

Les algorithmes de recommandation musicale et l'autonomie de l'auditeur. Analyse des écoutes d'un panel d'utilisateurs de streaming

Jean-Samuel Beuscart, Samuel Coavoux, Sisley Maillard

▶ **To cite this version:**

Jean-Samuel Beuscart, Samuel Coavoux, Sisley Maillard. Les algorithmes de recommandation musicale et l'autonomie de l'auditeur. Analyse des écoutes d'un panel d'utilisateurs de streaming. Réseaux : communication, technologie, société, 2019, 213, pp.17-47. 10.3917/res.213.0017 . halshs-01639788v2

HAL Id: halshs-01639788
https://shs.hal.science/halshs-01639788v2

Submitted on 24 Apr 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

CAIRN.INFO
CHERCHER, REPÉRER, AVANCER

LES ALGORITHMES DE RECOMMANDATION MUSICALE ET L'AUTONOMIE DE L'AUDITEUR

Analyse des écoutes d'un panel d'utilisateurs de streaming

Jean-Samuel Beuscart, Samuel Coavoux et Sisley Maillard

La Découverte | « Réseaux »

2019/1 n° 213 | pages 17 à 47

ISSN 0751-7971
ISBN 9782348042782

Article disponible en ligne à l'adresse :

https://www.cairn.info/revue-reseaux-2019-1-page-17.htm

Distribution électronique Cairn.info pour La Découverte.
© La Découverte. Tous droits réservés pour tous pays.

La reproduction ou représentation de cet article, notamment par photocopie, n'est autorisée que dans les limites des conditions générales d'utilisation du site ou, le cas échéant, des conditions générales de la licence souscrite par votre établissement. Toute autre reproduction ou représentation, en tout ou partie, sous quelque forme et de quelque manière que ce soit, est interdite sauf accord préalable et écrit de l'éditeur, en dehors des cas prévus par la législation en vigueur en France. Il est précisé que son stockage dans une base de données est également interdit.

Document téléchargé depuis www.cairn.info - Orange Labs - - 161.106.4.5 - 09/04/2019 10h55. © La Découverte

LES ALGORITHMES DE RECOMMANDATION MUSICALE ET L'AUTONOMIE DE L'AUDITEUR

Analyse des écoutes d'un panel d'utilisateurs de streaming

Jean-Samuel BEUSCART
Samuel COAVOUX
Sisley MAILLARD

DOI: 10.3917/res.213.0017

Document téléchargé depuis www.cairn.info - Orange Labs - - 161.106.4.5 - 09/04/2019 10h55. © La Découverte

Le streaming est devenu un mode d'accès à la musique important. Les données d'usage disponibles sont encore partielles, mais les enquêtes suggèrent qu'environ 80 % des internautes français ont une pratique de la musique en streaming, via les plateformes audio ou vidéo. En France, on compte 13 millions d'utilisateurs mensuels du streaming audio, dont 3,7 millions d'abonnés payants ; les usages gratuits comme payants ont été multipliés par trois entre 2013 et 2017 (SNEP, 2018). Le succès des plateformes spécialisées (Spotify, Deezer) ou généralistes (YouTube) se traduit par une part accrue des revenus du streaming dans le chiffre d'affaires de l'industrie musicale – même si de nombreux acteurs trouvent cette part encore insuffisante : ils représentent, sur le marché français, 85 % des revenus numériques, et 42 % de l'ensemble du marché. Ces indicateurs en forte hausse témoignent dans l'ensemble de l'installation du streaming comme régime dominant de consommation de la musique.

L'environnement sociotechnique du streaming met l'auditrice dans une position spécifique, à maints égards différente de celle de l'auditrice de disque ou de musique téléchargée. Il met à sa disposition un catalogue virtuellement illimité, ayant une prétention à l'exhaustivité : Deezer comme Spotify offrent plus de 30 millions de titres immédiatement accessibles, YouTube plus encore. Pour guider l'usager dans ce catalogue, les services proposent une variété d'outils algorithmiques suggérant ou choisissant des titres : depuis l'outil de filtrage collaboratif (« les gens qui ont aimé ce titre que vous avez écouté ont aussi aimé celui-ci »), de recommandation thématique (« les autres titres de cet artiste, de ce sous-genre du hip-hop que vous avez écouté ») jusqu'aux radios thématiques (« radio inspirée de cet artiste ») ou aux playlists complètement « inspirées de vos goûts » (le « flow » de Deezer, les « découvertes de la semaine » de Spotify), en passant par les nombreuses « ambiances » mises en playlists pour s'adapter aux humeurs des usagers. Ces algorithmes sont en constante évolution, et sont fortement mis en avant dans les discours des entreprises comme des commentateurs du secteur.

Les plateformes de streaming se situent, en matière d'agentivité, à mi-chemin entre deux techniques d'écoute musicale : la collection de disques et la

Document téléchargé depuis www.cairn.info - Orange Labs - - 161.106.4.5 - 09/04/2019 10h55. © La Découverte

radio. Elles construisent une figure d'auditeur potentiellement plus assistée et dirigée que l'écoute de disque. Alors que, pour reprendre l'expression d'Antoine Hennion, « on choisit deux fois un disque », lors de son achat et lors de son écoute sur la platine (Hennion, Maisonneuve *et al*., 2000, p. 233), il est possible qu'un morceau ou album écouté en streaming n'ait jamais été « choisi » – c'est-à-dire consciemment et intentionnellement sélectionné dans un univers de possibles – mais entièrement suggéré ou mis en playlist par le service de streaming en fonction de logiques plus ou moins transparentes. Tout comme à la radio, l'auditrice peut se trouver dans la position de choisir un flux auquel elle délègue le choix ; mais ce flux est plus individualisé que celui, de masse, de la radio. Alors que cette dernière se fonde sur la répétition d'un petit nombre de morceaux choisis pour correspondre aux goûts attribués à un public cible décrit par des indicateurs sociodémographiques grossiers, notamment l'âge, le flux de streaming puise dans un large catalogue et adapte sa proposition au profil individuel. Ce flux peut être interrompu et réorienté à tout moment par l'utilisateur. Le choix n'est cependant pas entièrement idiosyncratique : les algorithmes des services de streaming s'appuient le plus souvent sur les comportements des utilisateurs au profil similaire pour sélectionner des propositions.

Tout comme le format MP3 contient une figure d'un utilisateur accumulateur compulsif (Sterne, 2006), les plateformes et leurs algorithmes dessinent une figure particulière de l'amateur de musique : il s'agit d'un auditeur défini par ses traces d'activité et pris en charge par des machines, guidé au sein de l'océan musical par des robots infaillibles. L'exercice du jugement de goût est suspendu, ou réduit à des acquiescements et refus (passer le morceau proposé). Pour reprendre les termes de Karpik (2007), le consommateur est ici passif (il délègue le choix) et hétéronome (il endosse les goûts portés par le dispositif). Cette figure est à l'opposé de la figure classique de l'amateur de musique, activement engagé dans le choix des musiques permettant la félicité et attentif à la définition et au maintien de ses goûts personnels (Karpik, 2007, p. 144). Cette figure de l'auditeur fais l'objet de nombreuses critiques, puisqu'elle dépouille l'amateur du contrôle de ses goûts, de son rôle moteur dans l'interaction avec les œuvres, de son activité parfois virtuose dans l'entretien de ses passions.

