

La diligence des plateformes et l'expression « légale, mais préjudiciable »

Emily Laidlaw

Le gouvernement du Canada devrait présenter un projet de loi relatif à la sécurité en ligne au cours des prochains mois. Quelle que soit la proposition, il ne s'agit que du début. La prochaine décennie sera une période de frictions juridiques considérables, à mesure que les lois en matière de sécurité en ligne seront affinées par les tribunaux, les organismes de réglementation et les législateur·trice·s. Même si aucune loi n'est adoptée, nous serons toujours confronté·e·s à la même question : qu'est-ce qu'une plateforme responsable au sens juridique? Plusieurs casse-têtes juridiques me tiennent éveillée la nuit, mais il y en a un qui me tracasse tout particulièrement :

un modèle de gestion des risques peut-il et doit-il cibler l'expression « légale, mais préjudiciable »?

Tout indique du côté de [Patrimoine canadien](#) que la législation s'inspirera de ce que l'on peut qualifier de modèles de gestion des risques, d'obligation de diligence ou de diligence raisonnable. Dans le cadre de ces cadres juridiques, l'accent est mis sur les risques systémiques de préjudice des plateformes, comme les systèmes de recommandation ou d'autres types d'algorithmes, la modération du contenu, la publicité et les pratiques en matière de données. On traite les services de ces plateformes comme s'il s'agissait d'une question de sécurité des produits. L'obligation de l'entreprise est donc d'évaluer les risques de préjudice que présentent ses services du côté des applications frontales, d'effectuer une surveillance continue des risques et d'agir en conséquence. Le mécanisme de responsabilisation des plateformes allie des obligations de production de rapports relatifs à la transparence et une surveillance réglementaire. On peut observer cette approche dans le [Online Safety Bill](#) du Royaume-Uni et le [Règlement sur les services numériques](#) de l'Union européenne. Le Canada explore également cette approche dans le cadre du [devoir d'agir de manière responsable](#). Tous ces modèles, en substance, exigent que les plateformes mettent en œuvre des systèmes de prise de décision raisonnable.

Bien que j'aie préconisé ce modèle, l'analogie de la sécurité des produits a ses limites. Il y a une différence entre le fait de construire des voitures plus sécuritaires et la création d'espaces plus sûrs pour le discours, en raison de l'intersection avec des droits fondamentaux. La friction est d'autant plus évidente lorsqu'on examine les systèmes et le contenu aux confins de la sphère juridique.

Lorsque je discute de législation en matière de sécurité en ligne, la plupart des gens mentionnent la mésinformation, la désinformation, l'intimidation, les attaques de foules, le contenu haineux et d'autres types de préjudices. Le contenu de ce type est généralement légal, mais préjudiciable. La législation en matière de sécurité en ligne peut-elle et doit-elle traiter ce type de préjudice? Cette question a soulevé une énorme controverse au Royaume-Uni dans le contexte du *Online Safety Bill*. La plupart des dispositions allant dans ce sens [ont été retirées](#) du

projet de loi. Il y a pourtant une discussion importante à tenir au sujet des responsabilités et des obligations des plateformes dans un cadre constitutionnel.

Par exemple, en 2022, au Royaume-Uni, [l'enquête du coroner](#) sur le suicide d'une jeune fille, Molly Russell, a révélé que les systèmes de recommandation de Pinterest et d'Instagram mettaient en avant du contenu sur l'automutilation, le suicide et la dépression. Un modèle de gestion des risques qui se concentre uniquement sur l'expression illégale n'exigerait pas de diligence raisonnable de la part des plateformes dans ce domaine. Toutefois, le devoir de protéger les intérêts spéciaux des enfants et d'évaluer les risques associés au système qui leur recommande du contenu permettrait de cerner ce qui a été préjudiciable à Molly Russell. De même, la désinformation et la mésinformation sont généralement [légales](#), même si elles pourraient inciter certaines personnes à prendre des décisions liées à l'illégalité. Le *Règlement sur les services numériques* de l'Union européenne contourne de manière créative la question de l'expression légale, mais préjudiciable et de la désinformation en imposant des obligations de diligence raisonnable aux très grandes plateformes en ligne dans le but de gérer les risques systémiques qu'elles posent sans entrer dans le sujet épineux du contenu pris isolément ou de sa légalité.

