

CANADA
PROVINCE DE QUÉBEC
DISTRICT DE MONTRÉAL

(Chambre des actions collectives)
COUR SUPÉRIEURE

N° 500-06-001414-255

ANNE ROBILLARD, domiciliée et résidant
au 501-400, rue Serge-Pépin à Beloeil,
province de Québec, J3G 6Z2

Demanderesse

c.

OPENAI, INC., personne morale dûment
constituée, ayant un établissement au 3180,
18th Street, San Francisco, Californie,
États-Unis, 94110 et

et

OPENAI GP, L.L.C., personne morale
dûment constituée, ayant un établissement
au 3180, 18th Street, San Francisco,
Californie, États-Unis, 94110

et

OPENAI GLOBAL, LLC, personne morale
dûment constituée, ayant un établissement
au 3180, 18th Street, San Francisco,
Californie, États-Unis, 94110

et

OPENAI OPCO, LLC, personne morale
dûment constituée, ayant un établissement
au 3180, 18th Street, San Francisco,
Californie, États-Unis, 94110

et

OPENAI HOLDINGS, LLC, personne
morale dûment constituée, ayant un
établissement au 3180, 18th Street, San
Francisco, Californie, États-Unis, 94110

et

OPENAI, L.L.C., personne morale dûment
constituée, ayant un établissement au 3180,
18th Street, San Francisco, Californie,
États-Unis, 94110

et

OAI CORPORATION, personne morale dûment constituée, ayant un établissement au 3180, 18th Street, San Francisco, Californie, États-Unis, 94110

et

OPENAI STARTUP FUND I, L.P., société dûment constituée, ayant un établissement au 3180, 18th Street, San Francisco, Californie, États-Unis, 94110

et

OPENAI STARTUP FUND GP I, L.L.C., personne morale dûment constituée, ayant un établissement au 3180, 18th Street, San Francisco, Californie, États-Unis, 94110

et

OPENAI STARTUP FUND MANAGEMENT, LLC, personne morale dûment constituée, ayant un établissement au 3180, 18th Street, San Francisco, Californie, États-Unis, 94110

et

MICROSOFT CORPORATION, personne morale dûment constituée, ayant son siège social au 1, Microsoft Way, Redmond, Washington, États-Unis, 98052

et

MICROSOFT CANADA INC., personne morale dûment constituée, ayant un établissement au 450-2000, av. McGill College, à Montréal, province de Québec, H3A 3H3

Défenderesses

**DEMANDE POUR AUTORISATION D'EXERCER UNE ACTION COLLECTIVE ET
POUR ÊTRE REPRÉSENTANTE**
(Art. 574 et suivants CPC)

**À L'UN DES HONORABLES JUGES DE LA COUR SUPÉRIEURE, LA
DEMANDERESSE EXPOSE RESPECTUEUSEMENT CE QUI SUIT :**

1. APERÇU

1. La demanderesse sollicite l'autorisation d'exercer une action collective pour le compte des personnes faisant partie du groupe ci-après décrit, dont elle est elle-même membre, à savoir :

Toute personne au Québec titulaire du droit d'auteur sur une œuvre littéraire ou dramatique que les défenderesses ont reproduite sans autorisation dans le cadre du développement ou de la formation de grands modèles de langage (en anglais, « *large language models* »).

2. Le droit d'auteur sert l'intérêt public en stimulant la création et la diffusion d'œuvres artistiques et intellectuelles, tout en garantissant aux créateurs une rémunération équitable pour leur travail. La réalisation de ces objectifs est facilitée par l'existence, à l'échelle mondiale, de régimes législatifs et d'une convention internationale qui accordent des droits exclusifs aux créateurs et interdisent à toute personne d'exploiter leurs œuvres sans leur autorisation.
3. Les membres du groupe sont des auteurs et autres titulaires de droit d'auteur sur des œuvres littéraires et dramatiques. Ils n'ont jamais consenti à ce que les défenderesses se servent de leurs œuvres à des fins commerciales ni à toute autre fin.
4. Les défenderesses ont développé différents grands modèles de langage (en anglais, « *large language models* ») (« **LLM** »), y compris, mais sans s'y limiter la série de modèles connus sous le nom de « GPT », issue de leur partenariat, ainsi que le modèle connu sous le nom de « Megatron-Turing NLG 530B » (« **MT-NLG** »), développé indépendamment par les défenderesses Microsoft.
5. Plutôt que d'agir dans le cadre de la loi et dans le respect des droits des membres du groupe, les défenderesses ont délibérément choisi de former leurs LLM à partir d'ensembles de données contenant des copies illicites d'œuvres du monde entier, dont celles des membres du groupe.
6. Ce faisant, les défenderesses ont commis ce qui est sans aucun doute l'un des plus grands actes de piratage de l'histoire, exploitant à des fins commerciales des millions d'œuvres dans le but de développer des outils d'intelligence artificielle générative (« **IA générative** ») susceptibles de remplacer les créateurs dont elles ont usurpé le travail.
7. Les LLM des défenderesses n'existeraient pas sans le vaste corpus d'œuvres protégées par le droit d'auteur qui ont servi à les former. Or, ces mêmes LLM mettent aujourd'hui en péril la capacité des membres du groupe à gagner leur vie.

8. En effet, différents outils IA basés sur les LLM des défenderesses permettent à n'importe qui n'importe où de générer des textes pour lesquels ils paieraient autrement des auteurs.
9. En outre, ces outils IA permettent de produire des œuvres dérivées non autorisées, c.-à-d., du matériel basé sur, imitant, résumant ou paraphasant les œuvres protégées par le droit d'auteur des membres du groupe, et nuisant au marché de ces dernières et à leur exploitation normale.
10. La demanderesse, en son nom et pour le compte des membres du groupe, demande donc l'autorisation d'exercer une action collective en dommages-intérêts compensatoires et en restitution de profits ou, au choix, en dommages-intérêts préétablis, de même qu'en dommages-intérêts punitifs et en injonction.

2. LES PARTIES

2.1. La demanderesse

11. La demanderesse est une autrice à succès dont plusieurs livres ont figuré sur les listes de *best-sellers*.
12. Elle est considérée comme une figure incontournable de la littérature fantastique.
13. La demanderesse est connue notamment pour la saga *Les Chevaliers d'Émeraude*, dont le premier tome est paru en 2002. Cette série culte composée de 13 tomes a été vendue à plus de 5 millions d'exemplaires au Québec et en France et traduite en plusieurs langues.
14. Outre cette série, la demanderesse a écrit de nombreux autres livres, dont les séries *Les héritiers d'Enkidiev* (12 tomes), *Les Chevaliers d'Antarès* (12 tomes), *Légendes d'Ashur-Sîn* (7 tomes), *Les magiciens d'Enkidiev* (7 tomes), *Les Wilder* (3 tomes), *A.N.G.E.* (10 tomes), *Les ailes d'Alexanne* (10 tomes), *Les cordes de cristal* (10 épisodes), *Le retour de l'oiseau-tonnerre* (3 tomes), *La malédiction des Dragensblöt* (7 tomes), *Maynard Bennett, enquêteur de l'étrange* (5 tomes) et le roman *Les Chevaliers d'Épées*, tel qu'il apparaît de la liste des livres d'Anne Robillard publiée sur le site Web anne-robillard.com, **PIÈCE P-1**.
15. En reconnaissance de son succès, la demanderesse a notamment reçu le *Grand Prix littéraire Archambault* pour le cinquième tome de *Les Chevaliers d'Émeraude*, ainsi que le *Prix des lecteurs du Salon du livre de Trois-Rivières* en 2007.
16. La demanderesse est titulaire des droits d'auteur et moraux sur l'ensemble de son œuvre.