C'est à la réalité de ce modèle que s'intéresse notre recherche : dans quelle mesure la figure dessinée par les interfaces se réalise-t-elle, se traduit-elle dans les comportements réels des auditeurs ? Les utilisateurs suivent-ils les

Document téléchargé depuis www.cairn.info - Orange Labs - - 161.106.4.5 - 09/04/2019 10h55. © La Découverte

recommandations des algorithmes ? Dans quels contextes, avec quelle régularité, quels effets ?

Pour répondre à ces questions, nous nous appuyons sur un corpus de traces d'activité d'un échantillon aléatoire de 4000 utilisateurs d'une plateforme de streaming musical, observés pendant une période de cinq mois, au cours de laquelle ils ont réalisé collectivement 17 millions d'écoutes[1]. Pour environ la moitié d'entre elles, il est possible de caractériser le contexte de l'interface dans lequel elles ont été réalisées, notamment dans quelle mesure elles ont été induites par les algorithmes. À partir de ces données, nous commençons par décrire les usages qui sont faits du service de streaming, en termes d'intensité et de variété des écoutes. Nous suggérons ensuite que la figure de l'auditeur algorithmisé est encore (très) minoritaire dans les usages ; l'essentiel des écoutes est issu d'un choix explicite et conscient de l'auditeur, et les modes d'exploration actifs et autonomes dominent. L'enjeu est de préciser dans quels contextes, et par quels types d'auditeurs, les algorithmes de recommandation sont suivis : existe-t-il plutôt des auditeurs particulièrement passifs, ou des moments de passivité partagés par tous les auditeurs ? Notre étude suggère la primauté du contexte d'écoute dans le choix des dispositifs.

Cet article s'inscrit dans le cadre d'une réflexion plus globale sur l'agentivité des algorithmes, sur la capacité des individus à s'appuyer sur ces dispositifs sans pour autant leur abandonner le contrôle de leur expérience. Il s'agit de comprendre comment, dans un monde d'agentivité de plus en plus distribuée entre l'amateur et ses prothèses numériques, se construisent des situations d'autonomie et d'hétéronomie.

LA RECOMMANDATION ALGORITHMIQUE, ASSISTANCE OU CONTRÔLE ?

Dans la littérature, les algorithmes de recommandation tels que ceux mis en place par les plateformes de streaming sont traités de deux façons principales : les travaux en économie et en marketing s'efforcent de mesurer leur efficacité, tandis que les travaux issus des *media studies* mettent l'accent critique sur les potentialités de contrôle portées par les dispositifs.

1. Une première présentation de ces résultats figure dans la thèse de Sisley Maillard (2015).

Document téléchargé depuis www.cairn.info - Orange Labs - - 161.106.4.5 - 09/04/2019 10h55. © La Découverte

La mesure de l'efficacité des recommandations

Une première tradition d'analyse s'intéresse aux formes d'influence et de recommandation sur la consommation. Depuis les travaux de Katz et Lazarsfeld (2008), de nombreuses recherches explorent, dans le cas des biens culturels notamment, le rôle respectif de l'entourage, des critiques, de la publicité, etc., sur la décision d'acheter un disque ou d'aller voir un film (pour une synthèse, voir Beuscart et Mellet, 2012). Si le terme parfois utilisé d'« influence » peut laisser penser le contraire, ces travaux discutent peu la légitimité des différentes formes de sollicitation, et sont avant tout centrés sur la question de l'efficacité relative des différentes formes de « recommandation » qui sont faites au consommateur : il s'agit de montrer par exemple que la critique n'a d'effet sur l'audience des films que s'ils appartiennent à la catégorie « art et essai », tandis que la publicité accroît significativement les audiences des blockbusters (Prag et Casavant, 1994).

Dans cette perspective, les différentes formes de guidage numérique sont perçues comme de nouveaux outils de recommandation inscrits dans la continuité d'autres dispositifs. Du fait de la disponibilité des données, de nombreux travaux s'intéressent ainsi au « bouche-à-oreille numérique » (*electronic word-of-mouth*, eWoM) : ils étudient l'influence des avis de consommateurs sur les ventes de livres ou de disques (Chevalier et Mayzlin, 2006), l'impact des classements sur l'audience des morceaux de musique (Salganik, Dodds *et al*., 2006), les effets des mentions sur les médias sociaux sur les ventes (Bastard *et al.,* 2012), le lien entre le nombre de tweets à propos d'un film et les entrées au box-office (Asur et Huberman, 2010). En discussion plus ou moins explicite avec l'hypothèse de la longue traîne, la numérisation des pratiques culturelles est vue comme un double mouvement : d'un côté l'élargissement de l'offre, tendant vers l'illimité, de l'autre l'enrichissement des prises offertes aux consommateurs pour se faire un avis et naviguer dans cette abondance (Brynjolfsson, Hu *et al.*, 2011).

Les recherches sur la recommandation algorithmique proprement dite sont plus rares, du fait de la difficulté d'obtenir des données permettant d'évaluer leur efficacité. La technique de recommandation le plus souvent utilisée est le filtrage collaboratif, qui établit des proximités statistiques entre les clients selon leurs paniers d'achats, et propose un retour à la manière des dispositifs d'Amazon : « les clients qui ont aimé ce que vous venez d'acheter ont aussi aimé… ». Comment mesurer l'efficacité de ces systèmes ? Une partie non

Document téléchargé depuis www.cairn.info - Orange Labs - - 161.106.4.5 - 09/04/2019 10h55. © La Découverte

négligeable des achats sur les sites de e-commerce se font suite à des liens de recommandations algorithmiques (jusqu'à 15 % sur Amazon) ; mais il est plus difficile de mesurer quels sont, parmi ces achats, ceux qui n'auraient pas été conclus en l'absence d'algorithme (Sharma, Hofman *et al*., 2015). L'usage de recommandations affecte les ventes, et la recommandation algorithmique est meilleure que la recommandation humaine dans ce domaine (Senecal et Nantel, 2004). Les travaux évaluant les effets des recommandations algorithmiques portent surtout sur leurs effets sur la diversité des contenus : parce qu'il est fondé sur la cooccurrence des consommations, le filtrage collaboratif peut favoriser l'enfermement dans un portfolio de produits très similaires (Konstan et Riedl, 2012). Alors qu'il est pensé comme une façon de favoriser la découverte d'artistes peu connus, il peut aussi paradoxalement renforcer la place des stars (Celma et Cano, 2008 ; Fleder et Hosanagar, 2009).

Critique de la raison algorithmique

Un second ensemble de travaux prend pour point de départ le rôle croissant des algorithmes dans nos sociétés. Issus de la sociologie, de la communication, de la théorie des médias, ils documentent la prolifération des agents socio-informatiques dans différents domaines de la vie sociale (finance, culture, médias, consommation, travail). Une grande partie adopte une perspective inquiète quant au pouvoir pris par ces algorithmes. Ils soulignent l'absence de transparence, d'*accountability* de ces systèmes : ils agissent comme des boîtes noires, selon des principes qui ne sont ni clairs, ni connus, ni discutés, ni discutables (Pasquale, 2016 ; Gillespie, 2016). Ils insistent en outre sur leur effectivité : les algorithmes transforment le monde, ils représentent un pouvoir mal perçu sur les personnes et leurs environnements (Galloway, 2012 ; Lash, 2007).