Du point de vue juridique, il y a deux défis de taille. Tout d'abord, le point de départ pour cibler ce type de contenu, même à un niveau systémique, est qu'un système juridique fondé sur la liberté d'expression protège [l'expression](#) impopulaire, déplaisante et dérangeante, et cela devrait être le cas, car les fondements de la démocratie, de la découverte et de la vérité inhérents à l'expression occupent une place centrale dans notre société. Toutefois, un tel droit doit être renforcé et équilibré par d'autres droits constitutionnels, comme le droit à la vie privée et à l'égalité, les obligations fiduciaires de la Couronne envers les peuples autochtones et l'infrastructure unique d'Internet qui modifie les conditions sociales de la parole.

Par exemple, il existe un risque élevé de conséquences imprévues. Les systèmes de modération de contenu, malgré les bonnes intentions des personnes qui les ont mis en place, ont été utilisés abusivement pour cibler des groupes de personnes racisées et d'autres groupes de personnes marginalisées. Les systèmes algorithmiques basés sur des intrants indésirables produisent des extrants biaisés, et ainsi de suite. La réponse ne consiste pas à étendre la liberté d'expression ou la réglementation. Une approche beaucoup plus nuancée est nécessaire. Par exemple, il ne devrait pas y avoir d'obligation de supprimer le contenu s'apparentant à une expression légale, mais préjudiciable, car cela constituerait une interférence extraordinaire avec le droit à la liberté d'expression. Toutefois, certaines caractéristiques de conception pourraient être justifiables, notamment les fonctions redonnant un certain pouvoir aux utilisateur·trice·s, comme les outils de mise en sourdine, de filtrage du contenu ou de présentation d'autres sources d'actualité aux côtés du contenu signalé comme désinformation. Il pourrait être utile de tirer des leçons des jugements portant sur les restrictions relatives [aux heures, au lieu ou au mode d'expression](#) et sur [la publicité](#). Les obligations en matière de transparence pourraient comprendre la présentation de rapports sur les mesures de gestion des risques adoptées en vue de protéger les enfants et les droits fondamentaux.

Deuxièmement, il y a une question distincte qui se pose : ces types de modèles fondés sur la diligence raisonnable soulèvent-ils même des questions relatives aux droits de la personne? Dans le cadre d'un modèle fondé sur la gestion des risques, on ne dit pas aux entreprises ce qu'elles doivent faire, simplement que *quelque chose* doit être fait et qu'elles doivent être responsables de ce processus. Par conséquent, il est peu probable qu'un tribunal considère qu'une entreprise de médias sociaux entreprend une action gouvernementale et qu'elle est directement liée par la [Charte canadienne des droits et libertés](#) pour ses fonctions publiques.

Toutefois, les décisions qui découlent de cette évaluation des risques peuvent impliquer des droits. Plus la législation est précise, plus le risque est grand. Par exemple, les risques en matière de droits sont différents selon qu'une éventuelle législation exige qu'on réduise ou qu'on augmente l'importance d'un contenu dans le système de recommandation. On s'entend généralement sur le fait qu'il n'existe pas de droit à un public, et par conséquent, la réduction en importance d'un contenu pourrait être moins problématique du point de vue juridique, bien qu'il s'agisse d'une [zone grise](#). Cependant, si on exige qu'un contenu soit amplifié, cela pourrait constituer une violation des droits en matière de liberté d'expression de la plateforme, puisqu'il s'agirait d'un discours forcé.

Les exigences en matière d'avertissements et d'indicateurs posent des problèmes similaires. La liberté d'expression comprend le droit de [ne rien dire](#) ou de ne pas être obligé-e de dire certaines choses. Le risque en matière de droits est plus grand aux États-Unis, où la doctrine sur le discours forcé a fait l'objet de nombreux [recours judiciaires](#). Au Canada, l'examen de la proportionnalité de l'article 1 de la *Charte canadienne des droits et libertés* pèse en faveur de la constitutionnalité des messages d'avertissement, mais cela dépend toujours du contexte et nous n'avons que peu d'exemples à notre disposition ([ici](#) et [ici](#)).

Il y a de bonnes raisons pour que le gouvernement commence par une liste restreinte relative à l'expression illégale. S'il suit les traces de l'Union européenne, du Royaume-Uni et de [l'Australie](#), un nouveau système réglementaire verra le jour, accompagné des difficultés croissantes qui en découlent. Je veux toutefois être claire : une vaste série de préjudices resteront non réglés, ce qui signifie que nous continuerons de dépendre de l'autogouvernance des entreprises et des pratiques sélectives, voire absentes, en matière de transparence. Pour de nombreux types de préjudices, il s'agit précisément de l'approche qu'on devrait employer en soutien à notre engagement envers la liberté d'expression. Dans d'autres cas, les méfaits sont trop importants pour être ignorés. Les 10 prochaines années seront remplies de débats sur le domaine des droits garantis en matière d'expression, de préjudices et d'algorithmes.