Direction scientifique de la collaboration
avec les acteurs publics, Obvia

Dans l'œil de l'Obvia

JANVIER 2026

L'IA face à la diversité linguistique et culturelle : un équilibre à trouver

Mise en contexte

Au Québec, le numérique et l'Internet génèrent un débat face à la préservation de la langue française depuis de nombreuses années. Avec l'adoption massive de systèmes d'IA dans nos activités quotidiennes, entraînés sur des données majoritairement anglophones et produits par des entreprises étrangères, les craintes de voir la langue française au Québec encore plus menacée se sont multipliées.

Plus précisément, les difficultés du **français québécois** à être compris et traité par les algorithmes conçus ailleurs dans le monde sont notables. Pensons à la correction automatique d'un texto qui remplace « je suis tanné » par « talentueux » ou encore à un agent conversationnel qui comprend mal une requête formulée en français québécois. Les autres langues locales, c'est-à-dire les **langues autochtones**, sont d'autant plus marginalisées et mal desservies par les systèmes d'IA.

Dans le cas d'outils d'assistance simples comme la correction ou la transcription automatique, la performance médiocre d'un algorithme peut représenter, tout au plus, un désagrément. Or, pour d'autres sphères comme la musique ou l'audiovisuel, les algorithmes non adaptés au contexte local posent une réelle menace à la survie de la spécificité culturelle québécoise.

Puisque l'isolement face aux nouvelles technologies n'est pas une option viable, comment faire en sorte que ces outils deviennent des médiateurs qui soutiennent la diversité linguistique plutôt que de la menacer ? Il s'agit d'une question centrale abordée lors de la dernière édition des Dialogues interdisciplinaires, ayant eu lieu le 10 décembre 2025 à Québec. Cette note de breffage rapporte les points saillants discutés à l'occasion de cette rencontre et fait état des principaux enjeux qui touchent la diversité linguistique et la découvrabilité des contenus locaux à l'ère de l'IA.

Statistique

En 2024, au Québec, seulement 4,6 % des pistes les plus écoutées sur des plateformes de diffusion en continu provenaient d'artistes francophones québécois, alors que la proportion des ventes sur supports physiques était 4 fois plus élevée[1].

Citation de l'UNESCO

« Alors que les technologies numériques transforment la manière dont les sociétés communiquent et produisent des savoirs, veiller à ce que toutes les langues puissent s'épanouir — dont les langues autochtones et les langues à faibles ressources — est devenu une priorité mondiale urgente. »[2]

1 Institut de la statistique du Québec (2025). *Faits saillants sur la consommation d'enregistrements musicaux – Annuels.* https://statistique.quebec.ca/en/produit/publication/faits-saillants-consommation-enregistrements-musicaux

2 UNESCO (2025). *UNESCO and G20 promote Linguistic Diversity in Digital Age.* https://www.unesco.org/en/articles/unesco-and-g20-promote-linguistic-diversity-digital-age. Traduction libre

Résumé des connaissances

La complexité d'assurer la diversité linguistique au sein des systèmes d'IA est entre autres due à la problématique des langues dites à **faibles ressources** (*low-resource languages*), c'est-à-dire qui sont sous-représentées dans les contenus disponibles en ligne. Sans une masse critique de données dans la langue ciblée, les grands modèles de langage (GML) ont du mal à saisir les nuances, les expressions locales, le contexte ou encore l'humour et le sarcasme. Les langues étant des vecteurs de culture, d'histoire et de valeurs partagées, le risque que les systèmes d'IA omniprésents marginalisent encore plus les petites communautés linguistiques est bien réel. Lorsqu'un modèle est entraîné sur tout Internet, il faut comprendre que cet Internet est largement anglophone et alimenté de contenus américains, véhiculant ainsi des références plutôt homogènes.

« Au-delà des erreurs factuelles, l'IA a tendance à aplanir la richesse linguistique. Les agents conversationnels peinent à saisir les accents régionaux et les variations linguistiques, comme le français québécois ou les langues créoles parlées en Haïti et dans les Caraïbes francophones. »[3]

Même avec des systèmes conçus en français (parisien), le français québécois est souvent mal compris. D'un point de vue technique, cela est dû aux **lexèmes** (les séquences de caractères formant des unités porteuses de sens en français québécois) qui ne sont pas regroupés pour former les **jetons** nécessaires au traitement lexical par l'algorithme. C'est pourquoi certains experts et expertes vont recommander de miser sur des modèles de langage plus petits, mais adaptés au contexte québécois, pour obtenir de meilleurs résultats.