2.2. Les défenderesses et leurs outils IA

2.2.1. OpenAI

17. Les défenderesses OpenAI, Inc., OpenAI GP, L.L.C., OpenAI Global, LLC, OpenAI OpCo, LLC, OpenAI Holdings, LLC, OpenAI, L.L.C., OAI Corporation, OpenAI Startup Fund I, L.P., OpenAI Startup Fund GP I, L.L.C. et OpenAI Startup Fund Management, LLC sont une série de sociétés étroitement liées qui font partie d'une structure corporative complexe et opaque créée afin de développer, financer, exploiter et commercialiser des technologies d'intelligence artificielle avancées (collectivement, « **OpenAI** »). Ces sociétés sont toutes constituées en vertu des lois du Delaware et ont toutes un établissement à San Francisco, en Californie.
18. La défenderesse OpenAI, Inc. est une société à but non lucratif. Elle est propriétaire et contrôle, directement ou indirectement, les autres sociétés du groupe. Elle détient donc le pouvoir de gouvernance ultime sur les activités d'OpenAI.
19. La défenderesse OpenAI GP, L.L.C. est une société à responsabilité limitée qui est entièrement détenue et contrôlée par OpenAI, Inc. Elle assume le contrôle et la gestion de l'entreprise au quotidien.
20. La défenderesse OpenAI Global, LLC, est une société à responsabilité limitée, dans laquelle Microsoft détient un intérêt économique minoritaire. Il s'agit d'une entité opérationnelle d'OpenAI contrôlée par OpenAI GP, L.L.C., tel qu'il appert d'un extrait du site Web openai.com, **PIÈCE P-2**.
21. La défenderesse, OpenAI OpCo, LLC, est une société à responsabilité limitée, dont la gestion est assumée par OpenAI Global, LLC. Il s'agit d'une entité opérationnelle d'OpenAI.
22. La défenderesse OpenAI Holdings, LLC est une société à responsabilité limitée, dont la gestion est assumée par OpenAI GP, L.L.C.
23. La défenderesse OpenAI, L.L.C. est une société à responsabilité limitée, dont la gestion est assumée par OpenAI GP, L.L.C. Il s'agit d'une entité opérationnelle d'OpenAI.
24. La défenderesse OAI Corporation est une société à responsabilité limitée.
25. La défenderesse OpenAI Startup Fund I, L.P. est une société en commandite.
26. La défenderesse OpenAI Startup Fund GP I, L.L.C. est une société à responsabilité limitée. Agissant à titre de commanditée d'OpenAI Startup Fund I, L.P., elle en assume la gestion ainsi que l'exploitation de ses affaires et activités courantes.

27. La défenderesse OpenAI Startup Fund Management, LLC, est une société à responsabilité limitée.
28. En partenariat avec Microsoft, OpenAI a développé des LLM, dont la série GPT, ainsi que des outils IA dérivés, dont l'assistant IA « ChatGPT », qui reposent sur ces modèles.
29. ChatGPT est accessible au public via une interface Web et application mobile, avec plusieurs niveaux d'abonnement, offrant des fonctionnalités et limites d'usage croissantes selon le plan choisi.
30. ChatGPT est disponible au Québec et dans l'ensemble du Canada.
31. OpenAI propose également des interfaces de programmation (« **API** ») payantes qui permettent aux entreprises et aux développeurs d'intégrer directement ses LLM dans leurs propres applications ou produits.
32. OpenAI a également des partenariats commerciaux spécifiques par lesquels elle licence certains LLM ou produits dérivés en exclusivité ou pour des usages spécifiques.
33. En juillet 2025, OpenAI rapportait que ses produits ChatGPT comptaient environ 700 millions d'utilisateurs actifs hebdomadaires. Il était également estimé qu'elle générerait environ 12 milliards de dollars US en revenus annualisés, le tout tel qu'il appert de l'article intitulé « OpenAI hits \$12 billion in annualized revenue, The Information reports » daté du 30 juillet 2025 et publié sur le site Web [reuters.com](https://www.reuters.com), **PIÈCE P-3**.

2.2.2. Microsoft

34. La défenderesse Microsoft Corporation est une société incorporée dans l'État de Washington dont le siège social est situé à Redmond, dans l'État de Washington. Elle opère notamment le Microsoft Research Montréal Lab, un établissement de recherche IA situé au 400-6795, rue Marconi, Montréal, H2S 3J9, tel qu'il appert d'extraits du site Web [microsoft.com](https://www.microsoft.com), **PIÈCE P-4**.
35. La défenderesse Microsoft Canada Inc. est une société incorporée en Nouvelle-Écosse dont le siège social est situé à Halifax et ayant plusieurs établissements au Québec. Elle est entièrement détenue par Microsoft Corporation.
36. Les activités de Microsoft Corporation et Microsoft Canada Inc. (collectivement, « **Microsoft** ») sont inextricablement liées aux fins du développement de produits d'IA générative.

37. Depuis 2019, Microsoft a investi plus de 13 milliards de dollars US dans OpenAI et elle détiendrait une participation de 49 % dans les activités à but lucratif de l'entreprise.
38. Dans une entrevue accordée en 2023, le PDG de Microsoft, Satya Nadella, a insisté sur le rôle essentiel joué par Microsoft dans le cadre de ce partenariat avec OpenAI, soulignant son implication directe dans les opérations :
- [...] We have all the IP rights and all the capability.
- If OpenAI disappeared tomorrow, I don't want any customer of ours to be worried about it quite honestly, because we have all of the rights to continue the innovation. Not just to serve the product, but we can go and just do what we were doing in partnership ourselves. We have the people, we have the compute, we have the data, we have everything. [...]
- And that gives us significant rights as I said. And also this thing, it's not hands off, right? We are in there. We are below them, above them, around them. We do the kernel optimizations, we build tools, we build the infrastructure. So that's why I think a lot of the industrial analysts are saying, "Oh wow, it's really a joint project between Microsoft and OpenAI." The reality is we are, as I said, very self-sufficient in all of this.
- [...] After all, we had many engineers from Microsoft who are also working with OpenAI [...]
- Article intitulé « Satya Nadella on hiring the Most Powerful Man in AI When OpenAI threw Sam Altman overboard, Microsoft's CEO saw an opportunity » date du 21 novembre 2023 et publié sur le site Web nymag.com, **PIÈCE P-5** (p. 4 à 6).
39. Compte tenu de l'influence et de l'implication étroite de Microsoft dans les activités d'OpenAI, ainsi que de la collaboration avouée de ses ingénieurs avec ceux d'OpenAI, il est manifeste que Microsoft a été directement impliquée dans les contraventions alléguées aux présentes ou, à tout le moins, qu'elle les a directement autorisées.
40. En janvier 2025, Microsoft a annoncé que son partenariat avec OpenAI se poursuivrait jusqu'en 2030, tel qu'il appert d'un communiqué de Microsoft intitulé « Microsoft and OpenAI evolve partnership to drive the next phase of AI » daté du 21 janvier 2025 et publié sur le site Web blogs.microsoft.com, **PIÈCE P-6**.
41. Ce partenariat comprend notamment des ententes de partage de revenus réciproques permettant aux deux parties de tirer profits de l'utilisation des modèles développés.
42. En outre, le partenariat entre Microsoft et OpenAI a notamment mené à la création, au développement et à la maintenance de systèmes de supercalcul visant à soutenir le recherche d'OpenAI et le développement et la formation de ses

systèmes IA, tel qu'il appert notamment d'un communiqué de Microsoft intitulé « Microsoft announces new supercomputer, lays out vision for future AI work » daté du 19 mai 2020 et publié sur le site Web news.microsoft.com, **PIÈCE P-7**.

