, dépôt de pétitions, présentation de projet de loi, période de questions et de réponses orales, les motions sans préavis, etc.

Les critères d'évaluation ont été adaptés pour les tests de catalogage et d'indexation puisque plusieurs critères n'étaient pas applicables. Les critères ont été précisés afin de mesurer adéquatement les tâches qui ont été testées. Ainsi, seuls les critères de validité, de pertinence et de performance ont été utilisés pour les tests de catalogage. Pour le volet indexation, tous les critères ont été utilisés à l'exception de celui sur la cohérence.

Service de l'information

Les tests retenus pour le Service de l'information se divisent en deux parties : la recherche d'information et le développement de collections. La recherche d'information se base sur des demandes traitées par le Service avant 2021. Les essais sont établis en fonction de paliers de complexité déterminés par le Service lors du traitement des demandes. Les paliers s'échelonnent de 1 à 5, où 5 correspond au niveau de complexité le plus élevé. Les demandes de niveau 1 ont été exclues puisqu'elles se rapportent aux services offerts par le comptoir de la Bibliothèque comme le prêt de documents ou l'accès à des articles de revues spécialisées. Les essais ont couvert un large éventail de sujets et de types de demandes comme une bibliographie, la recherche de subventions et la recherche comparative. Les résultats obtenus ont été comparés avec celles réalisées par les bibliothécaires.



Le second volet portant sur le développement des collections consiste à demander à ChatGPT des titres d'ouvrages portant sur la politique québécoise, la santé et l'environnement pour une période donnée. Le but de cet exercice étant d'identifier de nouveaux ouvrages récents pour le catalogue de la Bibliothèque. Les résultats obtenus ont été comparés avec les acquisitions réalisées par la Bibliothèque au cours de la même période.

Certaines tâches ont été exclues de la phase expérimentale puisqu'il est impossible de les automatiser. Par exemple, le service à la clientèle et la gestion des documents de la Bibliothèque ne peuvent être effectués par ChatGPT. Les tâches de veille n'ont pas été testées par considération de temps.

Service de la recherche

Pour le Service de la recherche, les tests retenus touchent les domaines de spécialisation les plus fréquemment sollicités : l'histoire, le droit, les sciences et technologies et les sciences sociales. La diversification des domaines vise à mesurer la performance de ChatGPT sur différents sujets afin de cerner les forces et les faiblesses de l'outil. La phase expérimentale est scindée en quatre tests représentant le plus grand ensemble possible de tâches effectuées par le Service.


Les tâches sélectionnées sont la définition de concepts, la conception de plans pour la rédaction, la rédaction de notes et la synthèse d'information. Pour chaque essai, le résultat donné par ChatGPT a été comparé à un travail préparé par une ou un analyste du Service. Les travaux sélectionnés ont été rédigés avant septembre 2021 afin que ChatGPT ait accès à un niveau d'information comparable à celui de l'analyste au moment de la rédaction.

Pour la définition de concepts, les éléments de comparaison proviennent de l'[Encyclopédie du parlementarisme québécois](#). Chaque concept a fait l'objet de deux essais, soit une définition courte et une définition longue. Les concepts appartiennent à des domaines disciplinaires variés soit le droit, l'histoire du Québec, l'histoire générale, la procédure parlementaire et l'architecture.

Les essais réalisés pour la rédaction de plan et de notes d'information utilisent des thèmes similaires en y incluant les sciences et les technologies. Les sujets de la gestation pour autrui, la télémédecine, les voitures autonomes et l'exploitation sexuelle des jeunes Autochtones ont été retenus pour les tests du plan et de la rédaction d'une note d'information.

Finalement, la synthèse d'information a fait l'objet de deux essais consistant à demander à ChatGPT de résumer un mémoire produit à l'occasion des travaux parlementaires⁹. Les mémoires choisis proviennent [du mandat d'initiative sur les impacts des pesticides sur la santé publique et l'environnement](#) de la Commission de l'agriculture, des pêcheries, de l'énergie et des ressources naturelles et du [mandat](#)

⁹ Puisqu'il est impossible de verser un fichier dans ChatGPT dans la version utilisée lors de la phase expérimentale, la conduite de ce test s'est effectuée selon une autre méthode. Pour ce faire, le texte du fichier Word a d'abord été copié, puis collé dans la boîte de dialogue de ChatGPT. Ensuite, il a été demandé à ChatGPT de résumer le texte.




[d'initiative portant sur les enjeux de recyclage et de valorisation locale du verre](#) de la Commission des transports et de l'environnement. Les mémoires sélectionnés varient en longueur et en contenu. L'un porte sur une analyse approfondie réalisée par un organisme alors que l'autre, plus court, est un témoignage.

Certaines activités du Service de la recherche ont toutefois été exclues de la phase expérimentale. C'est notamment le cas pour les activités reliées à la Commission de l'administration publique et le travail administratif. Les analystes attitrés à cette commission réalisent en grande partie des tâches analytiques à l'aide de documents gouvernementaux. Considérant que ChatGPT ne peut accéder à ces contenus qui sont publiés sous la forme de fichiers PDF, ce secteur d'activités n'a pas été couvert par les tests. Cependant, ChatGPT pourrait être un outil de travail plus général pour appuyer les analystes de cette commission. Par exemple, en étant utilisé comme outil de vulgarisation de concepts ou pour la reformulation de phrases. Il pourrait être utilisé aussi pour générer de l'information sur un sujet donné, un peu à la manière d'un moteur de recherche.

Les tâches administratives ont aussi été exclues de la phase expérimentale, car il aurait été impossible de les tester sur ChatGPT. Plusieurs de ces tâches requièrent des logiciels spécifiques auxquels ChatGPT n'a pas accès.

Résultats



À la suite des tests, les grilles d'évaluation ont été compilées afin d'identifier les zones d'expertise et les lacunes de ChatGPT. Les résultats ont été compilés pour chaque type de test et pour chaque service. Pour ce faire, l'échelle de notation a été convertie en une échelle numérique pour permettre le calcul de moyennes. Ces moyennes ont ensuite été transformées dans les catégories de départ, soit d'Excellent à Très mauvais, en fonction du chiffre obtenu.

Les résultats présentés ne sont pas pondérés. Plusieurs raisons justifient ce choix méthodologique. La démarche entreprise par le groupe de travail vise une appréciation qualitative des forces et des faiblesses de ChatGPT. L'échelle de notation choisie offre des catégories de référence pour apprécier les résultats obtenus. La transformation des cotes en résultats a pour seule visée la simplification de la présentation des résultats. Ainsi, la mesure de la performance de l'outil demeure qualitative, et non quantitative. De plus, la pondération des critères impose une appréciation au cas par cas de chacun des tests réalisés. En conséquence, l'importance de chaque critère est susceptible de varier en fonction du type de tâches à accomplir et de la finalité de la demande. La pondération des résultats est intéressante, mais elle aurait ajouté un caractère subjectif à l'appréciation des résultats. Le biais humain étant déjà omniprésent dans l'évaluation de la performance de l'outil, l'ajout d'une pondération supplémentaire aurait posé le risque de réduire l'objectivité des résultats.

Service des collections

Équipe du développement des collections et du traitement documentaire

Pour l'Équipe du développement des collections et du traitement documentaire, les tests ont mobilisé deux personnes qui ont effectué l'ensemble des tests, soit quinze essais de catalogage et quinze essais d'indexation.

Catalogage

La qualité du travail accompli par ChatGPT pour les tâches de catalogage est variable. Occasionnellement, ChatGPT est en mesure de diriger les informations vers les bonnes zones de MARC21, mais il se trompe régulièrement quant aux champs employés et utilise des zones de codage périmées. L'outil fournit parfois les points d'accès alors que dans d'autres cas, ils font défaut. Un problème similaire a été relevé avec la norme RDA : ChatGPT ne transpose pas toujours les informations en RDA et lorsqu'il l'effectue, des erreurs sont présentes. Des erreurs similaires sont présentes dans la ponctuation des notices générées.