Peut-on faire confiance à la traduction par l'IA ?

Les limites linguistiques des IA n'ont pourtant pas empêché leur utilisation de plus en plus généralisée à des fins de traduction. S'il est vrai que les logiciels de traduction peuvent engendrer des économies de temps et d'argent importantes, une œuvre en français québécois qui subit une traduction automatique risque de voir ses subtilités et son sens culturel dilués. L'exercice devient pratiquement impossible dans le cas des langues autochtones, qui sont porteuses de savoirs très complexes et propres à leurs contextes : la traduction vers les langues indo-européennes induit une perte nette de sens. L'appauvrissement des nuances culturelles et l'effacement de certaines spécificités linguistiques par la traduction automatique s'insèrent dans le phénomène de « **technocolonialité** » observé par plusieurs spécialistes[4].

De plus, la traduction automatisée d'une langue neutre comme l'anglais vers des langues plus genrées comme le français fait apparaître des **biais algorithmiques** et reproduit des **stéréotypes**, par exemple dans le domaine de l'emploi. Un simple essai de ChatGPT (fait en janvier 2026) permet d'illustrer la problématique :

Requête

Traduire la phrase suivante :
« the doctor and the nurse treat the patient »

Réponse

« Le médecin et l'infirmière traitent le patient »

Il faut toutefois reconnaître que les plateformes d'IA générative peuvent faire mieux, puisqu'elles proposent parfois des alternatives non genrées, comme pour « *housekeeper* », où ChatGPT suggère « femme de ménage », mais aussi « l'employé(e) de maison » et « la personne chargée de l'entretien ménager ».

Ces imperfections, qui peuvent avoir de graves répercussions, justifient la présence d'un **humain dans la boucle** pour tout exercice de traduction. C'est le constat unanime des experts et expertes présentes aux *Dialogues interdisciplinaires : L'IA et la diversité linguistique* – les traducteurs et traductrices demeurent essentiels pour saisir les nuances de ton, les références culturelles d'une œuvre et la neutralité d'un texte, mais le métier est voué à en devenir un d'édition et de garant de la qualité.

Des secteurs et publics directement affectés

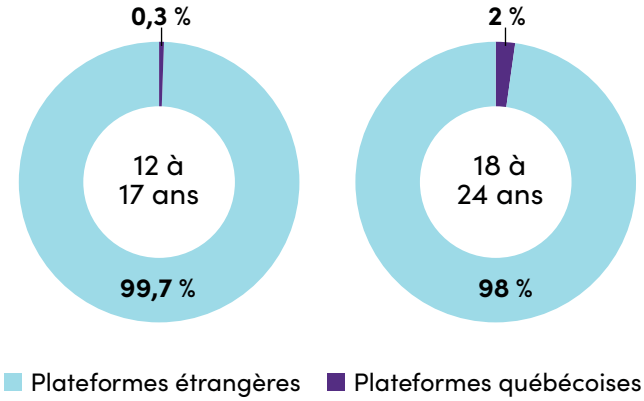
Question d'actualité depuis plusieurs années, le milieu de la culture est fortement préoccupé par l'effet des algorithmes de recommandation et la force de frappe des contenus musicaux et audiovisuels américains. Les organisations de la société civile comme la Coalition canadienne pour la diversité des expressions culturelles et la Fédération internationale des Coalitions pour la diversité culturelle (FICDC) sont d'ailleurs très mobilisées tant aux niveaux local, fédéral et international pour rappeler l'importance de favoriser la découvrabilité et la valorisation des contenus culturels francophones. En 2024 et 2025, leurs travaux ont contribué à la recommandation du Groupe de réflexion sur la diversité des expressions culturelles dans l'environnement numérique de l'UNESCO[5] d'ajouter un **Protocole spécifique à l'environnement numérique** dans la Convention de 2005 sur la protection et la promotion de la diversité des expressions culturelles.