43. De fait, Microsoft agit comme partenaire infonuagique d'OpenAI, et ses services cloud ont alimenté l'ensemble des charges de travail d'OpenAI, qu'il s'agisse de produits, de services d'API ou de recherche, tel qu'il appert d'un communiqué intitulé « Microsoft and OpenAI extend partnership » daté du 23 janvier 2023 et publié sur le site Web blogs.microsoft.com et d'un article intitulé « Microsoft extends OpenAI partnership in a 'multibillion dollar investment' » daté du 23 janvier 2023 et publié sur le site Web theverge.com, **PIÈCE P-8**, en liasse.
44. En 2020, le modèle principal de GPT-3 d'OpenAI a été lancé et a fait l'objet d'une licence exclusive accordée à Microsoft.
45. En novembre 2021, Microsoft a lancé son service Azure OpenAI, intégré à sa plateforme infonuagique Azure, qui permet aux entreprises et développeurs d'accéder aux LLM d'OpenAI via des API.
46. En outre, en 2023, Microsoft a lancé sa propre série d'outils IA sous l'appellation « Copilot », reposant sur les modèles GPT développés par OpenAI. Copilot est notamment intégré à la suite Microsoft 365 (Word, Excel, Outlook, Teams, etc.), au moteur de recherche Bing, et au système d'exploitation Windows.
47. Copilot est accessible au public gratuitement via une interface web et des applications mobiles et Windows, et propose plusieurs niveaux d'abonnement payants offrant des fonctionnalités accrues et des limites d'usage élargies selon le plan choisi.
48. Copilot est disponible au Québec et dans l'ensemble du Canada.
49. En juillet 2025, Microsoft rapportait que son écosystème d'applications Copilot comptait plus de 100 millions d'utilisateurs actifs hebdomadaires, tel qu'il appert d'un communiqué de Microsoft intitulé « Microsoft Copilot Surpasses 100 Million Users: Revolutionizing Enterprise AI in 2024 » daté du 31 juillet 2025 et publié sur le site Web windowsforum.com, **PIÈCE P-9**.
50. En dehors de son partenariat avec OpenAI, Microsoft a pris des initiatives indépendantes pour développer sa propre technologie d'IA générative, y compris le LLM, MT-NLG.

3. LA RESPONSABILITÉ DES DÉFENDERESSES

3.1. Introduction

51. Les technologies d'IA permettent aux ordinateurs et machines de simuler l'intelligence humaine et ses capacités de résolution.
52. Le terme « IA générative » fait spécifiquement référence à une catégorie d'IA capable de générer différents types de contenu (par ex., texte, image, musique, vidéo, audio ou code logiciel) supposément original en réponse aux instructions données par l'utilisateur.
53. L'IA générative s'appuie sur des modèles d'apprentissage profond (en anglais, « *deep learning* »), c.-à-d., des algorithmes sophistiqués qui identifient et codent les schémas et les relations dans d'énormes quantités de données puis utilisent ces informations pour comprendre et répondre aux requêtes des utilisateurs.
54. Les LLM sont une catégorie de modèles d'apprentissage profond pouvant servir à la génération de texte.
55. Pour développer un LLM, celui-ci doit être « formé » à partir d'énormes quantités de textes et de données provenant d'Internet ou d'autres sources, communément appelées « ensembles de données d'apprentissage » (en anglais, « *training dataset* »).
56. Ce processus est décrit ainsi dans un document de consultation du Gouvernement du Canada sur le droit d'auteur et l'IA générative :

La fouille de textes et de données (FTD) consiste en la reproduction et l'analyse de grandes quantités de données et d'informations, **y compris celles extraites de contenus protégés par le droit d'auteur**, afin d'identifier des tendances et de faire des prédictions. La FTD est une étape essentielle dans la formation de modèles d'apprentissage automatique. Cette technique permet au modèle « d'apprendre » à reconnaître et à reproduire des tendances qui lui permettront d'accomplir certaines tâches, notamment de générer de la poésie, de la musique ou des œuvres d'art. Au-delà de l'IA générative, la FTD peut également faire progresser la science et aider les entreprises à résoudre des problèmes, à innover et à créer davantage de valeur.

Suite à l'essor des outils d'IA générative tels que ChatGPT et DALL-E, **plusieurs titulaires de droits ont exprimé des inquiétudes croissantes concernant l'utilisation de contenus protégés par le droit d'auteur par les développeurs dans la formation de modèles d'apprentissage automatique, sans que les titulaires de droits y consentent ou reçoivent le crédit et la compensation qui leur sont dus.** Certains titulaires de droits ont commencé à proposer des licences pour IA, mais il reste difficile de faire valoir leurs droits et d'obtenir une rémunération pour

l'utilisation de contenus protégés par le droit d'auteur dans le cadre des activités FTD. [Nous soulignons]

- Publication du Gouvernement du Canada intitulée « Consultation sur le droit d'auteur à l'ère de l'intelligence artificielle générative », **PIÈCE P-10** (p. 8 et 9).

57. En outre, selon le United States Patent & Trademark Office, le processus implique nécessairement la reproduction d'œuvres protégées par le droit d'auteur :

The ingestion of copyrighted works for purposes of machine learning will almost by definition involve the reproduction of entire works or substantial portions thereof. [...]

[...]

When AI algorithms or processes "learn" their functions by ingesting copyrighted works, reproductions of those works are made in the process as the works are digitized and/or "read" by the AI algorithms or processes. [...]

[...]

- Rapport du United States Patent & Trademark Office intitulé « Public Views on Artificial Intelligence and Intellectual Property Policy », **PIÈCE P-11** (p. 24 et 25).

58. D'ailleurs, dans ses observations devant des instances publiques, OpenAI a défendu son usage d'œuvres protégées par le droit d'auteur dans le cadre du développement et de la formation de ses LLM :

[...] For this response, **we draw on our experience in developing cutting-edge technical AI systems**, including by the use of large, publicly available datasets **that include copyrighted works**. [...]

Modern AI systems require large amounts of data. For certain tasks, that data is derived from existing publicly accessible "corpora" (singular: "corpus") of data **that include copyrighted works**. By analyzing large corpora (**which necessarily involves first making copies of the data to be analyzed**), AI systems can learn patterns inherent in human-generated data and then use those patterns to synthesize similar data which yield increasingly compelling novel media in modalities as diverse as text, image, and audio.

[...]

[Note de bas de page 33] It is true that each work in a training corpus makes a minimal contribution to the overall model, and thus that loss of one work from the training corpus is not very detrimental to AI system quality. However, holding that training AI systems constitutes infringement would dramatically reduce the size of available corpora (by demanding removal or substantial redaction of copyrighted works in the corpora), and thus lead to significant reductions in model quality. [Notre gras]

- Observations d'OpenAI, L.P. au United States Patent and Trademark Office, **PIÈCE P-12** (p. 1, 2, 6, 7).