Les travaux plus empiriques soulignent néanmoins le caractère pluriel et difficile à saisir des algorithmes. S'ils ont un pouvoir, le lieu de ce pouvoir est souvent fuyant. Comme le rappellent Seyfert et Roberge (2016), il n'existe pas en général d'interprétation univoque de l'algorithme et de ses effets. Ils doivent toujours être dépliés, ils ne peuvent qu'être interprétés ; ceci d'autant plus qu'ils sont le plus souvent imbriqués avec d'autres, interconnectés, pris dans des contraintes techniques locales. Leurs effets sont dépendants des contextes, évoluent avec les usages qui en sont faits et en fonction des données de leur environnement. Leur performativité répond moins au modèle de la règle à laquelle on obéit qu'à celui d'une agentivité distribuée entre les lignes de codes et les différentes catégories d'utilisateurs (Beuscart et Mellet,

Document téléchargé depuis www.cairn.info - Orange Labs - - 161.106.4.5 - 09/04/2019 10h55. © La Découverte

2016). Seyfert et Roberge (2016) proposent de parler de « cultures algorithmiques » pour désigner ces ensembles de performativités et de significations qui émergent des usages des algorithmes.

Des travaux récents s'efforcent de produire une description empirique des principes de construction des algorithmes de recommandation musicale, qui permettent de mieux comprendre la diversité et l'évolution des figures de l'utilisatrice qui y sont construites. Prey (2018) propose ainsi une comparaison entre les systèmes de recommandation respectifs des services Pandora et Spotify. Le premier repose sur une définition plutôt essentialiste de la musique comme de l'auditeur : les morceaux y sont catégorisés par leurs « gènes » musicaux, au nombre de 150 à 500 selon les fichiers, gènes qui sont tous relatifs à un caractère purement musical (tonalité, distorsion de la guitare, rythmique, etc.). Symétriquement, l'auditeur est défini par sa réception positive à certains « gènes », c'est-à-dire par son appétence pour certaines caractéristiques purement musicales. À l'inverse, le système de recommandation central de Spotify se distingue par son caractère hybride, combinant caractéristiques du fichier musical et appréciations laissées par les internautes partout sur le web : la figure de l'internaute y est également plus fluide, plus susceptible d'évoluer en fonction des préférences de ses voisins. Seaver (2019) s'appuie quant à lui sur un travail ethnographique de plusieurs années auprès d'un service de recommandation musicale. Il décrit finement la façon dont la représentation et la mesure de la félicité de l'auditeur ont évolué au cours des années récentes : jusqu'au début des années 2010, les systèmes de recommandation étaient évalués par leur capacité à bien prédire le goût de l'auditeur, mesuré par la note qu'il attribue au produit. Progressivement, ce mode de représentation et de calcul a fait place à d'autres métriques centrées sur la capacité à retenir l'internaute ; l'absence de fuite vaut pertinence. De devins des goûts des utilisateurs, les systèmes de recommandation ont été reconfigurés, selon Seaver, en pièges à attention. Ces travaux soulignent ainsi à la fois la diversité des figures des cibles de la recommandation et l'évolution des métriques de pertinence vers la rétention de l'utilisateur.

Le projet des plateformes de streaming

Les principales plateformes de streaming offrent un rôle croissant aux algorithmes de recommandation. Au fur et à mesure de leur évolution, les services tels que Deezer, Spotify, LastFM développent des outils de suggestion

Document téléchargé depuis www.cairn.info - Orange Labs - - 161.106.4.5 - 09/04/2019 10h55. © La Découverte

de titres, les diversifient et leur accordent une place plus centrale dans le design du site. Alors que les pages d'accueil de Deezer ou de Spotify présentaient initialement les albums écoutés et archivés par l'utilisateur, à la manière d'une discothèque personnelle, elles organisent aujourd'hui une variété de suggestions d'artistes, d'albums et de playlists inspirés des écoutes passées des utilisateurs ; en outre, elles mettent en avant diverses modalités de flux musicaux (radio de l'artiste, découvertes de la semaine, ambiances, « flow », etc.) qui prennent complètement en charge le choix de la musique à diffuser. Sur YouTube, chaque écoute de musique est automatiquement suivie d'une autre sélectionnée par l'algorithme de recommandation, qui élabore la playlist de façon ininterrompue en l'absence d'action de l'utilisateur. Eriksson *et al*. (2019) documentent ainsi la façon dont, à partir de 2013, le format de la playlist a été mis en avant sur Spotify pour devenir progressivement le vecteur principal de recommandation algorithmique. Le centre de recherche de Spotify évoque le projet d'un « zero button music player », soit d'un service qui choisisse la musique en l'absence de toute action de l'utilisateur, à partir de la connaissance de ses habitudes, ses écoutes passées et des informations issues de son smartphone (Melançon, 2015). Se dessine ainsi une figure de l'utilisateur pris en charge par l'algorithme, déchargé de la tâche de choisir la musique du moment. Morris (2015), dans une étude du service de profilage musical the Echo Nest racheté par Spotify en 2014, décrit ainsi la figure d'un utilisateur dessiné de façon très précise à travers ses goûts, classé au sein de microsegments de marché, et nourri par des flux très individualisés. Il note néanmoins le maintien d'une figure complémentaire de la curation humaine, via les playlists partagées ; la culture algorithmique qui émerge sur ces sites reste – pour le moment du moins – hybride. Symétriquement, des études qualitatives se penchent sur l'expérience musicale des services de streaming, du moins sur les discours que ses utilisateurs portent sur leurs pratiques et les figures qu'ils dessinent. Hanrahan (2016) suggère, à partir de *focus groups*, que le modèle de l'écoute en streaming est celui du flux satisfaisant, de l'évitement de la mauvaise surprise, proche d'une radio très individualisée et capable de s'adapter au moment, mais à forte distance du modèle classique de l'amateur faisant l'expérience esthétique d'une œuvre ; autrement dit, les utilisateurs interrogés viennent habiter, au moins dans leurs discours en *focus group*, la figure proposée par les plateformes. Dans le même registre, Hagen (2016) évoque une pluralité des métaphores de l'usage du streaming, qui toutes s'éloignent de la figure de l'auditeur choisissant de prêter attention à une œuvre.

Document téléchargé depuis www.cairn.info - Orange Labs - - 161.106.4.5 - 09/04/2019 10h55. © La Découverte

L'écoute de musique apparaît ainsi comme un cas emblématique d'un questionnement plus large sur la performativité des algorithmes ; leur omniprésence sur les plateformes de streaming dessine une figure de l'auditeur calculé et pris en charge, en opposition à la figure de l'amateur explorant les œuvres. Au-delà de cette figure extrême, brandie tant par les développeurs des plateformes que par les critiques les plus virulents, il est très probable que la culture algorithmique qui émerge sur des usages est plurielle et composite.