Par ailleurs, ChatGPT produit certaines parties des notices en anglais, bien que les requêtes soient rédigées en français. Lorsque relancé sur son utilisation de l'anglais, il propose une nouvelle notice en remplaçant les champs en anglais par de nouvelles zones.

Les résultats des tests démontrent que ChatGPT est capable de différencier les types de documents. Par exemple, il peut repérer un chapitre d'un livre à partir des informations bibliographiques qui lui sont

fournies. Toutefois, il éprouve des difficultés à cataloguer des documents plus complexes et spécialisés comme une étude de crédits, un mémoire déposé en commission parlementaire ou un document numérisé.

Il convient de souligner que ChatGPT est conscient de son expertise limitée en catalogage. Il suggère même de s'adresser à un bibliothécaire professionnel ou à un catalogueur pour obtenir une transcription en MARC21 plus complète et précise. De plus, il mentionne que la norme RDA a évolué depuis sa dernière mise à jour en septembre 2021 et que certains codes et zones à enregistrer qu'il suggère peuvent varier selon les pratiques locales de catalogage de l'institution qui traite le document. Au cours de plusieurs essais, ChatGPT inclut de l'information descriptive sur les codes employés afin de mieux comprendre ce à quoi correspond chaque champ, et ce, même si ce n'était pas spécifié dans la requête initiale.

En somme, les tests démontrent que ChatGPT peut générer des notices bibliographiques à partir de données préalablement identifiées. Or, les notices générées ne sont pas uniformes sur plusieurs aspects, dont la présentation des informations, le choix des zones à utiliser, le codage des zones et des sous-zones et sur la ponctuation. Les règles de catalogage sont employées de façon différente selon le type de document. De plus, il ne permet pas de dériver de l'information provenant d'un catalogue source à partir d'un identifiant tel que l'ISBN.


Le travail fait par ChatGPT demeure minimal en ce qui a trait au catalogage. La validité et la pertinence de son travail sont acceptables, mais elles ne sont pas optimales. Quant à la performance et à la rapidité, la plupart des données sont valides dès la première requête. Le processus est donc rapide lorsque l'information est véridique. À lui seul, ChatGPT ne produit pas un travail équivalent à celui d'une personne de l'équipe du catalogage.

TABEAU 1
Résultats des tests de catalogage

Critère	Moyenne	Cote
Validité des éléments de réponse	3,00	Bon
Pertinence de la réponse	3,24	Bon
Performance et rapidité du processus	3,18	Bon
Total	3,14	Bon

Indexation

Pour le travail d'indexation, ChatGPT n'est pas uniforme dans ses réponses, autant dans la forme que dans le choix des concepts. Par exemple, il présente parfois les concepts entre guillemets. Pour une même série d'extraits, il utilise des mots-clés différents. Dans les extraits portant sur le dépôt d'une pétition à l'Assemblée nationale, ChatGPT a employé le terme pétition, pétition à l'Assemblée nationale et pétition citoyenne québécoise. Considérant que l'uniformité est une caractéristique essentielle du travail d'indexation, les réponses fournies par ChatGPT ne sont pas satisfaisantes à cet égard.



Par ailleurs, ChatGPT peut relever le ton et le contexte de l'extrait dans sa sélection de mots-clés. Lors d'un essai tiré de la période de questions à l'Assemblée nationale, ChatGPT a utilisé le terme « conflit entre députés » comme mot-clé. Quoique l'extrait démontre bel et bien l'opposition entre deux députés, cette expression fait référence davantage au contexte qu'au contenu. Cela démontre que ChatGPT comprend le ton et le contexte du débat, mais il indexe des éléments qui ne se réfèrent pas exclusivement qu'aux sujets abordés. Il manque donc de précision dans ses réponses.

De plus, la structuration des réponses de ChatGPT n'est pas optimale pour le travail d'indexation. En effet, il a de la difficulté à placer les termes dans un ordre cohérent et pertinent. Dans la plupart des essais réalisés, ChatGPT n'a pas été en mesure de fournir les mots-clés selon leur ordre d'importance. Même après certaines relances, il arrive qu'il manque toujours des concepts prioritaires ou que les termes ne soient toujours pas placés selon le bon ordre.

ChatGPT peine aussi à évaluer le nombre de mots-clés suffisant pour décrire un extrait de texte. Par exemple, si on lui demande entre trois et cinq termes, il en fournit toujours cinq même si certains sont moins pertinents et représentatifs de l'extrait. À l'inverse, si on lui demande de limiter le nombre de termes, il utilise les mêmes mots-clés sans forcément revoir le choix des mots-clés choisis.

Pour ces raisons, il est nécessaire de valider chaque réponse qu'il produit, ce qui peut prendre un temps considérable. ChatGPT fait moins bonne figure sur les critères de validité, de pertinence ainsi que de performance et rapidité.

En revanche, ChatGPT brille davantage dans les critères de neutralité et de qualité de la langue. Mis à part certaines répétitions constatées durant les tests, les termes employés par ChatGPT sont neutres et formulés en français standard. Les expressions constituent son point faible. Par exemple, il a utilisé les expressions « consommation en lieux publics et condamnation de l'abus des droits de l'Homme » dans l'indexation qu'il a réalisée. Ces expressions pourraient être reformulées afin d'être plus adéquates. Dans un autre test, ChatGPT a mélangé le français et l'anglais dans l'une des expressions utilisées.

Néanmoins, les tests démontrent que ChatGPT se démarque davantage dans les critères de forme que de fonds pour ce qui est du travail d'indexation. Notamment, il y a un écart important entre les résultats des critères de pertinence, d'une part, et ceux de neutralité et de qualité de la langue, d'autre part. Les valeurs élevées pour ces deux derniers critères ont pour effet d'augmenter la valeur totale de la moyenne pour l'ensemble des critères.

TABLEAU 2
Résultats des tests d'indexation

Critère	Moyenne	Cote
Validité des éléments de réponse	3,00	Bon
Pertinence de la réponse	2,87	Mauvais à Bon
Neutralité de la réponse	4,80	Très bon à Excellent
Performance et rapidité du processus	2,87	Mauvais à Bon
Structuration de la réponse	3,07	Bon
Qualité de la langue	4,40	Très bon
Total	3,50	Bon à très bon

Service de l'information

La phase test menée par le Service de l'information s'est scindée en deux volets : la recherche d'information et le développement des collections. Les tests ont mobilisé quatre personnes de ce service. Au total, 18 essais ont été menés, soit quinze essais pour la recherche d'information et trois essais pour le volet développement des collections.

Recherche d'information

Les résultats pour le volet de la recherche d'informations sont mitigés. La qualité et la pertinence de ses réponses varient grandement en fonction du sujet abordé et du type d'information demandé. Par exemple, ChatGPT réussit mieux les tests du niveau de complexité 2 qu'à ceux de niveau 5. Il peut offrir une réponse adéquate à des demandes de renseignements simples, mais il est limité par le fait qu'il ne peut accéder à des ressources plus spécialisées. En outre, il ne détient pas les capacités nécessaires à la production des bibliographies ou à plus de précisions sur les sources qu'il emploie. À plusieurs reprises, ChatGPT n'a pas lancé d'avertissement ou de mise en garde sur sa limite temporelle de 2021, ce qui pourrait induire l'utilisatrice ou l'utilisateur en erreur. De plus, ChatGPT a généré plusieurs informations erronées. Il est donc nécessaire de contre-vérifier systématiquement ses réponses.

Par ailleurs, ChatGPT obtient de meilleurs résultats pour ce qui est de la forme. La plupart des réponses sont bien structurées. Les propos s'enchaînent bien, sont cohérents entre eux. La qualité de la langue est également très satisfaisante. Les réponses sont rapides, mais il arrive parfois qu'elles ne répondent qu'à une partie de la question. Il est donc nécessaire de le relancer pour avoir plus de précisions ou d'informations sur un aspect de la demande. Afin d'améliorer l'efficacité de ChatGPT, il est parfois nécessaire de scinder la demande en plusieurs questions pour obtenir l'ensemble des éléments de réponse souhaités.

