

LATICCE
Laboratoire de recherche sur la découvrabilité
et les transformations des industries culturelles
à l'ère du commerce électronique



**MADE IN QUÉBEC - INDICE DE DÉCOUVRABILITÉ DE
LA MUSIQUE EN LIGNE**

**JEAN-ROBERT BISAILLON, DOCTORANT ET CO-
DIRECTEUR MUSIQUE LATICCE**

Recherche réalisée dans le cadre des travaux du LATICCE
sous la direction de Michèle Rioux financés par le programme Actions concertées
FRQSC-MCCQ Culture et numérique
*Mesure de la découvrabilité des
produits musicaux et audiovisuels
québécois sur les plateformes numériques*

AOUT 20192018

Remerciements

L'auteur remercie l'équipe du LATICCE pour leurs contributions aux efforts collectifs de recherche qui a alimenté le processus de rédaction de ce cahier de recherche ainsi que les organismes subventionnaires ayant contribué au financement de cette recherche.

**Fonds de recherche
Société et culture**

Québec



**Culture
et Communications**

Québec



**PLAN
CULTUREL
NUMÉRIQUE
DU QUÉBEC**



NOTRE CULTURE, CHEZ NOUS, PARTOUT

Les recherches sur la découvrabilité du LATICCE indiquent les faiblesses potentielles de la recommandation musicale dans un marché de niche

Typologie	5
Listes référentielles	6
Persona	6
Comportement	6
iD (Indice de découvrabilité)	6
P1 (Présence)	7
P2 (Présence)	7
V (Visibilité)	7
R (Recommandation)	7
Exemple	8
Questions et conclusions	Error! Bookmark not defined.
Annexe A - Tableau typologique	Error! Bookmark not defined.

Mots-clés : découvrabilité, musique, numérique, streaming, Apple Music, Spotify, Google Play, présence, visibilité, recommandation, rayonnement, profilage de goût.

Le LATICCE, laboratoire affilié au département de science politique de l'UQAM et au *Centre de recherches sur l'intégration et la mondialisation (CEIM)*, signifie Laboratoire de recherche sur la découvrabilité et les transformations des industries culturelles à l'ère du commerce électronique. Notre petite équipe d'environ 10 personnes travaille depuis 2 ans à la création d'un indice de découvrabilité permettant de mesurer les niveaux d'inventaire et la qualité des recommandations de répertoires de niche comme celui des artistes indépendants d'expression française au Québec. Nous pensons que nos données brutes peuvent favoriser des recherches plus approfondies sur la manière dont les algorithmes de profilage de goût peuvent générer une construction appropriée de listes de lecture de musique dans l'intérêt de la diversité des expressions culturelles. Nous pensons également que les plateformes de diffusion en continu doivent améliorer le ciblage des abonnés sur les marchés locaux pour conserver les niveaux d'adoption appropriés pour leurs produits, refléter les préférences des utilisateurs et leurs profils de goûts.

Le LATICCE a effectué des tests sur les services de streaming musical de mars à août 2019. Notre expérience singulière mesure la qualité de service des plateformes de diffusion transnationales dans un marché de niche substantiel, basé sur un répertoire de plus de 40 ans de production musicale indépendante.

La province canadienne du Québec possède une industrie du disque mature dans laquelle les labels et distributeurs multinationaux sont largement absents depuis les années 1980. Malgré cela, il a permis à diverses carrières internationales de se développer et de s'épanouir, notamment celles de Céline Dion, Cœur de Pirate, Grimes, Chromeo ou Patrick Watson. Un autre aspect de ce marché de niche est sa proportion majoritaire d'artistes d'expression française pouvant être isolés en termes de matériel de recherche et étudiés en tant que phénomène culturel spécifique. Très peu d'environnements culturels occidentaux ont un répertoire de musique populaire aussi circonscrit et largement adopté et consommé par ses habitants. Un article récent du quotidien *Le Devoir*¹ cite des sources qui mettent en garde les partis politiques fédéraux dans le contexte des prochaines élections canadiennes d'automne 2019 contre le risque d'érosion de cette culture spécifique. Solange Drouin, porte-parole de l'ADISQ, une association québécoise de labels indépendants, y affirme que les albums de musique sont passés de 13 millions d'unités vendues en 2005 à moins de 3 millions

¹ Delgado, « Ottawa appelé à protéger la culture ».

l'année dernière.