UN PANEL D'UTILISATEURS D'UNE PLATEFORME DE STREAMING

Nos données permettent d'en esquisser des grandes lignes. Elles décrivent la consommation d'un échantillon d'utilisateurs d'une plateforme de streaming musical, suivis pendant cinq mois. L'échantillon est constitué d'environ 4 000 utilisateurs inscrits en France, sélectionnés aléatoirement parmi ceux qui ont utilisé la plateforme au cours de la première semaine d'avril 2014. L'échantillonnage aléatoire assure la meilleure représentativité possible pour la population incluse : les personnes employant le service au moins une fois par semaine. Les usagers plus occasionnels sont eux sous-représentés dans notre échantillon. Dans la mesure où la consommation au moins hebdomadaire de musique concerne près des deux tiers de la population, et plus encore dans les classes d'âge les plus jeunes, jusqu'à 95 % des moins de 18 ans (Donnat, 2009), classes d'âge qui sont plus susceptibles d'employer des plateformes de streaming, ce biais d'échantillonnage n'est pas excessivement préoccupant. Il n'en reste pas moins que nos résultats valent d'abord pour les utilisateurs de streaming, et parmi eux pour les utilisateurs réguliers.

Leurs écoutes ont été enregistrées du 1er avril au 31 août 2014, aboutissant à une base d'un peu plus de 20 millions d'observations. Les données collectées nous renseignent à la fois sur le profil des utilisateurs (âge, sexe, lieu de résidence, date d'inscription, abonnement, titres et artistes favoris), sur leur consommation (date, durée, supports d'écoute) et sur les caractéristiques des contenus écoutés (artiste, titre, date de sortie). En particulier, la variable « contexte d'écoute » nous permet d'appréhender par quel type de dispositif, de recommandation ou de collection, un consommateur a sélectionné un titre à écouter.

La plateforme propose un catalogue de 35 millions de titres, et deux offres illimitées, sur le modèle économique du « freemium » : l'une est gratuite et

Document téléchargé depuis www.cairn.info - Orange Labs - - 161.106.4.5 - 09/04/2019 10h55. © La Découverte

financée par la publicité, l'autre est payante, à un tarif de 9,99 euros par mois. La formule payante permet à l'utilisateur d'écouter sa musique en haute qualité, sans publicité, et surtout en mode hors connexion. Une offre intermédiaire à 4,99 euros par mois, aujourd'hui disparue, n'offrait pas de mode hors connexion. Toutes les offres sont disponibles sur l'ordinateur et les terminaux mobiles, comme les smartphones et les tablettes (via des applications pour iOS, Android et Windows).

Après retraitement des valeurs aberrantes, notre échantillon compte précisément 3 813 utilisateurs inscrits en France. Il s'agit d'une population plutôt masculine, jeune et urbaine. On compte seulement 37 % de femmes au sein de l'échantillon. L'âge médian des utilisateurs est de 28 ans et 80 % de l'échantillon a entre 17 et 46 ans. Notre échantillon est très urbain. Parmi les 75 % d'utilisateurs à avoir renseigné leur lieu de résidence, 14 % déclarent résider à Paris (qui rassemble un peu plus de 3 % de la population française) et 73 % vivent dans un grand pôle urbain (contre 59 % de la population française). Ils sont dans l'ensemble plutôt expérimentés : la majorité des utilisateurs (53 %) ont rejoint le service depuis plus de deux ans, et 18 % depuis plus de cinq ans. Seul un quart des utilisateurs se sont inscrit assez récemment, il y a moins d'un an.

Le mobile et l'ordinateur sont les supports privilégiés d'écoute (la tablette et le navigateur web, les deux autres options possibles sont marginales). En moyenne, sur l'ensemble de l'échantillon, un utilisateur se sert du support mobile pour deux tiers de ses écoutes (63 %) et utilise l'ordinateur pour le tiers restant. Le type d'offre souscrit influence évidemment les supports utilisés. Parmi nos utilisateurs, 58 % ont souscrit à l'offre premium à 9,99 euros, qui permet l'écoute illimitée sur mobile et le mode écoute hors connexion, 27 % ont souscrit à l'offre premium à 4,99 euros, sans publicité, et 15 % à l'offre gratuite incluant de la publicité. Les abonnés aux offres premium, sans publicité, utilisent plus fréquemment, parfois même presque exclusivement, le mobile pour écouter de la musique, contrairement aux abonnés gratuits, qui utilisent plutôt l'ordinateur (tableau 1). Néanmoins, même pour les abonnés payants qui en bénéficient, l'option hors connexion, pour écouter la musique hors ligne, est relativement peu utilisée par les abonnés. Les écoutes qu'ils ont réalisées hors ligne représentent seulement 11 % de leurs écoutes, principalement sur le mobile.

Document téléchargé depuis www.cairn.info - Orange Labs - - 161.106.4.5 - 09/04/2019 10h55. © La Découverte

Tableau 1. Utilisation des supports en fonction du type d'abonnement

Offre	Support	Fréquence moyenne d'utilisation
Gratuite	Téléphone	11,80 %
	Ordinateur	83,10 %
	Tablette	2,60 %
	Navigateur web	2,50 %
	Total	*100,00 %*
Payante	Téléphone	70,40 %
	Ordinateur	27 %
	Tablette	2 %
	Navigateur web	0,60 %
	Total	*100,00 %*

Champ : tous les usagers.

Source : auteurs.

LA DISTRIBUTION DES ÉCOUTES EN *STREAMING*

Une première façon de caractériser les auditeurs, en regard notamment de la figure de l'utilisateur « captivé » (Seaver, 2019) par les plateformes et promené dans la variété du catalogue immense consiste à s'intéresser à l'intensité et la diversité de leurs consommations. Nous commençons par mesurer la diversité des écoutes des différents utilisateurs, avant d'observer dans quelle mesure elles se répartissent équitablement entre stars et artistes de la longue traîne.

Volume et diversité des écoutes

Au cours de la période d'observation de cinq mois, on enregistre plus de 17 millions d'écoutes par les usagers de l'échantillon. Cependant, ce chiffre doit être pondéré par la durée réelle d'écoute des titres sur la plateforme : seule un peu plus de la moitié des écoutes (56 %) vont jusqu'à la fin du titre, et un tiers des écoutes ne durent pas plus de 30 secondes. Ces écoutes ont porté sur près d'un million de titres différents. Parmi les titres écoutés au moins une fois plus de 30 secondes (81 %), un titre est écouté en moyenne 16 fois par notre

Document téléchargé depuis www.cairn.info - Orange Labs - - 161.106.4.5 - 09/04/2019 10h55. © La Découverte

échantillon de 3 813 personnes observées durant cinq mois. Cela signifie qu'un titre moyen est écouté une fois par mois pour chaque millier d'usagers actifs. La moyenne est cependant trompeuse, car la distribution du nombre d'écoutes est, sans surprise, très asymétrique : 39 % des titres ont été écoutés une seule fois par un seul utilisateur, tandis que 5 % seulement l'ont été plus de 50 fois. Comme attendu, le « top » des titres affiche des audiences considérables : les cinq titres les plus écoutés au cours de la période de cinq mois ont cumulé chacun plus de 25 000 écoutes et chacun a été écouté par presque la moitié des individus de l'échantillon. Cela correspond donc à une moyenne de 13 écoutes par usager ayant écouté le titre au moins une fois. Les titres concernés sont *Stolen Dance* (Milky Chance), *Sur ma route* (Black M), *Rather Be* (Clean Bandit), *Waves* (Mr. Probz) *et Liar Liar* (Cris Cab feat. Pharell Williams).