La plupart du temps, ChatGPT donne des réponses neutres. Or, quelques problèmes ont été rencontrés durant les tests. D'abord, un biais eurocentriste a été relevé lors d'un essai. L'ensemble des exemples qu'il a mentionnés sont européens alors que les bibliothécaires avaient relevé des cas provenant d'Amérique latine. De plus, il a proposé peu de références concernant les femmes alors que la recherche et la demande soumise insiste sur cet aspect. Toutefois, aucun élément péjoratif ou stéréotypé n'a été recensé.

Le logiciel produit des informations intéressantes pour amorcer une recherche documentaire. Il peut donc être un bon point de départ pour les bibliothécaires et les techniciennes et techniciens. ChatGPT pourrait être utilisé par le Service de l'information à la manière d'un site de recherche comme Wikipédia, pour se faire rapidement une idée du sujet traité. La validité de l'information générée n'est pas suffisante pour apporter une réponse satisfaisante aux demandes des usagères et des usagers de la Bibliothèque.

En résumé, les membres de l'équipe ont trouvé les potentialités de l'outil intéressantes. Toutefois, la qualité variable des réponses les empêche, pour l'heure, de l'utiliser de façon efficace à leur travail.

TABLEAU 3
Résultats des tests de recherche d'informations

Critère	Moyenne	Cote
Validité des éléments de réponse	2,67	Mauvais à Bon
Pertinence de la réponse	2,53	Mauvais à Bon
Cohérence de la réponse	3,40	Bon
Neutralité de la réponse	3,07	Bon
Performance et rapidité du processus	2,60	Mauvais à Bon
Structuration de la réponse	3,60	Bon à très bon
Qualité de la langue	3,60	Bon à très bon
Total	3,07	Bon

Développement de collections

ChatGPT comporte plusieurs lacunes de validité de l'information générée dans le contexte du développement des collections. Dans la plupart des tests, les titres d'ouvrages proposés par ChatGPT n'existent pas, et ce, même si le sujet et la période temporelle sont précisés dans les consignes qui lui sont données. Pourtant, les titres inventés semblent crédibles, car ils sont en cohésion avec le sujet demandé. Par exemple, l'un des tests consistait à réclamer dix livres publiés en 2020 sur la politique québécoise. Dans sa réponse, ChatGPT crée des titres et utilise des noms de personnalités publiques québécoises comme auteurs tels que Mathieu Bock-Côté, Lucien Bouchard et Alec Castonguay.

ChatGPT reconnaît que ses réponses peuvent comporter des inexactitudes ou des erreurs factuelles. Pour corriger, l'utilisatrice ou l'utilisateur doit relancer ChatGPT ou l'interpeller. Dans un tel cas, ChatGPT répond qu'il n'a pas accès à des bases de données de livres et de publications. Il poursuit en mentionnant qu'il génère des titres fictifs basés sur ses connaissances du sujet. Il recommande même de consulter des bibliothèques, des librairies ou des sites Web spécialisés pour obtenir l'information. Ainsi, ChatGPT est conscient de ses lacunes, mais il répond tout de même aux questions et aux instructions, même s'il ne dispose pas des connaissances nécessaires pour y répondre.

De plus, certains résultats sont transmis en anglais alors que la requête est rédigée en français. Sur les questions relatives aux titres d'ouvrages sur la politique québécoise, il suggère majoritairement des livres en anglais, qu'ils soient réels ou inventés.

Sur sa performance et sa rapidité, ChatGPT répond avec promptitude, mais les contenus doivent être systématiquement vérifiés. Malgré plusieurs relances, ChatGPT n'arrive pas à livrer une liste de références satisfaisante et composée de sources réelles. Les relances et les vérifications nécessaires alourdissent le processus. Sous sa forme actuelle, le logiciel n'est pas utile pour la réalisation de cette tâche.

Tout comme les tests d'indexation, les valeurs plus élevées des critères de structuration et de qualité de la langue ont pour effet d'augmenter la moyenne totale. En excluant ces critères, la moyenne descend de 2,38 à 1,9.

TABEAU 4
Résultats des tests de développement des collections

Critère	Moyenne	Cote
Validité des éléments de réponse	1,00	Très mauvais
Pertinence de la réponse	2,00	Mauvais
Cohérence de la réponse	2,50	Mauvais à Bon
Neutralité de la réponse	2,67	Mauvais à Bon
Performance et rapidité du processus	1,33	Très mauvais
Structuration de la réponse	3,17	Bon
Qualité de la langue	4,00	Très bon
Total	2,38	Mauvais

Service de la recherche

Le Service de la recherche a réalisé 22 essais répartis dans 4 types de tests : 12 essais ont été menés pour la définition de concepts, 4 pour la conception de plans, 4 pour la rédaction de notes d'information et 2 essais pour la synthèse d'information. Dix analystes du service ont contribué à la phase test.

Définition et vulgarisation de concepts

La première partie de la phase expérimentale avait pour objectif d'évaluer l'apport de ChatGPT dans le processus de rédaction d'entrées pour l'[Encyclopédie du parlementarisme québécois](#) de l'Assemblée nationale. Les termes retenus sont les suivants¹⁰ :

TABEAU 5
Termes choisis pour les tests relatifs à l'Encyclopédie du parlementarisme

Catégorie	Terme choisi
Juridique	Législature
Histoire	<i>Gerrymandering</i>
Histoire du Québec	Acte de Québec
Procédure parlementaire	Bâillon
Architecture	Édifice Jean-Antoine-Panet
Architecture	Citadelle de Québec

Afin de comparer la qualité des réponses de ChatGPT avec ce qui a déjà été produit par le service, les sujets retenus font déjà l'objet d'une fiche dans l'Encyclopédie du parlementarisme. Pour chaque terme, une définition courte et une définition longue ont été demandées telles que présentées.

Pour l'ensemble des essais, les définitions courtes étaient de meilleure qualité. À plusieurs reprises, le logiciel a transmis de fausses informations. Toutefois, la grande majorité d'erreurs factuelles est contenue dans les définitions longues. Dans d'autres cas, plusieurs éléments sont difficiles à vérifier, considérant que ChatGPT ne dévoile pas les sources de ses réponses. Ainsi, le processus d'utilisation de ChatGPT est beaucoup plus long pour les définitions longues que les courtes.

Les résultats des tests démontrent cet écart entre les définitions courtes et longues. La moyenne des essais pour les définitions courtes est de 4,03 contre 3,02 pour les longues. Les écarts sont perceptibles pour

¹⁰ Les termes choisis comportent tous une définition courte et une définition longue dans l'Encyclopédie du parlementarisme afin d'avoir des éléments comparatifs pour l'évaluation des tests.

plusieurs critères de fonds tels que la validité, la pertinence et la cohérence des éléments de réponse de ChatGPT. De plus, les notes plus élevées pour les définitions courtes ont une incidence positive sur les moyennes globales pour le test de définition et de vulgarisation de concepts.

Le test portant sur l'édifice Jean-Antoine-Panet a dû être remplacé par la citadelle de Québec puisque ChatGPT n'était pas en mesure de reconnaître le bâtiment.