Notre méthode repose sur six niveaux de données: Présence, Visibilité et Recommandation, qui repose elle-même sur la concordance, la pertinence et la nouveauté (niveau de nouvelles parutions). Nos listes de référence sont constituées des nouveautés officielles et des candidatures historiques de 40 ans de gala des Félix, l'équivalent québécois des Grammy's ou des Juno. Nous avons construit des personas et des récits d'utilisateurs et effectué des requêtes par API, du moissonnage de sites Web (*Web Scraping*) et des observations humaines sur trois grands services d'écoute de musique en flux numérique (*Streaming*) : Apple Music, Google Play Music et Spotify.

Nos découvertes les plus intéressantes concernent notre persona de référence. Ce persona incarne une femme du début de la trentaine, qui a ouvert des comptes payants (*Premium*) sur les trois services étudiés. Elle a acheté, ajouté à sa bibliothèque et écouté sa musique préférée sur des listes de lecture éditoriales sélectionnées ainsi que sa propre liste de lecture personnalisée pendant 21 semaines.

Malgré l'accent mis par les services de musique sur la qualité de l'expérience musicale personnalisée offerte à leurs abonnés, aucune des 21 semaines d'écoute n'a offert ce que notre personnage attendait. Les services proposés ont réagi de manière très variée aux problèmes rencontrés dans une situation de *Cold Start* (possibilité de recommander du contenu pertinent avec très peu d'historique des données). Notre pire résultat étant Apple, qui ne pouvait pas générer de liste de lecture automatique après six semaines d'abonnement payant. Le plus intrigant étant Google avec des propositions dès le premier jour. Globalement, les services ont pris de très grandes libertés avec le profil d'intérêt musical de notre cobaye virtuel. Notre jeune femme obéit à un scénario de comportement très balisé, et a volontairement écouté une sélection de musique très limitée et cohérente : elle n'écoute que des artistes québécois qui chantent en français. Elle écoute ce type de contenu, tant de la part d'artistes de longue date, que de jeunes artistes hip-hop québécois. Seule constance, elle n'écoute QUE des chansons québécoises en français.

Notre indice montre qu'on lui a proposé un contenu de découvrabilité de 50/100, bien inférieur à nos attentes. Dans l'ensemble, les performances des services peuvent être caractérisées comme suit: Apple *New Music Mix* a fait preuve de peu de discrimination en ce qui concerne la notoriété des artistes, creusant trop souvent dans de nouvelles propositions ne s'appuyant sur aucun historique et mélangeant les genres et les langues d'expression de manière très chaotique; Google *J'ai de la chance* a de son côté offert un très haut niveau de concordance avec notre liste d'échantillons initiale et revenait constamment à des artistes traditionnels très attendus, générant potentiellement un haut niveau de lassitude; enfin Spotify *Radar des sorties* (Radar Release) possédait les meilleures qualités générales, mais offrant de très nombreuses propositions hors de notre persona étroitement défini, beaucoup trop de titres aux qualités distantes de notre cible.

Aucun des services de *streaming premium* n'a répondu aux attentes très précises de notre auditrice de niche.

Typologie

Une des étapes préliminaires nécessaire à la création de l'indice fut de dresser une typologie de l'offre de service des trois plateformes d'écoute musicale en flux numérique (streaming) qui ont fait l'objet de nos tests. Nous devions en effet identifier des éléments d'observation présentant des caractéristiques similaires et avons cru essentiel de nous livrer à ce travail de caractérisation et de comparaison des trois services.

Dans le cadre de cet exercice, nous nous sommes notamment penchés sur l'existence et le nom spécifique de listes d'écoute (playlists) éditorialisées ou générées par profilage de goûts pouvant présenter des caractéristiques suffisamment comparables pour effectuer nos tests de visibilité et de recommandation de façon concluante.