L'intensité d'usage est assez élevée (tableau 2) : en moyenne, les utilisateurs se sont connectés à la plateforme 86 jours au cours des cinq mois (soit un jour sur deux), pour écouter près de 50 titres chaque jour de connexion. Ces moyennes cachent bien sûr des disparités très importantes, comme le résume le tableau ci-dessous : les 10 % les plus actifs ont réalisé plus de 10 000 écoutes au cours de la période, tandis que les 10 % les moins actifs en ont réalisé moins de 450.

Tableau 2. Intensité de la consommation musicale

Variables	Moyenne	Écart-type	Min	Premier décile	Premier quartile	Médiane	Troisième quartile	Neuvième décile	99e centile	Max
Nombre d'écoutes totales	4 617	5 064	101	422	1 180	3 006	6 202	10 795	24 212	41 210
Nombre de jours d'utilisation de la plateforme	86	46	1	17	46	93	127	145	153	154
Nombre moyen d'écoutes lors des jours d'utilisation	47	36	3	17	25	36	57	88	187	412
Consommation quotidienne de musique (durée d'écoute)	1 h 50 m	1 h 39 m	0 h 1 m	0 h 30 m	0 h 51 m	1 h 23 m	2 h 13 m	3 h 39 m	8 h 28 m	19 h15 m

Champ : tous les usagers.
Source : auteurs.

Document téléchargé depuis www.cairn.info - Orange Labs - - 161.106.4.5 - 09/04/2019 10h55. © La Découverte

Comment caractériser la diversité consommée sur les plateformes de streaming ? La promesse est que, lorsque l'offre est vaste, gratuite ou forfaitaire, et les outils de recommandation plus perfectionnés, la consommation musicale devrait être plus diversifiée sur les plateformes en streaming que sur les marchés existants du disque et du téléchargement. Notons d'abord que la diversité est difficile à mesurer et à résumer, tant il s'agit d'un concept multidimensionnel (Stirling, 2007). Dans cet article, nous ne nous intéressons qu'à une de ces dimensions, la variété, mesurée par le nombre d'artistes différents écoutés. Pour notre échantillon d'environ 4 000 utilisateurs, les écoutes portent sur 925 555 titres différents, soit 3 % des titres présents dans le catalogue. Elles correspondent également à un peu plus de 3 % des artistes et 6 % des albums proposés sur la plateforme (114 182 artistes et 226 939 albums différents). Ces chiffres portent cependant sur un échantillon. Une extrapolation grossière, à partir de ces résultats, laisse penser que la quasi-totalité du catalogue serait consommée à un moment ou à un autre par deux millions d'abonnés. Cependant, nous ne disposons pas de moyen de savoir combien de personnes sont réellement actives sur la plateforme et d'estimer ainsi la diversité globale de l'écoute. Donnat (2018), s'appuyant sur des données GfK, estime que sur l'ensemble du catalogue de Deezer, 2 millions de titres font l'objet de plus de 20 écoutes par les utilisateurs français. Ces estimations comme les nôtres renforcent le constat d'un utilisateur installé en régime d'abondance, pour qui « le caractère quasi illimité de l'offre […] se traduit par une variété consommée nettement supérieure à celle observée sur les autres modes d'accès à la musique ». Certes, la comparaison de cette variété avec d'autres médias est délicate. Les conditions d'écoute ne sont pas les mêmes : par rapport au disque ou au téléchargement, l'écoute d'un morceau en streaming n'implique pas de dépense supplémentaire ; par rapport à la radio, le rôle du consommateur dans le choix des morceaux est – jusqu'à preuve du contraire – plus actif, et l'offre est bien moins contrainte. Par ailleurs, les données issues des études de la diversité de l'industrie musicale (Bourreau, Moreau *et al*., 2011) concernent des flux (le nombre d'albums vendus par semaine), et non des stocks (le nombre d'albums possédés et écoutés) et sous-représentent donc la consommation effective. On constate cependant que le nombre de titres écoutés sur notre plateforme par seulement un peu moins de 4 000 utilisateurs au cours de cinq mois représente environ 9 fois plus de titres que la diffusion de l'ensemble des radios françaises au cours d'une année (Observatoire de la musique, 2015).

Document téléchargé depuis www.cairn.info - Orange Labs - - 161.106.4.5 - 09/04/2019 10h55. © La Découverte

La variété de la consommation varie fortement selon les individus (tableau 3). Le tableau ci-dessous décrit le nombre de titres et d'artistes différents écoutés par l'échantillon. Un utilisateur écoute en moyenne 1 149 titres différents ; 5 % en ont écouté le triple, alors qu'un quart de notre échantillon n'en a pas écouté plus de 359.

Tableau 3. Variété de la consommation musicale

Variable	Moyenne	Écart-type	Min	Premier décile	Premier quartile	Médiane	Troisième quartile	Neuvième décile	99e centile	Max
Nombre de titres différents écoutés	1 149	1 211	9	157	359	789	1 502	2 519	5 782	13 042
Nombre d'artistes différents écoutés	402	412	1	57	129	278	527	877	2 028	3 811

Champ : tous les usagers.

Source : auteurs.

L'examen des corrélations linéaires entre les indicateurs de variété et les variables dont nous disposons laissent deviner une inclination plus forte à la diversité de certains utilisateurs. Ces corrélations doivent être lues avec prudence, dans la mesure où il ne s'agit que d'analyses bivariées, qui portent sur des données agrégées contenant beaucoup de bruit. À titre exploratoire, elles suggèrent plusieurs liens entre profils des auditrices et variété consommée. D'un côté, la variété est corrélée positivement et significativement avec le temps écoulé depuis l'inscription sur la plateforme, avec l'intensité de la consommation musicale, avec la fréquence d'écoute des radios et avec le nombre de dispositifs de recommandation utilisés. L'ensemble de ces variables peuvent fonctionner comme des indices, sinon d'expertise, du moins d'engagement dans la pratique d'écoute musicale, désignant des personnes écoutant plus de musiques par plus de moyens ; il n'est pas surprenant qu'elles soient globalement associées avec une plus grande variété de consommation. La variété est également corrélée avec l'âge des utilisateurs, ce qui suggère que les consommateurs les plus jeunes ont des pratiques d'écoutes moins diversifiées. De l'autre côté, on observe une corrélation négative et significative de la variété consommée avec la fréquence d'écoute des nouveautés et des artistes stars. Autrement dit, la consommation la plus

Document téléchargé depuis www.cairn.info - Orange Labs - - 161.106.4.5 - 09/04/2019 10h55. © La Découverte

diversifiée est le fait d'utilisateurs plus intensifs, plus anciens dans leur usage de la plateforme, plus enclins à utiliser les outils de recommandation ; il s'agit d'utilisateurs plus âgés. Ils s'opposent tendanciellement à des utilisateurs plus jeunes, écoutant des artistes connus sur leur mobile. Ces résultats demeurent préliminaires : il s'agit de pistes dont l'exploration systématique dépasse le cadre de cet article.