TABLEAU 6
Résultats des tests de définition et vulgarisation de concepts

Critère	Définitions courtes	Définitions longues	Total	Cote
Validité des éléments de réponse	4,20	2,00	3,10	Bon
Pertinence de la réponse	3,40	2,60	3,00	Bon
Cohérence de la réponse	4,40	3,00	3,64	Bon à très bon
Neutralité de la réponse	3,80	4,20	4,00	Très bon
Performance et rapidité du processus	4,80	2,50	3,55	Bon
Structuration de la réponse	ND ¹¹	4,00	4,00	Très bon
Qualité de la langue	3,60	3,67	3,64	Bon à très bon
Total	4,03	3,02	3,48	Bon

Conception d'un plan

La deuxième partie de la phase test du Service de la recherche visait à évaluer les capacités de ChatGPT à rédiger le plan détaillé d'une note d'information. Trois notes du type [Matière à réflexion](#) ont été sélectionnées ainsi qu'une note traditionnelle. Les sujets abordent la gestation pour autrui, la télémédecine et les voitures autonomes. Ce type de travail se distingue des notes d'information traditionnelles par sa longueur (deux pages), sa disponibilité en ligne et sa structure. Une note d'information classique sur l'exploitation sexuelle des jeunes Autochtones a aussi été retenue en raison de la sensibilité du sujet.

Dans l'ensemble, les résultats des tests montrent que les plans suggérés par ChatGPT pourraient constituer une base de travail de très bonne qualité. Ce qui est proposé par l'outil est très bien structuré et équilibré. Les idées sont bien hiérarchisées et s'enchaînent de manière logique. En ce qui a trait au plan pour une note sur les véhicules autonomes, ChatGPT a cerné les principaux phénomènes entourant l'émergence de ces technologies. Il a aussi proposé certains points d'intérêts originaux, comme l'amélioration de la mobilité

¹¹ Étant donné la longueur des réponses courtes, ce critère n'a pas été évalué pour ce type de test.

pour certains groupes sociaux. De plus, l'outil est capable de structurer un plan selon les sections habituelles d'une note *Matière à réflexion* lorsqu'on le lui demande. Pour le plan d'une note sur la gestation pour autrui, ChatGPT a aussi ajouté des sous-sections et des idées qui pourraient être intégrées aux sections traditionnelles. Cependant, le plan proposé n'indique pas si ChatGPT a une connaissance fine du sujet. Malgré une demande pour un plan détaillé, ChatGPT s'en tient à des idées assez vagues plutôt qu'à des exemples précis.

TABLEAU 7
Résultats des tests de rédaction de plans

Critère	Moyenne	Cote
Validité des éléments de réponse	3,00	Bon
Pertinence de la réponse	3,25	Bon
Cohérence de la réponse	3,25	Bon
Neutralité de la réponse	3,25	Bon
Performance et rapidité du processus	3,75	Bon à Très bon
Structuration de la réponse	3,75	Bon à Très bon
Qualité de la langue	4,00	Très bon
Total	3,54	Bon à Très bon

Rédaction d'une note d'information

En ce qui a trait à la capacité de ChatGPT de rédiger des notes d'information, la performance de l'outil est plus faible. La validité des éléments de réponse constitue la principale lacune. La tendance de l'outil à inventer des sources, des faits et des données pose un sérieux problème pour la rédaction de notes d'information. À plusieurs reprises, l'outil a fourni des renseignements inexacts, comme des numéros d'articles de lois et la tenue de consultations particulières inexistantes. La vérification qui en découle s'avère laborieuse, car ChatGPT ne cite pas, ou cite incorrectement ses sources. Par exemple, il a donné comme référence le numéro douze d'un volume qui, après vérification, ne contient que dix numéros.

Outre ce problème majeur, les réponses de ChatGPT sont plutôt pertinentes, quoique parfois générales. En formulant une demande assez précise, l'outil fournit plusieurs éléments qui se rapprochent de la note initiale. Le fait de scinder la demande en plusieurs parties en suivant les sections prédéfinies d'un *Matière à réflexion* semble donner de meilleurs résultats. Dans tous les cas, ChatGPT fournit une réponse en très peu de temps. Les réponses sont cohérentes et les éléments présentés suivent un enchaînement logique. La qualité de la langue est généralement bonne, mis à part un vocabulaire parfois optimiste ou promotionnel.

En ce qui a trait au test concernant l'exploitation sexuelle des jeunes Autochtones, la description sommaire des nations autochtones est fidèle à la réalité. Il en est de même pour les résumés des rapports d'enquête

ou les éléments de réponse concernant les défis des Premières Nations et des Inuit en matière d'exploitation sexuelle. ChatGPT a démontré une bonne compréhension du sujet et a présenté des éléments de réponse pertinents, mais qui manquent généralement un peu de substance, d'exemples, de citations ou de faits tangibles pour appuyer et étoffer le propos. De plus, le ton, les termes utilisés et le choix des éléments factuels présentés par ChatGPT traduisent bien l'esprit dans lequel un ou une analyste du Service de la recherche aborderait un mandat sur la condition autochtone, soit avec sensibilité et impartialité.

En somme, l'outil semble revêtir une bonne utilité pour donner et structurer entre elles les idées générales à ne pas omettre dans une note d'information, mais moins pour développer chacune de ces idées de manière à constituer un argumentaire suffisant. Les résultats obtenus pour la rédaction d'une note d'information sont bons, mais insuffisants. La validité des éléments de réponse est un des critères essentiels pour mesurer la qualité d'une note d'information.

TABLEAU 8
Résultats des tests de rédaction d'une note d'information

Critère	Moyenne	Cote
Validité des éléments de réponse	2,38	Mauvais
Pertinence de la réponse	3,00	Bon
Cohérence de la réponse	3,25	Bon
Neutralité de la réponse	3,25	Bon
Performance et rapidité du processus	3,00	Bon
Structuration de la réponse	4,00	Très bon
Qualité de la langue	3,75	Bon à très bon
Total	3,23	Bon

Synthèse d'information

Parmi l'ensemble des tests effectués par le Service de la recherche, les plus satisfaisants ont trait à la synthèse de l'information. Deux mémoires présentés en commission parlementaire ont été sélectionnés dans le cadre d'auditions publiques pour les mandats, le premier sur les pesticides et le second sur le recyclage du verre. Ces mémoires, de 1045 mots et de 2360 mots respectivement, avaient déjà fait l'objet d'un résumé par un analyste du service. Puisqu'il n'est pas possible de soumettre à ChatGPT une pièce jointe, le texte a été copié directement dans l'espace de clavardage en précisant le résultat souhaité.

Pour le mémoire sur l'utilisation des pesticides¹², ChatGPT a repris fidèlement les éléments du texte. Il a synthétisé et fait des liens entre les différents points présentés, tout en conservant la structure et l'ordre du mémoire. Sa réponse était structurée en paragraphes et son ton était neutre et objectif. La qualité de la langue était excellente, si bien que la syntaxe du résumé de ChatGPT était meilleure que celle du texte original. Sur la valorisation du verre¹³, la validité des réponses de ChatGPT était tout aussi excellente. L'outil ne s'est pas écarté du texte soumis et l'utilisation de puces numérotées et d'une police en gras ont facilité la lecture.

En outre, ChatGPT est rapide pour générer des réponses, mais il est limité quant à la quantité de texte qu'il peut traiter. Compte tenu de la longueur du mémoire sur le recyclage, le texte a été scindé en deux parties pour obtenir une réponse satisfaisante. Pour cette raison, l'enchaînement logique des propos de ChatGPT a été plus difficile à évaluer. La plupart des réponses sont excellentes, mais l'outil atteint un certain plafond lorsqu'il est trop souvent relancé. De plus, le texte produit par ChatGPT reprend un ton similaire à celui utilisé par l'auteur du mémoire. Or, il convient de souligner que le travail de l'analyste doit être empreint de neutralité. En ce sens, l'analyste est amené à faire un résumé des faits énoncés dans un mémoire sur un ton davantage objectif.

Aussi, ChatGPT a dû suggérer des pistes de réflexion relatives aux deux mémoires. Les propositions se sont avérées très intéressantes. L'outil a notamment su relever les recommandations publiées sur les pesticides et sur le recyclage du verre.

En somme, ChatGPT performe de façon remarquable quand il s'agit de synthétiser un texte, particulièrement lorsque la longueur du texte est sous la limite de ses capacités.