Notre tableau typologique de l'offre de service de la musique en ligne est annexé en toute fin de l'article et consultable ici pour un meilleur confort de lecture :

<https://docs.google.com/spreadsheets/d/1FQ-1m2hzwxORNT0oxYf2V2uvMi3tXqPRNmt2y-jxnKI/edit?usp=sharing>

Listes référentielles

Notre première liste référentielle (A) contient les 498 **pistes individuelles** nommées et gagnantes de la chanson de l'année et du vidéoclip de l'année de 40 ans de gala de l'ADISQ :

<https://docs.google.com/spreadsheets/d/1xPqiEmwPZ8IQNLoR-VQWNSFBcDpC14Tm3kRrdpEzk8/edit?usp=sharing>

Notre seconde liste référentielle (B) est construite sur une base hebdomadaire depuis mai 2019 avec toutes les chansons figurant à la liste des **nouveautés du vendredi** de l'ADISQ (en excluant les « nouveautés numériques » spécifiques à Apple) : <https://palmaresadisq.ca/fr/nouveautes/chansons/>

Persona

Les Persona sont soumises aux comportements suivants :

LATICCIA (profil témoin)

<https://docs.google.com/spreadsheets/d/1vAHtfbukwnuCNx5DYfOET0IqVQQBsqHwgKykZ7RyDz0/e/dit?usp=sharing>

La bibliothèque musicale de LATICCIA est constituée de 50 pistes extraites des années les plus récentes de la liste A et inclut tant de la grande pop (Ginette Reno) que du folk, du rock ou du hip-hop (Alaclair Ensemble). Elle a acheté ces titres sur Apple et créé des listes d'écoute personnelle sur chacune des plateformes constituée de 30 de ces titres (LaticciaPlaylist). Enfin, elle s'abonne à une liste éditoriale francophone par plateforme.

STAN (profil variable)

<https://docs.google.com/spreadsheets/d/11nSVeb4ptLpPc18wNJP3q5f7fGvfaavI247OIO4wqgE/edit?usp=sharing>

La bibliothèque musicale de STAN est constituée de 25 pistes extraites de la liste familiale qu'il partage avec LATICCIA, il achète 25 pièces additionnelles de la liste Billboard Alternative Hits 2016 et créé des listes d'écoute personnelle sur chacune des plateformes constituées de 30 de ces titres (StanPlaylist) 15 francophones et 15 alternatifs. Enfin, il s'abonne à une liste éditoriale alternative par plateforme.

Comportement

1. Chacun s'abonne à 10 amis sur Apple Music et Spotify.
2. Chacun crée un profil sur Last.fm pour enregistrer l'ensemble de ses écoutes réalisées sur les trois services. <https://www.last.fm/user/laticciarioux> / <https://www.last.fm/user/StanRioux>
3. Chaque semaine LATICCIA et STAN écoutent de façon aléatoire 5 pièces de leur liste d'écoute personnelle par plateforme (MyPlaylist) = 15
4. Chaque semaine LATICCIA et STAN écoutent de façon aléatoire 5 pièces de leur liste éditoriale par plateforme = 15
5. Chaque semaine LATICCIA et STAN relèvent par écrit dans Google Sheets, les 10 premiers titres qui apparaissent dans leur liste algorithmique personnalisée et créent ainsi l'indicateur R fondé sur le critères décrits plus bas.

iD (Indice de découvrabilité)

$$iD = P/100 + V/100 + R/100 / 3$$

Dans nos outils de visualisation, les indicateurs P-V-R sont segmentés par plateforme et par semaine dans le but de les comparer les unes aux autres.

L'exercice actuellement terminé consistait à établir une preuve de concept. Au moment de préparer l'actuel rapport, les observations automatiques P et V se poursuivent et les observations par humains R sont suspendues pour validation.

Pour le P et le V, notre preuve de concept fut établie sur les services Apple Music (iTunes) et Spotify, puisque les coûts associés à l'automatisation de Google Play Music sont élevés et que ce service prévoit bientôt subir une mutation majeure.

Pour le R, notre preuve de concept fut établie sur les services Apple Music (iTunes), Google Play Music et Spotify. Nous avons cessé ces observations le 19 août faute de moyens et par souci de vérifier l'adhésion du milieu à nos méthodes avant de poursuivre l'exercice de façon formelle.