We believe that AI tools are at their best when they incorporate and represent the full diversity and breadth of human intelligence and experience. In order to do this, today's AI technologies require a large amount of training data and computation, as models review, analyze, and learn patterns and concepts that emerge from trillions of words and images. OpenAI's large language models, including the models that power ChatGPT, are developed using three primary sources of training data: (1) information that is publicly available on the internet, (2) information that we license from third parties, and (3) information that our users or our human trainers provide. Because copyright today covers virtually every sort of human expression—including blog posts, photographs, forum posts, scraps of software code, and government documents—it **would be impossible to train today's leading AI models without using copyrighted materials**. Limiting training data to public domain books and drawings created more than a century ago might yield an interesting experiment, but would not provide AI systems that meet the needs of today's citizens. [Nous soulignons]

- Observations d'OpenAI au *House of Lords Communications and Digital Select Committee* datées du 5 décembre 2023, **PIÈCE P-13** (p. 4).

59. La performance d'un modèle LLM dépend largement de la quantité et de la qualité des données utilisées pour son apprentissage.
60. Les œuvres écrites par des auteurs constituent une source de données particulièrement précieuse et convoitée pour ces fins. Entre autres, les œuvres littéraires et dramatiques permettent aux LLM de mimer la complexité de l'expérience humaine en reproduisant la manière dont des personnes, réelles ou fictives, raisonnent ou communiquent entre elles.
61. Il est possible d'obtenir de telles œuvres de façon légale. Les défenderesses ont plutôt choisi de former leurs LLM à partir d'ensembles de données d'apprentissage contenant des copies d'œuvres piratées prises de l'Internet, y compris des œuvres des membres du groupe.

3.2. Les contrefaçons liées à la série GPT

62. Les défenderesses ont reproduit sans permission des copies d'œuvres protégées par le droit d'auteur, y compris des œuvres des membres du groupe, dans le cadre du développement et de la formation de leurs LLM.
63. Sans limiter la généralité de ce qui précède, en juillet 2020, OpenAI a publié un document de recherche annonçant son modèle GPT-3, dans lequel elle mentionnait que l'ensemble de données d'apprentissage utilisé pour former GPT-

3 incluait deux corpus de livres pris d'Internet, nommés « Books1 » et « Books2 », tel qu'il appert du document de recherche intitulé « Language Models are Few-Shot Learners » publié le 22 juillet 2020, **PIÈCE P-14** (p. 8).

64. À l'époque, la provenance des œuvres contenues dans Books1 et Books2 était gardée secrète par OpenAI, rendant impossible toute vérification indépendante.
65. Cependant, il a longtemps été soupçonné que ces œuvres provenaient de ce qu'on appelle couramment des bibliothèques clandestines ou fantômes (en anglais, « *shadow libraries* »), c.-à-d. des plateformes en ligne qui permettent de télécharger des copies illicites d'œuvres protégées par le droit d'auteur. En effet, de telles bibliothèques clandestines sont les seules sources connues susceptibles de contenir une quantité suffisante de livres numérisés pour correspondre aux énormes volumes de données contenues dans Books1 et Books2.
66. Or, ces soupçons ont récemment été confirmés par un ancien employé d'OpenAI, qui, dans le cadre de l'affaire *Bartz et al. c. Anthropic PBC* (Case No. 3 : 24-cv-05417-WHA) devant la United States District Court, Northern District of California, a révélé qu'il avait personnellement participé au téléchargement d'œuvres à partir de « Library Genesis » (« **LibGen** ») dans le cadre du développement et de la formation de LLM pour OpenAI en 2019 :

Q. When you were at OpenAI, you personally were involving – involved in downloading a portion of Library Genesis, right?

A. Yes, I think that's fair to say.

Q. And that was around the 2019 time frame?

A. Yes.

[...]

Q. Why did you do fair use research in 2019?

A. At OpenAI, our plan was to use LibGen for training, and I wanted to understand what the legal implication of that might be.

- Extraits de l'interrogatoire hors cour de Benjamin Jacob Mann daté du 18 août 2025 dans l'affaire *Bartz et al. c. Anthropic PBC* (Case No. 3 : 24-cv-05417-WHA) devant la United States District Court, Northern District of California, **PIÈCE P-15** (p. 35 et 36).

67. Fondée en Russie, la bibliothèque clandestine LibGen contiendrait plus de 2,8 millions de livres de non-fiction, 2,5 millions de livres de fiction, 84 millions d'articles scientifiques, et 2 millions de *comics*, tel qu'il appert de la page « Library Genesis » sur le site Web fr.wikipedia.org, **PIÈCE P-16**.

68. Des recherches aléatoires du contenu de LibGen confirment que la bibliothèque contient de nombreuses œuvres d'auteurs québécois, tel qu'il appert de captures d'écran des résultats de recherche sur LibGen, **PIÈCE P-17** en liasse.
69. Il est reconnu qu'une très grande partie sinon la totalité des documents disponibles sur LibGen sont des œuvres partagées sans le consentement des titulaires de droits.
70. Il découle de ce qui précède que les livres qu'a reproduits OpenAI dans le cadre du développement et de la formation de ses LLM incluent nécessairement une très grande quantité de livres dont les droits sont détenus par des résidents québécois.
71. En outre, compte tenu de l'énorme quantité de données à télécharger de LibGen, il est vraisemblable qu'OpenAI a procédé au téléchargement du contenu par la voie de protocoles de partage de fichiers poste-à-poste (torrent).
72. En vertu de tels protocoles, l'utilisateur télécharge les données d'un même fichier depuis les postes de différents autres utilisateurs « pairs » du réseau. Dès que les parties du fichier souhaité sont téléchargées sur son poste, elles deviennent disponibles pour être redistribuées aux autres utilisateurs du réseau qui veulent se procurer le même fichier. On parlera alors de *seeding* :

In computing, and specifically peer-to-peer file sharing, seeding is the uploading of already downloaded content for others to download from. A peer, a computer that is connected to the network, becomes a seed when having acquired the entire set of data, it begins to offer its upload bandwidth to other peers attempting to download the file.

- Page « Seeding (computing) » sur le site Web en.wikipedia.org, **PIÈCE P-18**.

73. Ainsi, en téléchargeant des copies d'œuvres à l'aide de protocoles de partage de fichiers poste-à-poste, les défenderesses n'ont pas seulement reproduit des copies piratées pour leur propre usage; elles ont nécessairement distribué ces copies piratées à d'autres contrevenants.

3.3. Les contrefaçons liées à MT-NLG

74. En octobre 2021, Microsoft a publié un texte annonçant l'existence de son nouveau LLM, MT-NLG, qu'elle présentait comme le modèle le plus performant au monde, tel qu'il appert du texte intitulé « Using DeepSpeed and Megatron to Train Megatron-Turing NLG 530B, the World's Largest and Most Powerful Generative Language Model » daté du 11 octobre 2021 et publié sur le site Web microsoft.com, **PIÈCE P-19**.