TABLEAU 9
Résultats des tests de synthèse d'information

Critère	Moyenne	Cote
Validité des éléments de réponse	5,00	Excellent
Pertinence de la réponse	4,50	Très bon à Excellent
Cohérence de la réponse	4,50	Très bon à Excellent
Neutralité de la réponse	4,50	Très bon à Excellent
Performance et rapidité du processus	4,00	Très bon
Structuration de la réponse	5,00	Excellent
Qualité de la langue	4,50	Très bon à Excellent
Total	4,57	Très bon à Excellent

¹² Serge Giard, [Les maladies attribuées aux pesticides](#), 18 juillet 2019.

¹³ Réseau Environnement, [Mémoire de Réseau Environnement : les enjeux de recyclage et de valorisation locale du verre](#), 2 juillet 2019.

Conclusion

L'ensemble des résultats obtenus par les différents services et les équipes sont des indices révélateurs à plusieurs égards. ChatGPT est un outil performant, rapide et polyvalent. Il est en mesure de produire en quelques secondes un résultat plausible à un ensemble de questions portant sur un large spectre de sujets. Même interrogé sur des questions sensibles ou mis devant des interactions humaines qui nécessitent une compréhension fine du langage naturel, ChatGPT est en mesure d'adopter une posture adéquate dans ses résultats.

ChatGPT demeure néanmoins un outil généraliste. Il peut traiter d'une très grande diversité de sujets, comme démontré par les tests réalisés, mais il s'égare lorsque les informations demandées requièrent une certaine expertise. Sa performance est d'ailleurs optimale lorsque le résultat est court. L'outil a tendance à s'égarer dans le contenu ou démontre ses limites lorsqu'on lui soumet une question longue, alors qu'il répond avec les mêmes propos que pour une question courte.

Dans son état actuel, la performance générale de ChatGPT demeure insuffisante pour être utilisée par le personnel de la Bibliothèque de l'Assemblée. Ce logiciel comporte des lacunes majeures qui rendent la technologie inintéressante : son corpus ne va pas au-delà de septembre 2021 et la citation des sources est plutôt déficiente. De plus, malgré l'acquisition d'une version payante, plusieurs collègues de la phase test du logiciel ont rencontré des problèmes techniques. À différentes reprises, le compte était inaccessible ou un message d'erreur était affiché lorsque questionné.


Les résultats présentés doivent aussi être nuancés. Les moyennes n'étant pas pondérées, les résultats ne témoignent pas nécessairement de l'appréciation réelle du travail réalisé par ChatGPT comparé aux exigences de rigueur établies par la Bibliothèque de l'Assemblée. La qualité de la recherche, l'exactitude des informations ainsi que la pertinence des propos sont des éléments fondamentaux qui doivent orienter le travail des employées et employés. ChatGPT étant plutôt performant sur les critères de forme, les résultats obtenus offrent une vision optimiste des capacités réelles de l'outil en contexte professionnel.

Le mandat du Groupe de travail étant limité à l'évaluation de la performance de ChatGPT dans sa forme disponible au grand public, les résultats ne permettent pas de juger de la pertinence de l'acquisition d'une licence privée. L'issue tend toutefois à démontrer que les outils basés sur la reconnaissance et le traitement du langage naturel revêtent un intérêt pour les différents services de la Bibliothèque de l'Assemblée nationale.

Recommandations du Groupe de travail sur l'intelligence artificielle



ASSEMBLÉE
NATIONALE
DU QUÉBEC



Au fil de ses travaux, le Groupe de travail sur l'intelligence artificielle a recueilli un ensemble de commentaires sur l'usage de ChatGPT à la Bibliothèque de l'Assemblée nationale. Ces commentaires s'inscrivent dans une réflexion plus large que le mandat octroyé au Groupe de travail. La section ci-après vise à guider la réflexion pour les travaux à venir.

ChatGPT constitue un sujet d'intérêt non seulement pour la Bibliothèque de l'Assemblée nationale, mais aussi pour ses usagères et ses usagers. Sa disponibilité, sans frais, et sa rapidité d'exécution soulèvent des questions sur l'évolution des besoins d'information de ses clientèles. À la manière de Wikipédia et même de Google à une certaine époque, ChatGPT a su créer une onde de choc dans plusieurs milieux, notamment celui des institutions de documentation comme la Bibliothèque.

ChatGPT a tendance à générer des informations fausses ou trompeuses, les différents tests en témoignent. La présentation logique des idées porte toutefois à croire que l'information est exacte. À cet effet, le Groupe de travail recommande à la Bibliothèque de l'Assemblée nationale de réfléchir à son positionnement de la Bibliothèque de l'Assemblée quant à son rôle de pourvoyeur d'information fiables et vérifiables auprès de sa clientèle dans le contexte des technologies de pointe.

Recommandation 1

Poursuivre la réflexion sur le positionnement de la Bibliothèque de l'Assemblée nationale sur son rôle de documentaliste dans le contexte des nouvelles technologies.


Par ailleurs, les usagères et les usagers de la Bibliothèque de l'Assemblée sont susceptibles d'avoir des questions relatives à l'utilisation de ChatGPT. À cet effet, plusieurs membres de l'effectif ne se croient pas suffisamment outillés pour répondre à de très probables questions sur cet environnement. À leur avis, une formation devrait être donnée au personnel afin de mieux comprendre ChatGPT. Une formation devrait également être offerte à la clientèle de la Bibliothèque pour les sensibiliser aux limites de l'outil, mais aussi sur les bonnes pratiques en matière de recherche d'information.

Recommandation 2

Donner une formation au personnel de la Bibliothèque de l'Assemblée nationale sur l'usage de ChatGPT.

Recommandation 3

Offrir une formation aux usagères et aux usagers de la Bibliothèque sur les bonnes pratiques de recherche d'information de qualité sur ChatGPT.



Dans l'éventualité où la Bibliothèque ferait l'acquisition d'un logiciel de traitement du langage naturel comme ChatGPT, les tâches de son personnel seront appelées à évoluer. Au-delà des impératifs liés à la gestion du changement propre à tout type d'implantation technologique majeure, une réévaluation des tâches des employées et des employés actuels est essentielle. Pour certains, la complexité des tâches à réaliser risque d'être nettement augmentée par ChatGPT. Pour d'autres, la charge de travail ou le type de tâche risque d'être réduit ou simplifié. Au cours des entrevues semi-dirigées, plusieurs se sont dit inquiets de voir réduire leur poste à une simple responsabilité de validation de contenu. À cet égard, le Groupe de travail recommande de conserver une place prioritaire au personnel dans la réflexion sur la modernisation des tâches. De plus, une réévaluation du niveau de complexité des tâches et des postes touchés par l'outil choisi devra être réalisée.

Recommandation 4

Mettre le bien-être du personnel au cœur de la réflexion sur la modernisation des tâches menée avec l'intelligence artificielle.

Soulignons que la Bibliothèque de l'Assemblée nationale ne détient pas d'expertise technique particulière en matière d'intelligence artificielle. À l'heure actuelle, un seul employé est responsable des projets en technologies de l'information et des communications. De plus, une équipe composée de quatre personnes est chargée de la gestion intégrée de document. Or, l'intelligence artificielle repose sur une technologie complexe qui requiert une expertise avancée en informatique. Pour assurer la pérennité de la mise en œuvre d'une technologie de la sorte tout en évitant les bris de services, plus d'une ressource devrait être affectée exclusivement aux projets technologiques de la Bibliothèque. Entre autres, un ou des spécialistes en informatique et en intelligence artificielle devraient être embauchés.


Recommandation 5

Donner aux personnes responsables une formation avancée en gestion de projets technologiques reposant sur l'intelligence artificielle.

Recommandation 6

Embaucher une ou plusieurs ressources supplémentaires spécialisées dans les domaines de l'informatique et de l'intelligence artificielle.