Stars *vs* Longue Traîne

Sur la plateforme, l'audience des différents artistes est très hétérogène : la moitié des titres et des artistes ne sont écoutés que par un seul individu de l'échantillon, alors que le top 1 % des titres et des artistes est écouté par plusieurs centaines d'entre eux ; l'artiste le plus écouté l'a été par les trois quarts de l'échantillon. 10 % des artistes concentrent 90 % des écoutes, ce qui dessine une courbe de longue traîne tout à fait classique. Pour approfondir le phénomène, nous divisons notre population d'artistes en trois groupes :

- les stars : les 0,1 % les plus écoutés, soit 116 artistes écoutés chacun par au moins 25 % de l'échantillon ;
- la mid-tail, des 0,1 % aux 10 % d'artistes les plus écoutés, divisée en deux sous-catégories :
    - Haute : les 1 % d'artistes les plus écoutés, moins les stars, soit 1 038 artistes écoutés chacun par au moins 5 % de l'échantillon ;
    - Basse : des 1 % au 10 % d'artistes, les plus écoutés, soit 10 363 artistes écoutés par au moins 18 usagers différents, soit un peu moins de 0,5 % de l'échantillon ;
- la long-tail : les 90 % d'artistes les moins écoutés, qui ont pu n'être écoutés qu'une seule fois par un seul usager.

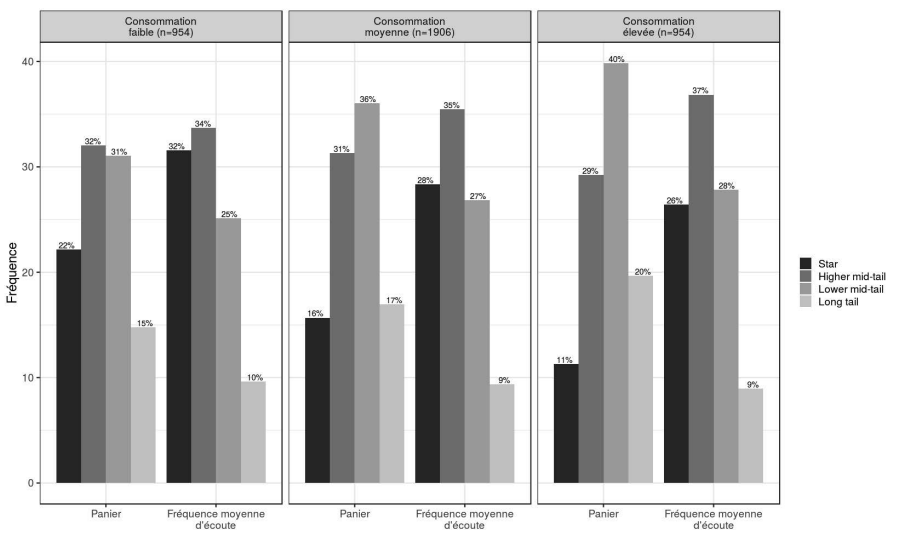
Sur la base de cette catégorisation, il est possible d'analyser le panier moyen du consommateur (figure 1). Le premier constat est que, au sein de l'échantillon, presque tous les individus consomment des artistes *stars* ; en outre, ces artistes constituent une part importante de leur consommation. Le panier « moyen » d'un consommateur surreprésente les artistes *stars* (17 % contre 0,1 % du catalogue) et de popularité moyenne (31 % pour 0,9 % du catalogue pour la higher-mid-tail, 35 % pour 9 % du catalogue pour la lower-mid-tail). Les artistes de niche ne représentent, eux, que 17 % de l'ensemble des artistes écoutés par un utilisateur. Ces écarts sont encore plus importants

Document téléchargé depuis www.cairn.info - Orange Labs - - 161.106.4.5 - 09/04/2019 10h55. © La Découverte

si l'on considère la fréquence d'écoute plutôt que le panier. Les artistes stars représentent alors 29 % des écoutes des auditeurs ; un utilisateur sur quatre n'écoute quasiment que des artistes stars ou de la higher-mid-tail – plus de 80 % des écoutes – alors qu'il est rare que des utilisateurs en écoutent peu – seuls 4 % des usagers consacrent moins de 20 % des écoutes aux 1 % d'artistes les plus écoutés. Au contraire, il est plus fréquent de ne consommer pas ou peu d'artistes de niche (53 % des utilisateurs consacrent moins de 5 % de leurs écoutes aux 90 % d'artistes les moins écoutés), mais il est rare d'en écouter dans plus de la moitié de ses écoutes (2% des utilisateurs).

Autrement dit, tout le monde écoute les stars (et c'est d'ailleurs ce qui les définit). La différence entre les utilisateurs intensifs et les autres réside dans le fait que les stars dominent plus le panier des utilisateurs dont la consommation musicale est faible (inférieure au premier quartile) alors que ceux dont la consommation est forte (supérieure au troisième quartile) font une part un peu plus importante aux artistes de la lower-mid-tail. Ces écarts, cependant, concernent surtout le panier et pas la fréquence moyenne d'écoute : les utilisateurs intensifs ont des pratiques d'écoute ponctuellement plus diversifiées, mais réécoutent les mêmes artistes que les autres.

Figure 1. Consommation et fréquence d'écoute d'artistes stars, mid-tail et tail en fonction de l'intensité d'usage



Champ : toutes les écoutes.

Source : auteurs.

Document téléchargé depuis www.cairn.info - Orange Labs - - 161.106.4.5 - 09/04/2019 10h55. © La Découverte

LES USAGES DU GUIDAGE

Dans quelle mesure les auditeurs se laissent-ils guider dans leur écoute, et comment ? Sur quelles formes de recommandations s'appuient-ils ? L'originalité de nos données est de pouvoir observer l'usage des différents dispositifs offerts par la plateforme à travers une variable de *contexte*, qui représente la page à partir de laquelle le titre a été écouté. Si l'écoute d'un titre se fait sur la page d'une playlist éditoriale, il est raisonnable de penser que l'utilisateur suit les recommandations de cette playlist fournie par la plateforme ; si au contraire l'écoute a lieu depuis une page d'album mis en favori par l'utilisateur, on peut estimer qu'il s'agit d'une écoute volontaire d'un artiste déjà connu de l'auditeur, de sa discothèque virtuelle personnelle, à la manière d'un disque posé sur la platine. Pour pouvoir interpréter ainsi les contextes, nous avons créé différentes catégories qui agrègent les dispositifs proches ou les modalités redondantes. Nous identifions ainsi plusieurs grands types de recommandations. En premier lieu figurent les recommandations algorithmiques personnalisées, au centre des analyses sociologiques évoquées plus haut, qui s'appuient sur les écoutes passées des auditrices pour calculer leurs préférences et suggérer des titres adéquats. Parmi elles, nous pouvons distinguer les recommandations algorithmiques ponctuelles, qui peuvent porter sur un artiste (vous avez aimé X, vous devriez aimer Y), un album, une playlist, de celles qui sont composées en un flux algorithmique, qui enchaîne les titres recommandés en l'absence d'interruption de l'auditeur : flux continu de recommandations personnalisées, « radio de l'artiste » qui enchaîne des titres censément proches de ceux de l'artiste. À côté de ces outillages calculatoires, la plateforme fournit également des formes de guidage plus classiques dans les univers de la consommation culturelle : des classements (les artistes les plus écoutés, les morceaux les plus streamés cette semaine) et des sélections éditoriales (la playlist élaborée par tel artiste, tel DJ, tel partenaire commercial). Enfin, la plateforme fournit des outils de navigations contextuels permettant aux utilisatrices de rechercher un artiste dans le moteur de recherche, de naviguer une discographie, etc. Une dernière forme de recommandation, très peu utilisée, repose sur les liens explicites construits entre les utilisateurs « amis ». Bien sûr, le classement des outils au sein de ces catégories peut toujours être discuté ; par exemple, certaines playlists éditoriales font l'objet d'une recommandation algorithmique (« vous devriez aimer cette playlist concoctée par… »), ce qui combine deux formes distinctes de guidage.