P1 (Présence)

Le **P1 (Gala de l'ADISQ)** pourra s'améliorer ou se dégrader. Il s'agit de vérifier par procédé automatique, la présence de la liste référentielle A de semaine en semaine.

*Calculer la moyenne sur 100. **Disons 73,64***

P2 (Présence)

Le **P2 (Nouveautés courantes)** pourra s'améliorer ou se dégrader. Il est cumulatif. Il s'agit de vérifier par procédé automatique, la présence des nouveautés chanson de l'ADISQ (liste B) en continu et de noter la semaine d'apparition d'une nouveauté à l'inventaire. La semaine d'apparition ne fera toutefois pas l'objet du calcul.

*Calculer la moyenne sur 100 i.e. le nombre de nouveautés présentes cumulées / nouveautés totales cumulées = **89,74***

*Le P, une note sur 100, représentera le total de $P1 + P2 / 2$ – soit **73,64 + 89,74 / 2 = 81,69***

V (Visibilité)

Le **V (Nouveautés courantes dans les playlists éditorialisées)** pourra aussi fluctuer. Il n'est pas cumulatif.

Il consiste à mesurer par procédé automatique, de semaine en semaine, l'inclusion des nouveautés chansons (liste B) dans les listes éditoriales de nouveautés de Apple et Spotify la semaine suivant le vendredi de parution (le lundi suivant par exemple). Si les playlists éditoriales sont infinies, nous observons seulement les 30 premières positions.

Apple Music Accueil > Nouveau franco <https://music.apple.com/ca/playlist/nouveau-franco/pl.2e2171ec3ceb4e8aabcd696e7d8b41b6?l=fr>

Spotify Accueil > Nouveautés de la semaine Canada
(<https://open.spotify.com/playlist/37i9dQZF1DX9SvXmR7wQty>)

*Extrapolé sur 100 = **14,26** (notons la faiblesse de cette note)*

R (Recommandation)

Le **R (observations manuelles)** n'est pas un procédé automatique (mais pourrait le devenir une fois la preuve de concept améliorée et validée), il est hebdomadaire et établi sur les 10 premiers titres figurant aux listes d'écoute fondées sur le profilage des abonnés, par PNEL – soit :

Apple **New Music Mix** / Google **J'ai de la chance** (*I'm feeling lucky Dice App*) et Spotify **Radar des sorties** (*Release Radar*) – Voir article *Research on discoverability at LATICCE point to potential weaknesses of music recommendation in niche market* sur Medium

(<https://medium.com/@tgitjr/research-on-discoverability-at-laticce-point-to-potential-weaknesses-of-music-recommendation-in-da2be59f5363>)

Les trois indicateurs relevés pour chaque titre sont :

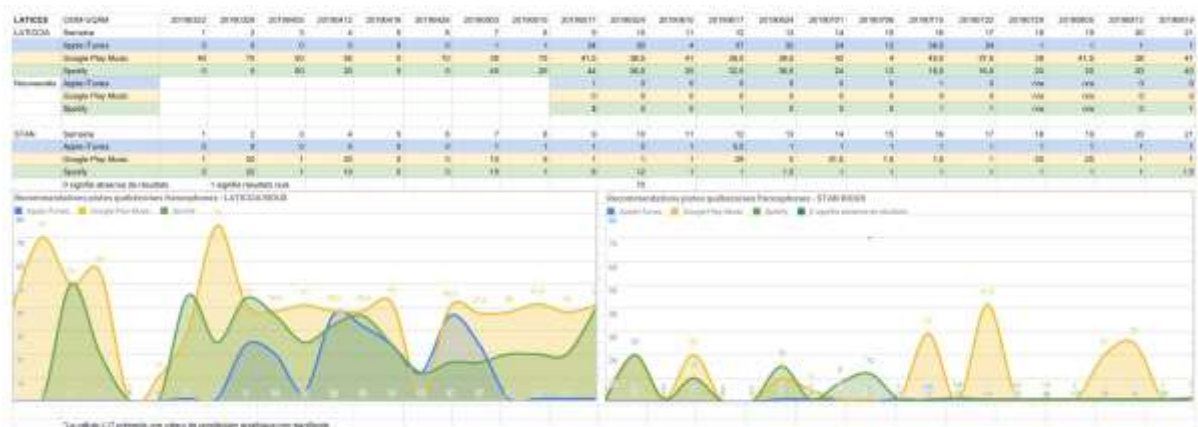
(c) concordance (lorsque l'interprète principal figure déjà au profil de l'abonné) = valeur maximale de 5 (10/2). La valeur maximale est limitée à 5 pour refléter le caractère lassant d'une trop forte concordance.