L'acquisition et l'intégration de technologies basées sur l'intelligence artificielle soulèvent des questions de nature humaine, mais aussi éthique et juridique. Le patrimoine informationnel détenu par la Bibliothèque de l'Assemblée nationale est composé d'information de nature sensible, susceptible de faire intervenir un



ensemble de lois. La protection de ce patrimoine contre les usages malveillants, la cybersécurité et la pérennité des informations détenues sont essentielles et doivent entrer en jeu dans la prise de décisions.

Recommandation 7

Évaluer les enjeux éthiques et les risques pour la protection des renseignements personnels et la cybersécurité dans la démarche de réflexion pour l'acquisition d'une technologie d'intelligence artificielle.

OpenAI étant une entreprise américaine, plusieurs ont manifesté un inconfort à utiliser une technologie dont les mécanismes reposent sur l'assujettissement à des lois différentes de celles du Québec. Le flou sur l'utilisation, le traitement et l'hébergement des données à l'extérieur du Québec soulève des questions éthiques et juridiques. L'hébergement local des données et le recours aux services d'une entreprise québécoise ou canadienne devrait avoir la priorité dans la recherche d'une technologie adaptée à la Bibliothèque de l'Assemblée.

Recommandation 8

Accorder la priorité aux fournisseurs québécois et canadiens dans la sélection d'une technologie d'intelligence artificielle pour la Bibliothèque de l'Assemblée.

Annexe I – Questionnaires des groupes de discussion semi-dirigés

Scénario général

- Remerciements et accueil.
- Présentation du Groupe de travail sur l'intelligence artificielle, de ChatGPT et de l'intelligence artificielle.
- Qui connaît ChatGPT? Si oui, pouvez-vous nous dire ce que c'est?
- Réalisation d'un test avec ChatGPT.
- À votre avis, parmi les tâches que vous effectuez, lesquelles vous semblent les plus pertinentes pour une éventuelle collaboration avec un outil d'IA? Pourquoi?
- Quelles sont les tâches qui seraient impossibles à réaliser en collaboration avec l'intelligence artificielle? Pourquoi?
- À quel moment l'IA pourrait intervenir dans vos processus de travail?
- Quelles seraient vos principales craintes face à l'utilisation d'un logiciel tel que ChatGPT?
- À votre avis, quelles seraient les balises nécessaires à mettre en place pour intégrer de manière diligente l'intelligence artificielle dans nos méthodes de travail?
- Quelles seraient vos principales attentes par rapport à l'utilisation de l'intelligence artificielle dans vos tâches courantes? (ex. : formation, intégration graduelle, sécurisation d'emploi, bonification des tâches.)
- De manière générale, seriez-vous à l'aise, autant du point de vue technique que du point de vue éthique, d'utiliser l'intelligence artificielle dans le cadre de votre travail?

Question spécifique au Service de l'information

- Considérant que votre rôle est de fournir de l'information fiable et objective aux usagères et aux usagers, de quelles façons percevez-vous l'apport de l'IA dans les méthodes de recherche?

Questions spécifiques à l'Équipe du développement des collections et du traitement documentaire

- Dans le contexte de vos activités de catalogage, l'intelligence artificielle pourrait-elle bonifier votre travail actuel? De quelle manière?

- › Croyez-vous que votre travail en traitement documentaire pourrait être bonifié? De quelle manière?

Questions spécifiques au Service de la recherche

- › Considérant que votre rôle est de fournir de l'information fiable et objective aux usagères et aux usagers, de quelles façons percevez-vous l'apport de l'IA dans les méthodes de recherche?
- › Seriez-vous confortable à l'idée d'utiliser et de citer un résultat obtenu par ChatGPT?

Annexe II – Recension des tâches des services de la Bibliothèque

Administration de la Bibliothèque de l'Assemblée

Technicienne en administration

Assister la direction dans la coordination des activités de la Direction de la Bibliothèque

- › Assurer la continuité des communications au sein des équipes de gestion et des employés
- › Rédaction de procès-verbaux

Suivi et gestion des échéanciers

- › Préparation d'un calendrier périodique de reddition de comptes institutionnel ainsi que du calendrier de production du rapport annuel de la direction

Réaliser les tâches administratives courantes liées à la Direction de la Bibliothèque :

- › Suivi de l'agenda, de la boîte de réception des courriels et de la correspondance de la directrice
- › Saisie, mise en forme et rédaction de divers documents
- › Suivi avec la technicienne en administration des demandes de paiement de factures
- › Suivi mensuel avec la Direction des ressources humaines
- › Mise à jour de bases de données
- › Élaborer des tableaux et rapports statistiques (coûts, achats, inventaire, statistiques, bilans, etc.)
- › Vérifier la conformité des dépenses encourues et préparer les documents pour le paiement des factures
- › Suivi des échéanciers de paiement

Effectuer les tâches administratives en appui aux employés de la direction

- › Assurer le contrôle de l'assiduité des employés
- › Venir en appui dans la gestion de certains fichiers tableur
- › Élaborer des outils de suivi

Réaliser des activités de suivi budgétaire

- › Recueillir les données financières pour la préparation de la revue de programme annuelle de la direction
- › Effectuer le suivi budgétaire et, à cette fin, analyser et préparer les données nécessaires à la confection et l'élaboration de rapports, tableaux et graphiques qui sont présentés aux autorités de l'Assemblée
- › Effectuer le suivi des dossiers relatifs à SAGIR et à la gestion des ressources humaines

Service des collections

Équipe des archives et de la numérisation

L'Équipe est composée d'analystes en gestion intégrée des documents, de techniciens en documentation, d'archivistes, d'agents de bureau et d'une technicienne principale en arts appliqués et graphismes. Leurs tâches se répartissent selon les secteurs suivants :

Gestion documentaire

- › Effectuer la recherche d'informations à la demande des usagers
- › Faire la classification, la codification et le tri des documents en fonction du calendrier de conservation
- › Élaborer et maintenir à jour le recueil de classification, des procédures et des politiques
- › Offrir un soutien technique pour les outils informatiques SharePoint et Ultima

Archives

- › Effectuer la recherche d'informations à la demande des usagers
- › Faire la classification et la codification des documents
- › Description des fonds d'archives et identification des objets patrimoniaux et des photographies
- › Trier la masse documentaire selon le calendrier de conservation et des procédures
- › Diffuser des archives (expositions, balados, médias sociaux, etc.)

Numérisation et graphisme

- › Effectuer la reconnaissance optique de caractères (ROC) des documents numérisés
- › Effectuer la recherche de visuels
- › Élaborer et mettre en forme des formulaires
- › Effectuer la numérisation de documents

Équipe du développement des collections

L'Équipe est composée de bibliothécaires, de techniciennes et d'une agente de bureau. Leurs tâches se répartissent selon les secteurs suivants :

Acquisition

- › Repérer des titres pour certaines catégories de documents
- › Faire le suivi des abonnements et des commandes
- › Mettre à jour la bibliographie des publications sériées et des périodiques
- › Effectuer la facturation des commandes
- › Faire l'élégage des publications gouvernementales et de la collection générale

Traitement documentaire

- › Effectuer une veille des documents électroniques
- › Effectuer un catalogage des publications gouvernementales et de la collection générale
- › Effectuer une classification des documents

Indexation

- › Faire l'indexation de la collection générale (attribution de vedettes-matière)
- › Faire l'indexation des séances en Chambre et en commissions parlementaires (analyse du texte et création de rubriques à l'aide du thésaurus)
- › Produire la liste et l'index des documents déposés en Chambre
- › Rédiger et mettre à jour les politiques et procédures
- › Réaliser l'index de certains ouvrages

Service de l'information

Agent et agente de bureau

Contrôler la circulation des documents de la Bibliothèque

- › Enregistrer les prêts, les retours et les réservations des documents
- › Mettre à jour les listes d'abonnés, envoyer les avis de retards, compiler les statistiques de prêts
- › S'assurer quotidiennement du bon fonctionnement et du bon usage du système informatique du prêt en maîtrisant l'ensemble des fonctionnalités du système
- › S'assurer de la transmission des documents aux usagers par Poste Canada ou messagerie