3 Robinet, F. (2025, 23 mars). Mind your language: The battle for linguistic diversity in AI. *UN News*. https://news.un.org/en/story/2025/03/1161406. Traduction libre
4 Concept évoqué par Julie Gramaccia lors d'un panel de Dialogues interdisciplinaires: L'IA et la diversité linguistique. Pour aller plus loin, voir : Mejias, U. A., & Couldry, N. (2024). *Data grab: The new colonialism of big tech and how to fight back*. University of Chicago Press.
5 Comité intergouvernemental pour la protection et la promotion de la diversité des expressions culturelles (2024). *Recommandations du Groupe de réflexion sur la diversité des expressions culturelles dans l'environnement numérique*. https://unesdoc.unesco.org/ark:/48223/pf0000392215

Les unions d'artistes francophones au Canada doivent déjà dédier du temps et des ressources importantes à la promotion de leurs œuvres sur le terrain (ex. : dans les salles de spectacle et les festivals). Désormais, la nécessité de répliquer ces efforts pour l'environnement numérique représente une tâche colossale. Ces organisations reconnaissent l'importance de **nourrir les plateformes** (éthiques) de diffusion en continu de leurs contenus et d'assurer leur **représentation dans l'espace infonuagique**, mais elles déplorent de devoir porter ce fardeau seules. Notamment, les fonds d'accélération comme Scale AI sont disponibles uniquement pour les entreprises privées, sans soutien pour le secteur culturel[6].

À long terme, le manque de représentativité des contenus locaux dans les algorithmes de recommandation pourrait avoir de lourdes conséquences. Des études récentes ont révélé que la place des plateformes locales dans les habitudes de consommation des jeunes et des enfants était marginale :

Volume d'écoute des services de diffusion continue (% de la durée totale) au Québec selon la propriété de la plateforme[7]



Plusieurs experts et expertes craignent que des générations entières ne soient plus du tout exposées aux œuvres locales, menant à un déclin des expressions culturelles québécoises. Les diffuseurs publics peinent à concurrencer avec des plateformes dont les contenus algorithmiques sont conçus pour accrocher les plus jeunes et créer une dépendance[8].

En 2024, le Conseil de la radiodiffusion et des télécommunications canadiennes (CRTC) a ordonné aux plateformes de diffusion étrangères de verser 5 % de leurs revenus canadiens annuels afin de soutenir le système de radiodiffusion canadien. **Les paiements ont depuis été suspendus, en attendant que la Cour d'appel fédérale se prononce sur l'affaire[9].**

Le projet de loi 109 adopté : une nouvelle ère?

L'Assemblée nationale a adopté, en décembre 2025, le projet de loi 109, *Loi affirmant la souveraineté culturelle du Québec* et édictant la *Loi sur la découvrabilité des contenus culturels francophones dans l'environnement numérique*[10]. La Loi inscrira dans la Charte des droits et libertés de la personne du Québec le droit à la découvrabilité des contenus culturels d'expression originale de langue française et à l'accès à de tels contenus. Le gouvernement créera également le Bureau de la découvrabilité des contenus culturels au sein du ministère de la Culture et des Communications (MCC) pour veiller au respect de la Loi et de ses éventuels règlements.

Cette nouvelle loi permettra au gouvernement d'adopter des règlements qui viendront clarifier plusieurs dispositions, notamment les suivantes :

- Les critères d'enregistrement des plateformes numériques auprès du MCC
- Les critères permettant de déterminer ce qui constitue du contenu culturel d'expression originale de langue française
- L'établissement de quotas relatifs aux contenus francophones
- Les règles en matière de découvrabilité de ces contenus

Il pourrait donc s'agir d'un nouveau souffle pour la protection et la promotion des contenus culturels québécois.

Des questions en suspens

1. L'encadrement de la radiodiffusion et des plateformes de diffusion continue relève traditionnellement de la **compétence fédérale** et est régi par le CRTC. Les prochains mois nous diront si le gouvernement du Québec réussira à appliquer des règlements provinciaux en la matière.
2. La question des **langues autochtones**, qui sont pourtant des langues locales et propres au territoire, est complètement absente de la nouvelle loi.