75. Dans ce document, Microsoft admet que les ensembles de données d'apprentissage utilisés pour former MT-NLG incluent le corpus de livres connu sous le nom de « Books3 » pris de « The Pile ».
76. The Pile est un ensemble de données d'apprentissage créée par EleutherAI, un groupe de recherche sur l'IA.
77. Dans des publications sur l'application X (anciennement Twitter), le créateur de Books3, Shawn Presser, a confirmé que ce corpus de livres était constitué de copies de tous les livres disponibles sur « Bibliotik » (soit 196 640 livres), tel qu'il appert de publications X de Shawn Presser datées du 25 octobre 2020, **PIÈCE P-20** en liasse.
78. Les membres d'EleutherAI ont également admis que Books3 correspond au contenu de Bibliotik dans un document de recherche qu'ils ont publié pour présenter The Pile :
- Books3 is a dataset of books derived from a copy of the contents of the Bibliotik private tracker made available by Shawn Presser (Presser, 2020). Bibliotik consists of a mix of fiction and nonfiction books and is almost an order of magnitude larger than our next largest book dataset (BookCorpus2). **We included Bibliotik because books are invaluable for long-range context modeling research and coherent storytelling.** [Nous soulignons]
- Document de recherche intitulé « The Pile: An 800GB Dataset of Diverse Text for Language Modeling », **PIÈCE P-21** (p. 3 et 4).
79. Comme LibGen, Bibliotik est une bibliothèque clandestine composée de copies illicites d'ouvrages protégées par le droit d'auteur reproduites et partagées sans l'autorisation des titulaires des droits.
80. En août et septembre 2023, le magazine *The Atlantic* a obtenu le contenu de Books3 et développé un moteur de recherche qui permet de vérifier les œuvres qui s'y retrouvent, tel qu'il appert des articles intitulés « Revealed : The Authors Whose Pirated Books are Powering Generative AI » et « These 183,000 Books are Fueling the Biggest Fight in Publishing and Tech » datés du 25 septembre 2023 et publiés sur le site Web theatlantic.com, **PIÈCE P-22** en liasse.
81. Une enquête de CBC/Radio-Canada a révélé qu'au moins 2 500 livres écrits par plus de 1 200 auteurs canadiens se trouveraient dans Books3, tel qu'il appert de l'article intitulé « CBC News Analysis Finds Thousands of Canadian Authors, Books in Controversial Dataset Used to Train AI » daté du 7 décembre 2023 et publié sur le site Web cbc.ca, **PIÈCE P-23**.

82. Un autre reportage de CBC/Radio-Canada rapporte aussi que plusieurs auteurs montréalais ont appris que leurs œuvres se trouvaient sans leur permission dans Books3, tel qu'il appert de l'article intitulé « Montreal Writers Want More Protections as AI Sucks up Their Stories » daté du 25 octobre 2023 et publié sur le site Web cbc.ca, **PIÈCE P-24**.
83. De fait, des recherches aléatoires sur le moteur de recherche développé par *The Atlantic* confirment que Books3 contient de nombreuses œuvres d'auteurs québécois, tel qu'il appert de captures d'écran des résultats de recherche du contenu de Books3 dans *The Atlantic*, **PIÈCE P-25** en liasse.
84. Avant la publication de ce moteur de recherche, il n'était pas possible pour les titulaires de droits de savoir que leurs œuvres se trouvaient dans Books3 et qu'elles avaient donc été reproduites sans autorisation par Microsoft.
85. Lorsqu'elle a pris la décision de télécharger le contenu de Books3 et de s'en servir pour former MT-NLG, Microsoft savait pertinemment que celui-ci était constitué de copies d'œuvres protégées par le droit d'auteur reproduites sans autorisation, puisque ce fait avait déjà été divulgué publiquement par Shawn Presser et EleutherAI.

3.4. Les manquements des défenderesses

86. En vertu de la *Loi sur le droit d'auteur*, L.R.C. (1985), ch. C-42 (la « **LDA** »), les membres du groupe détiennent les droits exclusifs de reproduction sur les œuvres que se sont approprié les défenderesses dans le cadre du développement et de la formation de leurs LLM. Ces droits comprennent en outre le droit d'autoriser telle reproduction, tel qu'il apparaît de l'article 3(1) a) et 3(1) *in fine* LDA.
87. Les défenderesses ont accompli ou autorisé que soient accomplis à leur insu de nombreux actes réservés aux membres du groupe, portant ainsi de nombreuses atteintes aux droits d'auteur de ceux-ci en vertu de l'article 27(1) LDA.
88. Ainsi, sans restreindre la généralité de ce qui précède, en reproduisant des œuvres des membres du groupe sans leur consentement, les défenderesses ont commis des actes réservés aux membres du groupe en vertu du paragraphe 3(1) a) LDA.
89. Les défenderesses auraient pu négocier des accords de licences avec les membres du groupe.
90. Elles ont plutôt délibérément choisi d'obtenir des copies piratées depuis des bibliothèques clandestines ou autres sources illégitimes.

91. Les défenderesses ont également participé à la distribution illégale d'œuvres piratées par leur participation à des réseaux de partage de fichiers poste-à-poste. Ce faisant, elles ont commis ou autorisé que soient commis des actes réservés aux membres du groupe en vertu du paragraphe 3(1)a) et 3(1) *in fine* LDA.
92. Par ailleurs, en donnant accès à leurs LLM aux développeurs IA et aux détenteurs de licences à travers le monde, les défenderesses ont mis et continuent de mettre en circulation les œuvres des membres du groupe de façon à leur porter préjudice, commettant ainsi des violations à une étape ultérieure en vertu du paragraphe 27(2) b) LDA.
93. De même, les défenderesses ont commis et continuent de commettre des violations à une étape ultérieure en vertu du paragraphe 27(2) LDA en développant et en commercialisant des outils IA qui permettent à leurs utilisateurs de générer du contenu en violation du droit d'auteur des membres du groupe.
94. De plus, en incorporant les œuvres des membres du groupe dans leurs LLM et leurs outils IA sans aucune attribution, les défenderesses portent atteinte aux droits moraux des auteurs membres du groupe en vertu des art. 14.1 et 28.1 LDA.
95. En outre, en s'appropriant sans droit des œuvres protégées par le droit d'auteur pour développer des outils IA qui peuvent générer du contenu qui fait concurrence au travail des auteurs et met en péril leur capacité à gagner leur vie, les défenderesses se livrent à une concurrence parasitaire et déloyale qui engage leur responsabilité extracontractuelle.
96. D'ailleurs, les plateformes de vente de livres en lignes telles qu'Amazon pullulent maintenant d'œuvres générées par des outils d'IA générative comme ChatGPT et CoPilot, tel qu'il appert de l'article intitulé « Les livres écrits avec ChatGPT envahissent Amazon » daté du 10 mai 2024 et publié sur le site Web tvouvelles.ca, et de l'article intitulé « Canadian authors warn readers that AI dupes of their books are popping up on Amazon » daté du 25 janvier 2025 et publié sur le site Web cbc.ca, **PIÈCE P-26** en liasse.
97. Par ces actes, les défenderesses se sont enrichies sans justification aux dépens des membres.
98. Finalement, chacun des actes précités constitue une atteinte illicite et intentionnelle aux droits des membres à la dignité (art. 4) et à la libre jouissance et à la libre disposition de leurs biens (art. 6) en vertu de la *Charte des droits et libertés de la personne*, L.R.Q., ch. C-12 (la « **Charte** »).