Service aux usagères et usagers

- › Accueillir les visiteurs externes. Recevoir et diriger, s'il y a lieu, les usagers et les visiteurs vers les personnes ou les services appropriés
- › Effectuer des travaux de photocopie et de numérisation pour différentes clientèles
- › Effectuer le traitement des communications téléphoniques vers les personnes ou les services appropriés
- › Saisir et compiler des statistiques de fréquentation de la Bibliothèque

Classer les documents de la Bibliothèque

- › Gestion, tri et classement des documents consultés sur place, ceux récemment acquis et ceux de retour de la section du traitement ainsi que ceux de la photocopie ou du prêt
- › Maintenir en bon ordre les collections désignées et effectuer quotidiennement la vérification du classement des documents
- › Trier, classer et enregistrer (Kardex) les journaux courants, effectuer sans délai les rappels auprès des fournisseurs de journaux
- › Vérifier régulièrement les indications sur l'emplacement des collections et les corriger au besoin
- › Tenir quotidiennement les statistiques sur les documents reçus et consultés

- › Faire de la lecture de rayons, participer au déplacement et élagage des collections

Technicien en documentation

Répondre aux demandes d'information et de recherche (en personne, par téléphone ou par courriel)

- › Effectuer une garde au comptoir de référence et répondre ou redistribuer les demandes d'information ou de recherche
- › Effectuer les recherches dans le catalogue de la Bibliothèque et dans toutes les sources documentaires pertinentes : bases de données, index, répertoires, et ce, dans tous les types de collections
- › Répondre aux demandes de prêt entre bibliothèques
- › Fournir une assistance aux clientèles prioritaires et aux usagers externes qui veulent utiliser les collections de la Bibliothèque (public, étudiants, chercheurs, etc.) :
 - Orienter les usagers vers les sources documentaires susceptibles de les aider
 - Aider à l'interrogation du catalogue et des bases de données offertes en libre accès
 - Aider à la recherche dans le site web de l'Assemblée et celui de la Bibliothèque, les index et répertoires et les autres outils de recherche
- › Remplir, pour chaque demande, le formulaire électronique de gestion des demandes de référence
- › Préparer des guides, des outils de référence pour l'aide à la recherche ou la mise en valeur des collections
- › Assister le responsable du développement des collections et les bibliothécaires dans la gestion des collections en faisant du classement, de la lecture de rayons, du choix, de l'élagage, des mises à jour ou de la gestion de reliure ou en s'occupant de la redistribution des documents élagués
- › Effectuer des visites de la Bibliothèque pour des groupes spécialisés, le personnel politique et administratif ou toutes autres personnes désignées
- › Seconder les préposés à l'accueil, à la circulation et à l'organisation des collections dans l'accueil des visiteurs, la circulation des documents, la réception et la redirection des communications téléphoniques de la Bibliothèque

Bibliothécaire

Recevoir et répondre aux demandes d'information et de recherche documentaire

- › Faire l'entrevue de référence pour évaluer la complexité de la demande soumise
- › Concevoir et préparer la stratégie de recherche la plus appropriée à la demande
- › Récupérer, rassembler et transmettre au demandeur dans les meilleurs délais et par les moyens les plus efficaces la documentation répondant à sa demande.
- › Assister et/ou former les usagers dans la consultation des sources documentaires disponibles à la Bibliothèque ou en ligne.
- › Choisir les nouveaux documents pour le développement des collections de la Bibliothèque.

Contribuer à la préparation et au suivi de veilles informationnelles sur différents sujets touchant le parlementarisme et la politique québécoise

- › Assumer la responsabilité de surveillance de certains thèmes en scrutant la production documentaire courante à l'aide d'outils logiciels, par des lectures de périodiques (revues, journaux, bulletins imprimés ou numériques) ou par l'examen de sites web
- › Contribuer au développement des ressources documentaires de la Bibliothèque en participant au choix et à l'élégage des collections
- › Produire de la documentation pertinente aux besoins des usagers prioritaires telles que des bibliographies thématiques ou des listes d'acquisitions récentes
- › Préparer des guides sur l'utilisation des sources documentaires disponibles ou en usage à la Bibliothèque
- › Contribuer à développer des outils de diffusion web de la Bibliothèque tels que le site web, des banques de données ou autres

Service de la recherche

Le Service de la recherche de la Bibliothèque de l'Assemblée nationale est composé d'un ensemble d'analystes spécialisés dans différents domaines. Il effectue principalement des travaux de recherche et d'analyse pour les personnes élues et leur personnel politique, les commissions parlementaires, les délégations parlementaires, le personnel administratif de l'Assemblée, le lieutenant-gouverneur et son personnel ainsi que les personnes désignées par l'Assemblée.

Le Service de la recherche contribue aussi au développement et à la diffusion des connaissances sur l'histoire et l'évolution du parlementarisme et des institutions parlementaires.

Analystes

Rédiger divers types de documents (notes d'information et d'allocution)

- › Analyser les rapports annuels de gestions et les plans stratégiques des ministères et organismes québécois
- › Analyse, interprétation et traitement de sources divers (bases de données, lois, rapports)

Soutenir les travaux des commissions parlementaires

- › Assister aux séances de travail et aux auditions publiques des commissions parlementaires de l'Assemblée, l'étude de pétitions, les mandats conférés par une loi, les mandats d'initiative, les mandats d'imputabilité et la surveillance d'organismes publics
- › Concevoir un questionnaire pour une consultation en ligne et analyser les réponses
- › Rédiger des documents de soutien (vulgarisation d'un sujet, proposition de questions sur des enjeux qui mériteraient d'être abordés lors du processus de consultation, etc.)
- › Rédiger des rapports (synthèse des auditions, formulation d'observations, de conclusions et de recommandations, etc.)

- › Résumer des mémoires

Participer à l'organisation de divers événements (tables rondes, Prix du livre politique, etc.)

- › Soutenir la coordination de l'événement et des acteurs impliqués
- › Rédiger des notes d'allocution
- › Proposer un thème, des personnes pour l'animation et des intervenants

Historiens

- › Recherche et développement des connaissances en histoire parlementaire
- › Rédaction de notes d'information, d'articles scientifiques et d'ouvrages
- › Développement et mise à jour de l'Encyclopédie du parlementarisme
- › Services-conseils en patrimoine
- › Diffusion de la connaissance lors d'événements scientifiques ou à l'occasion de rencontres de groupes

Révision linguistique

- › Effectuer la révision de documents produits par les analystes et les historiens

Technicienne de recherche

- › Faire le suivi et l'archivage des travaux
- › Compiler des statistiques sur les travaux effectués par le Service
- › Faire la mise en forme de documents
- › Mettre à jour la liste des personnes entendues en commission depuis les dernières législatures
- › Mettre à jour le tableau des mandats conférés par des lois et le tableau des organismes et ministères soumis au pouvoir de surveillance des commissions parlementaires
- › Extraire des données du système électronique d'appel d'offres du gouvernement

Annexe III – Grilles d'évaluation

Service des collections

Catalogage

Critères	Note	Explications
Validité des éléments de réponse Les éléments de la réponse générés par ChatGPT sont-ils véridiques? La réponse est-elle fiable et valide? Validité de l'encodage MARC et cohérence avec les RDA.		
Pertinence de la réponse La réponse obtenue correspond-elle au résultat attendu? ChatGPT répond-il adéquatement aux questions ? Son niveau de connaissance dans la discipline est-il suffisant pour générer une réponse satisfaisante? La réponse est-elle exhaustive? Les éléments d'encodage fournis sont-ils complets?		
Performance et rapidité du processus ChatGPT répond-il aux commandes rapidement et adéquatement à ce qui lui est demandé? Faut-il le relancer à plusieurs reprises pour obtenir une réponse satisfaisante? Le travail de vérification est-il rapide?		