6 Exemple proposé par Sven Buridans, représentant la Fédération culturelle canadienne-française, lors d'un panel de Dialogues interdisciplinaires: L'IA et la diversité linguistique

7 Groupe de travail sur l'avenir de l'audiovisuel au Québec (2025). *Souffler les braises : Raviver le feu de notre culture et transmettre la flamme.* https://cdn-contenu.quebec.ca/cdn-contenu/adm/min/culture-communications/documents/GTAAQ/Rapport_Groupe_travail_audiovisuel_Quebec.pdf Figure adaptée du Graphique 9, p.44-45.

8 Témoignage de Marie-Claude Robichaud, vice-présidente Stratégie multiplateforme, recherche et innovation à Télé-Québec lors d'un panel de Dialogues interdisciplinaires: L'IA et la diversité linguistique.

9 Karadeglija, A. (2025, 18 septembre). Le CRTC examinera les exigences de contenu canadien dans les services en continu. *La Presse Canadienne.* https://www.lapresse.ca/affaires/medias/2025-09-18/le-crtc-examinera-les-exigences-de-contenu-canadien-dans-les-services-en-continu.php

10 Gouvernement du Québec (2025). *Le Québec adopte le projet de loi nº 109 affirmant sa souveraineté culturelle.* https://www.quebec.ca/nouvelles/actualites/details/le-quebec-adopte-le-projet-de-loi-no109-affirmant-sa-souverainete-culturelle-67602

Des usages bénéfiques de l'IA pour la diversité linguistique

La finalité des différentes critiques faites à l'IA vis-à-vis les langues n'est pas de l'écarter complètement ou d'interdire son utilisation. Au contraire, si la volonté est présente, les systèmes d'IA peuvent être conçus de manière à nous exposer à davantage de diversité linguistique et culturelle. Ils peuvent devenir des outils de revitalisation linguistique et faire revivre des langues menacées ou éteintes, par exemple en archéologie. Il s'agit donc d'une technologie qui a le potentiel de valoriser plusieurs langues et expressions culturelles, à condition que les projets soient menés avec et par les communautés elles-mêmes, en respectant la propriété intellectuelle des données et d'autres considérations culturelles.

Le cas de l'Islande

L'Islande, un pays d'environ 400 000 personnes, est au front de la question des langues à faibles ressources dans les grands modèles de langage. Depuis déjà plusieurs années, le gouvernement et des entreprises islandaises se sont associés à OpenAI pour entraîner des GML (en commençant par GPT-4) en islandais, comme tout premier projet pilote en la matière. Les retombées escomptées dépassent les frontières de l'Islande, en aidant « **à planifier et à orienter le travail pour d'autres langues moins parlées, en montrant qu'avec des efforts ciblés, même des langues ayant relativement peu de locuteurs peuvent s'épanouir à l'ère numérique** ».[11]

Principaux bénéfices attendus en Islande:

- **Accès élargi aux services numériques** traditionnellement offerts seulement en anglais
- **Inclusion des locuteurs non natifs et des touristes** en permettant des traductions de l'islandais vers d'autres langues
- **Équité concurrentielle pour les entreprises**, qui peuvent opérer dans la langue locale sans désavantage.
- **Soutien à l'apprentissage de l'islandais chez les enfants**, notamment par la traduction en temps réel de contenus vidéo et de jeux
- **Meilleure accessibilité pour les personnes en situation de handicap**, grâce à des outils d'assistance linguistique dans leur langue native.

D'autres initiatives d'adaptation de systèmes d'IA pour le public islandais, comme un partenariat entre Anthropic et le ministère de l'Éducation[12], pourront également informer nombre d'autres petites communautés linguistiques, dont les francophones québécois et canadiens, des clés du succès pour répliquer l'expérience.