(p) pertinence (lorsque la piste est québécois et francophone - i.e. rencontre les attentes de notre abonné) = valeur maximale de 10.

(n) nouveauté (lorsqu'une proposition est une nouvelle parution recensée pour la période donnée) = valeur maximale de 10 (Remarque : Sommes-nous biaisés sur les nouveautés en utilisant que les nouveautés chansons ADISQ. Ajout des extraits des nouveaux albums?)

Total maximal de 25 - Le tout multiplié par 4 pour un résultat sur 100

$$= \text{SUM}((\text{CONCORDANCE})/2) + \text{SUM}(\text{PERTINENCE}:\text{NOUVEAUTÉ}) * 4$$



Visualisation R de l'été 2019

<https://docs.google.com/spreadsheets/d/1jkfXu3w5VtVul6dTWg0HsXOGsESSU8kbWiCJi7QGhxQ/edit?usp=sharing>

Exemple

ID (Indice de découvrabilité) pour la plateforme X et la semaine d'observation Y =

Spotify semaine du 12-08-2019 : $((P) 81,69 + (V) 14,26 + (R) 20) / 3 = (ID) 38,66$ (visualisation provisoire) https://docs.google.com/spreadsheets/d/123zIAYR3PRv74vXpWF_KulCUBvy2xeKma-wnGqi5rBA/edit?usp=sharing

Conclusion

Cet article d'abord publié en anglais et sommairement sur Medium, souligne que les services ne parviennent en aucun cas à témoigner de leur compréhension du comportement de LATICCIA qui ne souhaite écouter QUE du contenu de niche québécois francophone. Les ID calculés sont complètement chaotiques tant pour STAN que LATICCIA et ne remplissent pas la promesse commerciale faite aux abonnés de pouvoir leur proposer de la musique fondée sur leur profil de goût. Les services utilisent encore trop les méthodes basées sur la moyenne et le filtrage collaboratif de base pour créer les listes recommandées.

La première objection soulevée jusqu'ici est à l'effet qu'un profil comme celui de LATICCIA n'est pas "réaliste" fournissant ainsi un prétexte pour expliquer les failles des algorithmes. Est-ce le cas ou l'expérience doit-elle être poursuivie?

Les tests que nous n'avons pas faits consistent à observer comment l'ajout d'interactions comme les j'aime, j'ajoute à une playlist, les "skip" pourraient avoir une incidence sur la remontée de titres

francophones pour STAN. De même, nous devons observer comment l'ajout de métadonnées enrichies comme les paroles, la localisation territoriale ou les identifiants uniques pourrait influencer sur la qualité des recommandations pour LATICCIA.

Ainsi, (1) nous croyons avoir su créer une **méthodologie d'observation constante** basée sur les paramètres de service de l'offre en ligne;

Cette méthodologie permet effectivement (2) la **création d'un indice qui permet désormais de mesurer l'influence de mesures d'action industrielle** sur la remontée de titres québécois et de nouveautés québécoises dans l'offre.

Le projet du consortium à but non-lucratif MétaMusique consiste à la création d'outils de conditionnement de contenus et d'une base de métadonnées enrichies qui ont pour objectif d'activer une meilleure découvrabilité des propositions musicales québécoises dans l'offre en streaming ainsi qu'améliorer les mécanismes de reddition de comptes de l'industrie de l'enregistrement sonore.

Le Laticce estime que sa méthodologie de mesure de l'indice de découvrabilité peut permettre une validation des efforts industriels entrepris par MétaMusique. Cette validation permettrait la mise en place et l'ajustement progressif des protocoles industriels de conditionnement des contenus préalables à leur distribution et ainsi accroître la découvrabilité des enregistrements sonores sur les plateformes. Nous devons maintenant nous pencher sur la question des mesures à mettre en place pour pérenniser nos observations.