Document téléchargé depuis www.cairn.info - Orange Labs - - 161.106.4.5 - 09/04/2019 10h55. © La Découverte

Une limite à l'utilisation de la variable *contexte* est néanmoins que la moitié des écoutes (52 %) ne peut pas être reliée à un dispositif particulier : que celui-ci ne soit pas identifié à cause d'un problème technique (« inconnu ») ou qu'il soit caché derrière la navigation de l'utilisateur dont nous n'observons que le résultat final (« indéterminé »), c'est-à-dire l'écoute d'un titre à partir de la page d'un album ou d'une playlist sans que nous puissions déterminer par quel chemin l'utilisateur est parvenu sur cette page (par une recommandation de la plateforme, une exploration, ou une navigation « au hasard »). Nous travaillerons donc, dans la suite, sur les 50 % d'écoutes qualifiées, pour lesquelles la variable est renseignée.

Écoutes autonomes et hétéronomes

Le premier résultat est que, au sein des écoutes qualifiées (8,8 millions), l'écoute « de stock » – par laquelle l'auditeur va chercher quelque chose qu'il connaît et a stocké – est nettement dominante : elle représente 5,1 millions d'écoutes, soit 59 % des écoutes qualifiées. Il s'agit pour l'essentiel de l'écoute d'artistes placés en favoris ; plus marginalement, de playlists personnelles et d'albums archivés. Symétriquement, les écoutes « guidées » d'une manière ou d'une autre représentent 41 % des écoutes qualifiées, soit 3,7 millions. Notons qu'elles ne concernent pas pour autant nécessairement des artistes inconnus de l'usager : les recommandations qui partent des goûts établis peuvent y ramener.

Pour autant, l'ensemble des écoutes guidées ne correspondent pas à la figure d'un consommateur hétéronome pris en charge par les algorithmes. Elles se répartissent de la façon décrite dans le tableau 4.

Au sein des dispositifs de guidage, une partie renforce plutôt la figure d'un utilisateur autonome, explorant activement et consciemment le catalogue : c'est le cas du moteur de recherche et des outils contextuels tels que la discographie des artistes (qui permet de naviguer l'ensemble de l'œuvre d'un artiste) ou le top des artistes (qui propose les titres les plus écoutés de cet artiste). Pour ces trois dispositifs, il est raisonnable de penser que l'utilisateur les mobilise volontairement dans l'intention d'explorer un univers musical. Ces dispositifs de l'exploration autonome comptent pour 1,5 million d'écoutes, soit 17 % des écoutes qualifiées. Ce sont donc, avec les écoutes de stock, 76 % des écoutes qui relèvent de la figure de l'auditeur autonome.

Document téléchargé depuis www.cairn.info - Orange Labs - - 161.106.4.5 - 09/04/2019 10h55. © La Découverte

Tableau 4. Distribution des écoutes qualifiées entre les différents dispositifs

Dispositifs	Nb. écoutes	Fréquence	Fréquence hors stock	Nb. usagers distincts	Écoutes moyennes par usager
Stock	5 172 732	58,70 %	N/A	3 535	1 463
Exploration autonome					
Discographie de l'artiste	502 611	5,70 %	13,80 %	1 908	263
Top de l'artiste	536 030	6,10 %	14,70 %	3 526	152
Moteur de recherche	418 133	4,70 %	11,50 %	3 030	138
Recommandations traditionnelles					
Radio éditoriale	656 651	7,40 %	18 %	1 970	333
Reco. éditoriales	9 141	0,10 %	0,30 %	466	20
Classements	270 641	3,10 %	7,40 %	1 693	160
Nouveautés	16 670	0,20 %	0,50 %	440	38
Recommandations personnalisées : vous devriez aimer ce/cette…					
Playlist	541 612	6,10 %	14,90 %	2 442	222
Album	264 872	3 %	7,30 %	2 691	98
Titre	49 904	0,60 %	1,40 %	2 655	19
Radio algorithmique	24 694	0,30 %	0,70 %	1 192	21
Flux algorithmiques					
Radio éditoriale « feed »	128 443	1,50 %	3,50 %	1 469	87
Radio « flow »	107 952	1,20 %	3 %	763	142
Radio algorithmique	97 194	1,10 %	2,70 %	1 031	94
Autres					
Player (par défaut)	16 647	0,20 %	0,50 %	1 228	14
Recommandations sociales	1 154	0,01 %	0,02 %	190	6
Total	8 815 081	100 %	100 %		

Champs : écoutes qualifiées (variable dispositif renseignée).

Source : auteurs.

Document téléchargé depuis www.cairn.info - Orange Labs - - 161.106.4.5 - 09/04/2019 10h55. © La Découverte

Restent trois formes d'écoutes guidées où l'agentivité des dispositifs est plus marquée. La première est la forme numérique des intermédiaires traditionnels, que nous pouvons qualifier de recommandation éditoriale. Elle représente environ une écoute sur dix. La deuxième relève de la recommandation proprement dite, de la suggestion : il s'agit des propositions algorithmiquement construites de titres, d'artistes, de playlists, qui jouent un rôle dans 8 % des écoutes. Selon l'interprétation que l'on fait du clic de l'auditeur sur la recommandation, ces dispositifs peuvent aussi bien être lus comme des formes de recommandation pertinentes suivies volontairement par un consommateur averti, que comme des outils de prise en charge d'un consommateur peu compétent. La dernière regroupe les écoutes de flux, radios d'un artiste et radios algorithmiques : dans cette configuration, l'essentiel de l'agentivité revient à la plateforme, qui sélectionne les titres et organise leur enchaînement, sans aucune intervention de l'utilisateur. L'écoute guidée représente ainsi un peu moins d'un quart des écoutes, dont une moitié d'écoute de flux.

Tableau 5. Distribution des durées d'écoute par dispositifs

	Écoutes de moins de 30 secondes	**Écoutes de plus de 30 secondes arrêtées avant la fin du morceau**	**Écoutes jusqu'à la fin du morceau**	***Total***
Non guidées	25,80 %	17 %	57,20 %	*100 %*
Guidées	33,80 %	11,70 %	54,50 %	*100 %*

Champ : écoutes qualifiées (variable dispositif renseignée).

Source : auteurs.