3.5. Le préjudice subi par les membres du groupe et les remèdes

99. Les membres du groupe détiennent des droits exclusifs sur leurs œuvres.
100. Non seulement investissent-ils énormément de temps, d'énergie et de ressources dans leur travail; en tant qu'artistes, ils investissent tout leur être dans leur expression créative.
101. Il existe un marché permettant aux titulaires de droit d'être rémunérés équitablement pour ce travail. À titre d'exemple, en 2024, Microsoft Corporation a conclu un accord de licence avec la maison d'édition HarperCollins, lequel accord confère à Microsoft Corporation le droit d'utiliser une œuvre donnée pendant trois ans pour former ses LLM en contrepartie d'un paiement de 5 000 \$, réparti également entre l'auteur et l'éditeur, tel qu'il appert d'un article intitulé « HarperCollins, Microsoft AI deal sets first public price for training data » publié sur le site Web transparencycoalition.ai, **PIÈCE P-27**.
102. Or, dans le présent dossier, au lieu de respecter les droits exclusifs des membres du groupe, les défenderesses ont délibérément choisi de reproduire leurs œuvres sans leur consentement et sans rémunération et de s'en servir pour développer et former leurs LLM. Elles ont de plus tout mis en œuvre pour tenter de camoufler ce fait.
103. En contournant ce marché de licences légitime, les défenderesses ont non seulement porté atteinte aux droits des membres, mais elles ont aussi sapé les fondements d'un écosystème équitable et concurrentiel de rémunération des créateurs.
104. En outre, les LLM qu'elles ont formés grâce à cette violation intentionnelle des droits des membres du groupe servent à développer des produits IA qui peuvent générer des textes pouvant sembler équivalents et pouvant faire concurrence aux œuvres créées par les membres du groupe, mettant du même coup en péril leur capacité à gagner leur vie.
105. En conséquence directe de leurs gestes, les défenderesses ont profité et profitent indûment de la propriété intellectuelle des membres du groupe au mépris de leurs droits, leur causant des préjudices matériels et moraux.
106. Cette exploitation massive des œuvres des membres du groupe pour des fins commerciales est manifestement inéquitable. Ayant absolument besoin de ces œuvres pour développer des LLM plus performants que la concurrence, les défenderesses ont conclu qu'il valait mieux se tourner vers des sources pirates

plutôt que de négocier de bonne foi l'utilisation des œuvres avec les détenteurs de droits.

107. Le développement des technologies d'IA générative est au cœur des ambitions commerciales des défenderesses.
108. Or, le succès des opérations des défenderesses en matière d'IA générative repose en grande partie sur le pillage délibéré de la propriété intellectuelle des membres du groupe.
109. Sans restreindre la généralité de ce qui précède, les membres du groupe ont subi un préjudice matériel découlant des fautes des défenderesses, y compris, mais sans s'y limiter, des pertes de revenus et de marché.
110. Ils ont également subi un préjudice moral en raison de la douleur, de la tristesse et de l'anxiété de voir les fruits de leurs efforts et de leur expression créative usurpés et instrumentalisés par des entreprises multimilliardaires qui développent des produits susceptibles de les concurrencer, et ce, sans leur accorder le moindre crédit.
111. Par conséquent, les membres du groupe ont droit à des dommages-intérêts et à la restitution des profits en vertu de l'article 35 LDA ou, au choix, à des dommages-intérêts préétablis au montant 20 000 \$ par œuvre en vertu de l'article 38.1 LDA.
112. Les membres du groupe ont également droit à des dommages-intérêts punitifs en vertu de la Charte.
113. Finalement, les membres du groupe ont droit à une injonction pour faire cesser les violations de leurs droits et toute activité résultant de telles violations.

4. LES CONDITIONS REQUISES POUR L'AUTORISATION

4.1. Les faits donnant ouverture au recours individuel de la demanderesse

114. La demanderesse n'est pas parvenue à son succès facilement ou par hasard.
115. Diplômée de l'Université de Montréal en français et en traduction, elle a exercé le métier d'adjointe juridique durant 27 ans puis travaillé trois ans comme traductrice à la pige.
116. Pendant ces 30 années, elle a inlassablement soumis ses romans à diverses maisons d'édition qui, bien qu'appréciant leur qualité, ne croyaient pas suffisamment à l'existence d'un marché pour la littérature de genre fantastique au Québec.

117. Son destin a changé en 2002 avec la parution de son roman *Le feu dans le ciel*, premier tome de la série *Les Chevaliers d'Émeraude*.
118. L'univers des *Chevaliers d'Émeraude* a rapidement rejoint son lectorat, jeunes et jeunes de cœur, et est devenu un véritable phénomène littéraire.
119. Après des années d'acharnement, la demanderesse pouvait finalement vivre de sa plume.
120. Elle a depuis écrit plus de 100 livres, chacun étant le fruit de son imagination fertile et du travail acharné qu'elle consacre sans relâche à son art.
121. LibGen contient plus d'une centaine de copies piratées de livres de la demanderesse, tel qu'il appert de captures d'écran des résultats de recherche pour « Anne Robillard » dans le contenu de LibGen, **PIÈCE P-28** en liasse.
122. La demanderesse, qui détient les droits d'auteur et moraux sur l'ensemble de son œuvre, n'a jamais consenti à ce que des copies de ses livres soient mis à disposition pour téléchargement sur LibGen, ni sur aucune autre bibliothèque clandestine.
123. Elle n'a jamais été approchée par les défenderesses et n'a jamais consenti à ce que celles-ci reproduisent ses œuvres pour quelque fin que ce soit, et encore moins, pour développer des LLM et outils IA.
124. La demanderesse a récemment appris que les défenderesses avaient téléchargé le contenu de bibliothèques clandestines, y compris celles où se retrouvent illégalement ses œuvres, dans le cadre du développement et de la formation de leurs LLM.
125. Elle est outragée que des entreprises multimilliardaires comme les défenderesses aient fait le choix d'usurper ainsi sa propriété intellectuelle et celle de centaines de milliers d'auteurs et de créateurs dans une course effrénée à la domination du marché de l'IA.

4.2. Les questions de fait ou de droit identiques, similaires ou connexes

126. Les demandes des membres soulèvent les questions de droit ou de fait identiques, similaires ou connexes suivantes :
 - (a) Les défenderesses ont-elles reproduit sans autorisation des œuvres protégées par le droit d'auteur dont les membres du groupe sont les titulaires?

- (b) Les défenderesses ont-elles formé leurs LLM avec des ensembles de données d'apprentissage reproduisant des œuvres protégées par le droit d'auteur dont les membres du groupe sont les titulaires?
- (c) Les défenderesses ont-elles violé de quelque façon le droit d'auteur des membres du groupe?
- (d) Les défenderesses ont-elles violé de quelque façon les droits moraux des membres du groupe sur leurs œuvres?
- (e) Les défenderesses ont-elles commis une faute envers les membres du groupe en se livrant à une concurrence parasitaire et déloyale?
- (f) Les défenderesses se sont-elles enrichies sans justification aux dépens des membres?
- (g) Les défenderesses ont-elles violé les droits des membres en vertu de la *Charte des droits et libertés de la personne*?
- (h) Quels sont les remèdes et dommages auxquels ont droit les membres du groupe?