11 Heimisdóttir, L. (2024, 24 septembre). The Icelandic Approach: Preserving and Revitalizing Linguistic and Cultural Diversity in AI. *Miðeind*. https://mideind.is/en/greinar/islenska-adferdin-hvernig-studla-ma-ad-fjolbreytni-tungumala-og-menningar-a-svidi, traduction libre

Rôles que peuvent jouer les acteurs publics

1. Inclure explicitement le secteur des arts et de la culture francophones dans les **investissements fédéraux en intelligence artificielle**[13]
2. Agir au sein des forums internationaux **économiques** et non seulement culturels
3. Miser sur des **projets pilotes d'adaptation d'outils d'IA** à la spécificité culturelle et linguistique québécoise
4. Créer des **espaces de dialogue entre les différentes communautés linguistiques** du Québec et du Canada et considérer leurs besoins et recommandations
5. Viser un **équilibre** entre la protection de nos données et de nos langues et la nécessité de demeurer évolutifs et ouverts sur le monde (éviter un repli sur soi)
6. Adopter des règlements qui exigent une forme **de discrimination positive dans les algorithmes**, mais aussi encourager la population à **s'intéresser aux contenus locaux et à les rechercher activement**
7. **Soutenir les créateurs et créatrices à la base** pour valoriser leur production culturelle
8. Reconnaître que les **langues autochtones** font partie intégrante de l'histoire et du territoire québécois et les inclure dans les lois et règlements visant à protéger et à promouvoir l**es langues québécoises**

Pour davantage de recommandations sur la souveraineté culturelle dans l'environnement numérique, consultez les actions prioritaires détaillées dans ce document.

12 Bryant, M. (2025, 15 novembre). Icelandic is in danger of dying out because of AI and English-language media, says former PM. *The Guardian*. https://www.theguardian.com/world/2025/nov/15/icelandic-is-in-danger-of-dying-out-because-of-ai-and-english-language-media-says-former-pm

13 Obvia (2025). Agir maintenant pour renforcer notre souveraineté culturelle face aux enjeux liés aux technologies numériques et aux systèmes d'intelligence artificielle - Actions prioritaires. https://www.obvia.ca/ressources/agir-maintenant-pour-renforcer-notre-souverainete-culturelle-face-aux-enjeux-lies-aux-technologies-numeriques-et-aux-sytemes-dintelligence-artificielle-actions-prioritaires

Messages clés pour la politique et la pratique

1 **La majorité des systèmes d'IA ne sont pas conçus pour desservir des langues à faibles ressources. Ceux-ci s'alimentent de contenus largement anglophones et occidentaux et ont du mal à saisir les nuances culturelles et linguistiques des communautés moins bien représentées dans l'espace numérique.**

2 **Cette problématique de fonctionnement des GML empêche les locuteurs de langues à faibles ressources ou de dialectes locaux de profiter pleinement des bénéfices attendus des outils d'IA.**

3 **Dans le domaine de la traduction, la contribution humaine demeure essentielle pour éviter la perte de sens des contenus culturels et pour préserver les nuances de ton et les régionalismes.**

4 **Un meilleur encadrement des algorithmes des grandes plateformes numériques est nécessaire pour garantir la découvrabilité des contenus locaux. Dans le cas du Québec, cela doit inclure les expressions culturelles en français québécois et en langues autochtones, propres à notre territoire.**

Restez à l'affût

Une synthèse complète de l'évènement « Dialogues interdisciplinaires: L'IA et la diversité linguistique » sera disponible prochainement sur le site web de l'Obvia.

Pour aller plus loin

Consulter la vidéo

À vos agendas

2e édition du colloque de la direction Collaboration avec les acteurs publics de l'Obvia

La souveraineté numérique au Québec : concilier stratégie, éthique et intérêt public

15 avril 2026
En présentiel à Québec

Détails et inscription

L'Obvia, au service des acteurs publics et de la société

L'Obvia identifie les enjeux sociétaux de l'intelligence artificielle et du numérique et contribue à des solutions qui placent les êtres vivants et la biosphère au centre de leur cycle de développement et d'utilisation. Notre communauté de recherche produit des connaissances ouvertes qui renforcent les capacités individuelles et collectives, en collaboration avec la société civile, les acteurs publics, l'industrie et les développeurs.

obvia

Pour nous contacter :

Observatoire international sur les impacts sociétaux de l'IA et du numérique

Pavillon Charles-De Koninck, local 2489
1030, avenue des Sciences-Humaines
Université Laval
Québec (Québec) G1V 0A6

collaboration@obvia.ca
418.656.2131 poste 401234

Pour en savoir plus

obvia.ca

Pour consulter les autres notes de breffage

https://doi.org/10.61737/VTLD2456 ISBN : 978-2-925581-29-1