Indexation

Critères	Note	Explications
Validité des éléments de réponse Les éléments de la réponse générés par ChatGPT sont-ils véridiques? La réponse est-elle fiable et valide?		
Pertinence de la réponse La réponse obtenue correspond-elle au résultat attendu? ChatGPT répond-il adéquatement aux questions ? Son niveau de connaissance dans la discipline est-il suffisant pour générer une réponse satisfaisante?		
Neutralité de la réponse Les termes utilisés par ChatGPT ont-ils une connotation péjorative ou stéréotypées? Formule-t-il une opinion en répondant à ce qui lui est demandé?		
Performance et rapidité du processus ChatGPT répond-il aux commandes rapidement et adéquatement à ce qui lui est demandé? Faut-il le relancer à plusieurs reprises pour obtenir une réponse satisfaisante? Le travail de vérification est-il rapide?		
Structuration de la réponse Le texte est-il bien structuré (paragraphe, titres de section, utilisation de puces ou de numéros etc.)? La réponse comporte-t-elle une introduction, un développement, une conclusion (si applicable)?		
Qualité de la langue Appréciation de la qualité du vocabulaire, de la syntaxe et des expressions utilisés par ChatGPT. Le niveau de langue utilisé est-il approprié? ChatGPT fait-il des fautes de français dans sa réponse?		

Service de l'information

Critères	Note	Explications
Validité des éléments de réponse Les éléments de la réponse générés par ChatGPT sont-ils véridiques? La réponse est-elle fiable et valide?		
Pertinence de la réponse La réponse obtenue correspond-elle au résultat attendu? ChatGPT répond-il adéquatement aux questions ? Son niveau de connaissance dans la discipline est-il suffisant pour générer une réponse satisfaisante?		
Cohérence des propos Le texte est-il cohérent? Y-a-t-il un enchaînement logique dans les propos de ChatGPT?		
Neutralité de la réponse Les termes utilisés par ChatGPT ont-ils une connotation péjorative ou stéréotypées? Formule-t-il une opinion en répondant à ce qui lui est demandé?		
Performance et rapidité du processus ChatGPT répond-il aux commandes rapidement et adéquatement à ce qui lui est demandé? Faut-il le relancer à plusieurs reprises pour obtenir une réponse satisfaisante? Le travail de vérification est-il rapide?		
Structuration de la réponse Le texte est-il bien structuré (paragraphes, titres de section, utilisation de puces ou de numéros etc.)? La réponse comporte-t-elle une introduction, un développement une conclusion (si applicable)?		
Qualité de la langue Appréciation de la qualité du vocabulaire, de la syntaxe et des expressions utilisés par ChatGPT. Le niveau de langue utilisé est-il approprié? ChatGPT fait-il des fautes de français dans sa réponse?		

Service de la recherche

Critères	Note	Explications
Validité des éléments de réponse Les éléments de la réponse générés par ChatGPT sont-ils véridiques? La réponse est-elle fiable et valide?		
Pertinence de la réponse La réponse obtenue correspond-elle au résultat attendu? ChatGPT répond-il adéquatement aux questions ? Son niveau de connaissance dans la discipline est-il suffisant pour générer une réponse satisfaisante? La réponse est-elle exhaustive? Les éléments d'encodage fournis sont-ils complets?		
Cohérence des propos Le texte est-il cohérent? Y-a-t-il un enchainement logique dans les propos de ChatGPT?		
Neutralité de la réponse Les termes utilisés par ChatGPT ont-ils une connotation péjorative ou stéréotypées? Formule-t-il une opinion en répondant à ce qui lui est demandé?		
Performance et rapidité du processus ChatGPT répond-il aux commandes rapidement et adéquatement à ce qui lui est demandé? Faut-il le relancer à plusieurs reprises pour obtenir une réponse satisfaisante? Le travail de vérification est-il rapide?		
Structuration de la réponse Le texte est-il bien structuré (paragraphes, titres de section, utilisation de puces ou de numéros etc.)? La réponse comporte-t-elle une introduction, un développement une conclusion (si applicable)?		
Qualité de la langue Appréciation de la qualité du vocabulaire, de la syntaxe et des expressions utilisés par ChatGPT. Le niveau de langue utilisé est-il approprié? ChatGPT fait-il des fautes de français dans sa réponse?		

Annexe IV – Répertoire des essais

Service	Test	Essai
Collections	Catalogage	Monographie – Publication gouvernementale
		Monographie – Collection générale
		Monographie – Collection spéciale
		Périodique (terminé)
		Périodique (courant)
		Étude des crédits
		Mémoire déposé en commission parlementaire
		Numéro de publication en série traitée en monographie
		Description analytique – 1
		Description analytique – 2
		Enregistrement vidéo, diffusion en continu
		Enregistrement vidéo DVD
		Site web
		Monographie (RCAA2)
		Numérisation (RCAA2)
	Indexation	Déclaration de député – Semaine thématique
		Déclaration de député – Jour commémoratif
		Déclaration de députés – Mot de félicitations
		Dépôt de pétition – Encadrer l'usage du cannabis
		Dépôt de pétition – Amélogénèse
		Dépôt de pétition – Patrimoine du mont Saint-Hilaire
		Présentation d'un projet de loi – Régime de retraite
		Présentation d'un projet de loi – Neutralité religieuse
		Présentation d'un projet de loi – Modification du sexe sur l'acte de naissance
		Questions et réponses orales – Conditions de vie des Autochtones
		Questions et réponses orales – Gaz de schiste et acceptabilité sociale
		Questions et réponses orales – Maintien à domicile des aînés et crédit d'impôt
		Motion sans préavis – Libération de Raïf Badawi
		Motion sans préavis – Accès à l'avortement
		Motion sans préavis – Nomination de juges à la Cour supérieure du Québec

Information	Recherche d'information	Demande de renseignements sur l'adoption d'un article d'une loi – Niveau 2
		Recherche d'un article publié dans la revue Protégez-vous – Niveau 2
		Recherche d'un article sur Pascal Bérubé dans Le Voir en 2001 – Niveau 2
		Recherche de subventions sur une résidence pour personnes âgées – Niveau 3
		Recherche de subventions pour un tournoi de hockey scolaire – Niveau 3
		Bibliographie sur la conciliation travail-famille Québec-Canada – Niveau 3
		Recherche d'information sur la dissolution de l'Assemblée et les projets de loi au feuillet – Niveau 3
		Bibliographie sur les centres d'orientation, de francisation et d'intégration – Niveau 4
		Bibliographie sur les changements climatiques – Niveau 4
		Recherche comparative sur le taux de suicide et les pratiques de prévention – Niveau 4
		Recherche comparative sur les processus d'étude des crédits dans les parlements canadiens et de tradition britannique – Niveau 4
		Recherche d'information factuelle sur la politique publique ou les lois touchant le développement régional – Niveau 4
		Recherche comparative sur les parlements et les médias sociaux – Niveau 5
		Recherche comparative sur le pouvoir d'enquête de l'Ombudsman au Canada – Niveau 5
		Bibliographie sur la ventilation, changement d'air et COVID-19 – Niveau 5
	Développement des collections	Livres – Politique québécoise
		Livres – Environnement
		Livres - Santé
Recherche	Définitions courtes	Législature
		Acte de Québec
		Gerrymandering
		Baïllon
		Édifice Jean-Antoine-Panet
		Citadelle de Québec

Recherche	Définitions longues	Législature
		Acte de Québec
		Gerrymandering
		Bâillon
		Édifice Jean-Antoine-Panet
		Citadelle de Québec
	Rédaction d'un plan	Gestation pour autrui
		Télémédecine
		Voitures autonomes
		Exploitation sexuelle des jeunes Autochtones
	Rédaction d'une note d'information	Gestation pour autrui
		Télémédecine
		Voitures autonomes
		Exploitation sexuelle des jeunes Autochtones
	Synthèse d'information	Résumé d'un mémoire déposé en commission parlementaire - Pesticides
		Résumé d'un mémoire déposé en commission parlementaire - Recyclage du verre



**Notre
maison
citoyenne**

assnat.qc.ca