4.3. La composition du groupe

- 127. La composition du groupe rend difficile ou peu pratique, voire impossible, l'application des règles sur le mandat d'ester en justice pour le compte d'autrui ou sur la jonction d'instance.
- 128. Le secret entretenu par les défenderesses sur les œuvres qui ont servi et servent dans le cadre du développement ou de la formation de leurs LLM fait en sorte qu'il est impossible d'identifier et de retracer tous les membres du groupe.
- 129. Toutefois, on peut raisonnablement estimer que le groupe comprend des milliers de personnes, incluant des auteurs, éditeurs, cessionnaires et ayant droits.
- 130. Ainsi, selon Statistique Canada entre 6 800 et 8 485 nouveaux ouvrages d'auteurs canadiens sont publiés annuellement au Canada, tel qu'il appert du document de Statistique Canada intitulé « Tableau 21-10-0207-01 Éditeurs de livres, nouveaux ouvrages édités selon la catégorie commerciale et l'origine des auteurs », **PIÈCE P-29**.
- 131. Dans ces circonstances, l'action collective est une voie appropriée pour que les membres du groupe puissent faire valoir leurs droits et aient accès à la justice.

4.4. La représentation adéquate

132. La demanderesse est en mesure d'assurer une représentation adéquate des membres du groupe.
133. Elle est prête à consacrer le temps nécessaire au dossier.
134. Elle a donné mandat à ses avocats d'obtenir tous les renseignements pertinents au dossier et se tiendra informée des développements.
135. Elle est de bonne foi et entreprend ces procédures dans l'unique but de faire en sorte que ses droits et ceux des membres du groupe soient reconnus et respectés et qu'il soit remédié aux préjudices que chacun a subis.

5. LA COMPÉTENCE DU TRIBUNAL

136. Les autorités québécoises sont compétentes pour entendre l'action collective du fait que des fautes ont été commises en tout ou en partie au Québec et que les membres du groupe ont subi et subissent un préjudice au Québec.
137. Les fautes des défenderesses, notamment la collecte, la reproduction et la distribution de textes et de données provenant de l'Internet, dont des copies d'œuvres québécoises, ont nécessairement eu lieu au moins en partie au Québec.
138. Les membres du groupe ont tous subi ou subissent au Québec un préjudice résultant de la violation intentionnelle de leurs droits par les défenderesses.
139. Sans limiter la généralité de ce qui précède, en s'accaparant et en reproduisant sans autorisation les œuvres dont les droits d'auteurs sont détenus par les membres, les défenderesses ont privé ces derniers de leur droit d'obtenir, au Québec, une juste compensation pour leur travail.
140. De plus, les défenderesses Microsoft ont des établissements au Québec et la contestation est relative à leurs activités au Québec.
141. Finalement, les défenderesses savaient que leurs ensembles de données d'apprentissage contenaient des œuvres piratées du monde entier, y compris d'auteurs québécois. Elles pouvaient donc prévoir que des préjudices découlant de leur décision d'exploiter ces œuvres sans autorisation pour ses fins commerciales se manifesteraient au Québec.
142. La demanderesse propose que la présente action collective soit exercée dans le district judiciaire de Montréal puisqu'un nombre important des membres du groupe et les avocats soussignés y sont domiciliés, et que les défenderesses Microsoft y ont des établissements.

POUR CES MOTIFS, PLAISE À LA COUR

- A. **ACCUEILLIR** la présente demande;
- B. **AUTORISER** l'exercice de l'action collective, dont la nature est celle d'une action en dommages-intérêts, restitution de profits et injonction;
- C. **DÉSIGNER** Anne Robillard à titre de représentante du groupe des personnes ci-après décrit :

Toute personne au Québec titulaire du droit d'auteur sur une œuvre littéraire ou dramatique que les défenderesses ont reproduite sans autorisation dans le cadre du développement ou de la formation de grands modèles de langage (en anglais, « *large language models* »).

- D. **IDENTIFIER** comme suit les principales questions de droit et de fait qui seront traitées collectivement :
 - (a) Les défenderesses ont-elles reproduit sans autorisation des œuvres protégées par le droit d'auteur dont les membres du groupe sont les titulaires?
 - (b) Les défenderesses ont-elles formé leurs LLM avec des ensembles de données d'apprentissage reproduisant des œuvres protégées par le droit d'auteur dont les membres du groupe sont les titulaires?
 - (c) Les défenderesses ont-elles violé de quelque façon le droit d'auteur des membres du groupe?
 - (d) Les défenderesses ont-elles violé de quelque façon les droits moraux des membres du groupe sur leurs œuvres?
 - (e) Les défenderesses ont-elles commis une faute envers les membres du groupe en se livrant à une concurrence parasitaire et déloyale?
 - (f) Les défenderesses se sont-elles enrichies sans justification aux dépens des membres?
 - (g) Les défenderesses ont-elles violé les droits des membres en vertu de la *Charte des droits et libertés de la personne*?
 - (h) Quels sont les remèdes et dommages auxquels ont droit les membres du groupe?

- E. **IDENTIFIER** comme suit les conclusions recherchées qui s'y rattachent :

- A. ACCUEILLIR la présente demande;
- B. **CONDAMNER** les défenderesses à payer une somme à être déterminée à titre de dommages-intérêts punitifs, avec intérêt au taux légal et l'indemnité additionnelle à compter du jugement;

- C. CONDAMNER solidairement les défenderesses OpenAI et les défenderesses Microsoft à payer à chaque membre du groupe un montant de VINGT MILLE DOLLARS (20 000 \$), sauf à parfaire, par œuvre dont elles se sont servies sans autorisation dans le cadre du développement ou de la formation de la série GPT (ou de tout autre LLM développé conjointement par elles), à titre de dommages-intérêts compensatoires, avec intérêt au taux légal et l'indemnité additionnelle à compter de l'assignation; et

CONDAMNER solidairement les défenderesses Microsoft à payer à chaque membre du groupe un montant de VINGT MILLE DOLLARS (20 000 \$), sauf à parfaire, par œuvre dont elles se sont servies sans autorisation dans le cadre du développement ou de la formation de MT-NLG (ou de tout autre LLM développé par elles indépendamment des défenderesses OpenAI), à titre de dommages-intérêts compensatoires, avec intérêt au taux légal et l'indemnité additionnelle à compter de l'assignation;

- D. CONDAMNER les défenderesses à payer, en sus des dommages-intérêts, une proportion équitable des profits réalisés en commettant les violations du droit d'auteur des membres du groupe, avec intérêt au taux légal et l'indemnité additionnelle à compter de l'assignation;

- E. AU CHOIX, EN REMPLACEMENT DES CONCLUSIONS C ET D : CONDAMNER solidairement les défenderesses OpenAI et les défenderesses Microsoft à payer à chaque membre du groupe un montant de VINGT MILLE DOLLARS (20 000 \$) par œuvre dont elles se sont servies sans autorisation dans le cadre du développement ou de la formation dans le cadre du développement ou de la formation de la série GPT (ou de tout autre LLM développé conjointement par elles), à titre de dommages-intérêts préétablis, avec intérêt au taux légal et l'indemnité additionnelle à compter de l'assignation; et

CONDAMNER solidairement les défenderesses Microsoft à payer à chaque membre du groupe un montant de VINGT MILLE DOLLARS (20 000 \$) par œuvre dont elles se sont servies sans autorisation dans le cadre du développement ou de la formation de MT-NLG (ou de tout autre LLM développé par elles indépendamment des défenderesses OpenAI), à titre de dommages-intérêts préétablis, avec intérêt au taux légal et l'indemnité additionnelle à compter de l'assignation; et

- F. ENJOINDRE aux défenderesses de cesser toute commercialisation, distribution, ou publication de LLM ou de tout autre outil d'IA formé sur des œuvres des membres du groupe sans leur consentement et de détruire toute copie de ces œuvres ou données qu'elles ont en leur possession;

- G. ORDONNER le recouvrement collectif des réclamations;
 - H. ORDONNER aux défenderesses de déposer au greffe de cette cour la totalité des sommes susdites, ainsi que les intérêts et l'indemnité additionnelle;
 - I. LE TOUT avec frais de justice incluant les frais d'experts, de publication d'avis et d'administration des réclamations;
- F. **DÉCLARER** qu'à moins d'exclusion, les membres du groupe seront liés par tout jugement à intervenir sur l'action collective de la manière prévue par la loi;
 - G. **FIXER** le délai d'exclusion à trente (30) jours après la date de l'avis aux membres, délai à l'expiration duquel les membres du groupe qui ne se seront pas exclus seront liés par tout jugement à intervenir;
 - H. **ORDONNER** la publication d'un avis aux membres dans les termes et par les moyens à déterminer;
 - I. **RÉFÉRER** le dossier au juge en chef pour détermination du district dans lequel l'action collective devra être exercée et désignation du juge pour l'entendre;
 - J. **ORDONNER** au greffier de cette cour, pour le cas où l'action doit être exercée dans un autre district, de transmettre le dossier, dès décision du juge en chef, au greffier de cet autre district;
 - K. **LE TOUT** avec frais de justice incluant les frais de publication de l'avis.

MONTREAL, le 4 septembre 2025



Audren Rolland s.e.n.c.r.l.
Avocats de la demanderesse Anne Robillard

AVIS D'ASSIGNATION
(Art. 145 et suivants CPC)

Dépôt d'une demande en justice

Prenez avis que la demanderesse a déposé au greffe de la Cour supérieure du Québec, district judiciaire de Montréal, la présente Demande pour autorisation d'exercer une action collective et pour être représentante.

Réponse à cette demande

Vous devez répondre à cette demande par écrit, personnellement ou par avocat, au palais de justice de Montréal situé au 1, rue Notre-Dame Est, Montréal (Québec) H2Y 1B6, district de Montréal dans les 15 jours de la signification de la présente demande ou, si vous n'avez ni domicile, ni résidence, ni établissement au Québec, dans les 30 jours de celle-ci. Cette réponse doit être notifiée à l'avocat de la demanderesse.

Défaut de répondre

Si vous ne répondez pas dans le délai prévu, de 15 ou de 30 jours, selon le cas, un jugement par défaut pourra être rendu contre vous sans autre avis dès l'expiration de ce délai et vous pourriez, selon les circonstances, être tenu au paiement des frais de justice.

Contenu de la réponse

Dans votre réponse, vous devez indiquer votre intention, soit :

- de convenir du règlement de l'affaire;
- de proposer une médiation pour résoudre le différend;
- de contester cette demande et, dans les cas requis par le Code de procédure civile, d'établir à cette fin, en coopération avec la demanderesse, le protocole qui régira le déroulement de l'instance. Ce protocole devra être déposé au greffe de la Cour du district mentionné plus haut dans les 45 jours de la signification du présent avis. Toutefois, ce délai est de 3 mois en matière familiale ou si vous n'avez ni domicile, ni résidence, ni établissement au Québec;
- de proposer la tenue d'une conférence de règlement à l'amiable.

Cette réponse doit mentionner vos coordonnées et, si vous êtes représenté par un avocat, le nom de celui-ci et ses coordonnées.

Lieu et dépôt de la demande en justice

- P-20** Publications X de Shawn Presser datées du 25 octobre 2020, en liasse
- P-21** Document de recherche intitulé « The Pile: An 800GB Dataset of Diverse Text for Language Modeling »
- P-22** Articles intitulés « Revealed : The Authors Whose Pirated Books are Powering Generative AI » et « These 183,000 Books are Fueling the Biggest Fight in Publishing and Tech » datés du 25 septembre 2023 et publiés sur le site Web theatlantic.com, en liasse
- P-23** Article intitulé « CBC News Analysis Finds Thousands of Canadian Authors, Books in Controversial Dataset Used to Train AI » daté du 7 décembre 2023 et publié sur le site Web cbc.ca
- P-24** Article intitulé « Montreal Writers Want More Protections as AI Sucks up Their Stories » daté du 25 octobre 2023 et publié sur le site Web cbc.ca
- P-25** Captures d'écran des résultats de recherche du contenu de Books3 dans *The Atlantic*, en liasse
- P-26** Article intitulé « Les livres écrits avec ChatGPT envahissent Amazon » daté du 10 mai 2024 et publié sur le site Web tvanouvelles.ca, et article intitulé « Canadian authors warn readers that AI dupes of their books are popping up on Amazon » daté du 25 janvier 2025 et publié sur le site Web cbc.ca, en liasse
- P-27** Article intitulé « HarperCollins, Microsoft AI deal sets first public price for training data » publié sur le site Web transparencycoalition.ai
- P-28** Captures d'écran des résultats de recherche pour « Anne Robillard » dans le contenu de LibGen, en liasse
- P-29** Document de Statistique Canada intitulé « Tableau 21-10-0207-01 Éditeurs de livres, nouveaux ouvrages édités selon la catégorie commerciale et l'origine des auteurs »

La copie des pièces vous sera fournie sur demande, à la suite de la notification de votre réponse.

MONTREAL, le 4 septembre 2025



Audren Rolland s.e.n.c.r.l.

Avocats de la demanderesse Anne Robillard

**ATTESTATION D'INSCRIPTION AU
RÉPERTOIRE NATIONAL DES ACTIONS COLLECTIVES**
(Art. 55 du *Règlement de la Cour supérieure du Québec en matière civile*)

Les avocats soussignés attestent que la Demande pour autorisation d'exercer une action collective et pour être représentante sera inscrite au Répertoire national des actions collectives.

MONTRÉAL, le 4 septembre 2025



Audren Rolland s.e.n.c.r.l.
Avocats de la demanderesse Anne Robillard

AUDREN ROLLAND

(Chambre des actions collectives)
COUR SUPÉRIEURE
DISTRICT DE MONTRÉAL
N° 500-06-

ANNE ROBILLARD

Demanderesse

c.

OPENAI, INC

et

OPENAI GP, L.L.C.

et

OPENAI GLOBAL, LLC

et

OPENAI OPCO, LLC

et

OPENAI HOLDINGS, LLC

et

OPENAI, L.L.C.

et

OAI CORPORATION

et

OPENAI STARTUP FUND I, L.P.

et

OPENAI STARTUP FUND GP I, L.L.C.

et

OPENAI STARTUP FUND MANAGEMENT, LLC

et

MICROSOFT CORPORATION

et

MICROSOFT CANADA INC.

Défenderesses

**DEMANDE POUR AUTORISATION D'EXERCER
UNE ACTION COLLECTIVE ET POUR ÊTRE
REPRÉSENTANTE**

(Art. 574 et suivants CPC)

ORIGINAL

AUDREN | ROLLAND

Audren Rolland s.e.n.c.r.l.

393, rue St-Jacques, bureau 248

Montréal (Québec), H2Y 1N9

Tél. 514.974.3145; Téléc. 514.284.7771

Me Marie Audren, Ad. E.

Me Marc-André Grou

Me Emmanuelle Rolland

maudren@audrenrolland.com

mgrou@audrenrolland.com

erolland@audrenrolland.com

BA1391