Nous avons déjà noté que 56 % des écoutes seulement se poursuivaient jusqu'à la fin du morceau. Ce chiffre décroît légèrement pour les écoutes guidées, qui ne sont que 54,5 % à être écoutés en entier (tableau 5). Cet écart est cependant faible et le guidage ne peut être tenu pour responsable de l'importance des écoutes partielles, qui est grande même pour les musiques activement choisies. Ce constat peut cacher deux phénomènes contradictoires, qui s'annulent : d'une part, les écoutes guidées sont plus susceptibles de guider vers des musiques moins appréciées que le stock, et donc plus souvent interrompues ; mais d'autre part, ces écoutes se font plus souvent en contextes de multiactivité, durant lesquels la musique peut n'être qu'un arrière-fond : en somme, on choisit une playlist justement pour ne pas avoir à décider

Document téléchargé depuis www.cairn.info - Orange Labs - - 161.106.4.5 - 09/04/2019 10h55. © La Découverte

consciemment de l'écoute. En l'absence de données plus précises sur les contextes d'écoute, nous n'avons pas de moyen de mesurer l'importance de ces deux hypothèses.

Guidage et découverte

Les différentes formes d'exploration et de guidage permettent-elles de renforcer la diversité de la consommation (ce qui est une de leurs promesses), ou conduisent-elles au contraire à renforcer l'uniformité de l'écoute musicale ? Nos données ne permettent pas de répondre clairement à cette question : notamment, la durée d'observation de cinq mois n'est pas suffisante pour identifier avec certitude des découvertes, et nous n'avons aucune information sur les autres formes d'écoute et d'information musicale mobilisées par les utilisateurs en dehors de la plateforme. À titre exploratoire, nous avons cependant repéré, au sein des écoutes guidées, celles qui relevaient d'une « première écoute » : la première occurrence des titres non écoutés le premier mois et non placés en favori avant la période d'observation. On obtient ainsi un ensemble de 639 319 écoutes, que l'on peut considérer sinon comme des découvertes, en tout cas comme des morceaux amenés ou ramenés à la conscience de l'utilisateur via une modalité d'exploration sur le site. Nous pouvons voir si les morceaux ainsi (re)-découverts via les dispositifs sont plutôt le fait d'artistes stars, mid-tail ou longue traîne (tableau 6).

Le premier constat est que, en dehors des dispositifs de classements du type « les plus écoutés », la plupart des systèmes de recommandation orientent tendanciellement les auditeurs des titres de la mid-tail et de la longue traîne ; les premières écoutes en général sont moins concentrées sur les artistes les plus écoutés (seules 10,4 % de ces écoutes concernent des artistes stars, contre 27,2 % du total des écoutes qualifiées).

Mis à part les classements, seules les recommandations éditoriales de la plateforme et la chanson présentée à l'usager au lancement du service orientent plus que la moyenne vers les artistes star ou de la higher mid-tail. L'orientation vers la longue traîne, cependant, varie beaucoup d'un dispositif à l'autre. Elle est la plus forte avec les recommandations algorithmiques, puis dans un second temps avec l'exploration autonome de la discographie des artistes et des nouvelles sorties. Contrairement à ce que l'on pourrait penser, dans les données agrégées, les dispositifs d'exploration autonome ne sont pas les plus « pointus » musicalement, même s'ils proposent une ouverture plus large que

Document téléchargé depuis www.cairn.info - Orange Labs - - 161.106.4.5 - 09/04/2019 10h55. © La Découverte

Tableau 6. Dispositifs d'exploration et longue traîne

Dispositifs	Effectif	Fréquence	Long tail	Lower mid-tail	Higher mid-tail	Star	Total
Exploration autonome							
Discographie de l'artiste	20 469	3,20 %	***50,60 %***	33,30 %	12,30 %	3,80 %	100 %
Top de l'artiste	78 874	12,30 %	23,60 %	29,30 %	30,60 %	*16,50 %*	100 %
Moteur de recherche	80 163	12,50 %	***42,10 %***	30,30 %	20,60 %	7,00 %	100 %
Recommandations traditionnelles							
Radio éditoriale	53 301	8,30 %	30,90 %	42,80 %	19,90 %	6,40 %	100 %
Reco. éditoriales	3 707	0,60 %	15,20 %	40,60 %	17,50 %	***26,70 %***	100 %
Classements	39 409	6,20 %	**2,10 %**	**8,20 %**	14,30 %	***75,30 %***	100 %
Nouvelles sorties	6 110	1,00 %	***42,10 %***	37,10 %	16,20 %	4,50 %	100 %
Recommandations personnalisées : vous devriez aimer ce/cette…							
Playlist	190 730	29,80 %	29,90 %	36,20 %	23,70 %	10,30 %	100 %
Album	40 842	6,40 %	*33,60 %*	46,40 %	15,20 %	4,80 %	100 %
Titre	18 086	2,80 %	*33,90 %*	39,10 %	18,80 %	8,20 %	100 %
Radio algorithmique	15 358	2,40 %	***75,60 %***	21,50 %	**2,60 %**	**0,40 %**	100 %
Flux algorithmiques							
Radio éditoriale « feed »	41 502	6,50 %	32,30 %	*53,10 %*	**11,40 %**	3,10 %	100 %
Radio « flow »	11 619	1,80 %	38,80 %	38,20 %	16,80 %	6,20 %	100 %
Radio algorithmique	30 636	6,20 %	***68,70 %***	24,30 %	**5,60 %**	**1,40 %**	100 %
Autres							
Recommandations sociales	400	0,10 %	*37,00 %*	27,00 %	19,00 %	*17,00 %*	100 %
Player (par défaut)	8 113	1,30 %	**0,10 %**	**14,00 %**	***54,10 %***	***31,80 %***	100 %
Ensemble premières écoutes	639 319	100 %	21,10 %	42,50 %	26,10 %	10,40 %	100 %

Champ : premières écoutes sur la période de titres non présents dans la collection, écoutes qualifiées (variable dispositif renseignée).

Source : auteurs.

Document téléchargé depuis www.cairn.info - Orange Labs - - 161.106.4.5 - 09/04/2019 10h55. © La Découverte

les recommandations éditoriales et les classements. De ce point de vue, le dispositif d'écoute tient en partie sa promesse d'orienter les utilisateurs vers les parties moins connues du catalogue.

La durabilité de ces « découvertes » est néanmoins très variable. On peut mesurer cette durabilité au nombre de réécoutes des artistes ainsi découverts. Or ce sont les artistes découverts via leur page qui sont le plus susceptibles d'être réécoutés (18 réécoutes en moyenne sur la période pour les titres découverts sur la page de discographie de l'artiste, 15 réécoutes pour les titres découverts dans le top de l'artiste, contre 8 pour l'ensemble des dispositifs), puis ceux découverts via les recommandations éditoriales (14 réécoutes en moyenne) ou via des recommandations du réseau de l'utilisateur (13 réécoutes en moyenne). À l'inverse, les titres découverts sur la « smart radio » ne donnent lieu qu'à 4 réécoutes moyennes sur la période. En somme, les formes d'exploration autonome[2] et les recommandations humaines, par opposition aux recommandations algorithmiques, semblent produire des relations plus durables avec les musiques découvertes. On note enfin que les recommandations humaines sont légèrement plus durables que l'exploration autonome pour les stars, mais le sont nettement moins pour les artistes de la mid-tail, où l'exploration est plus importante.

Des genres de guidage ?

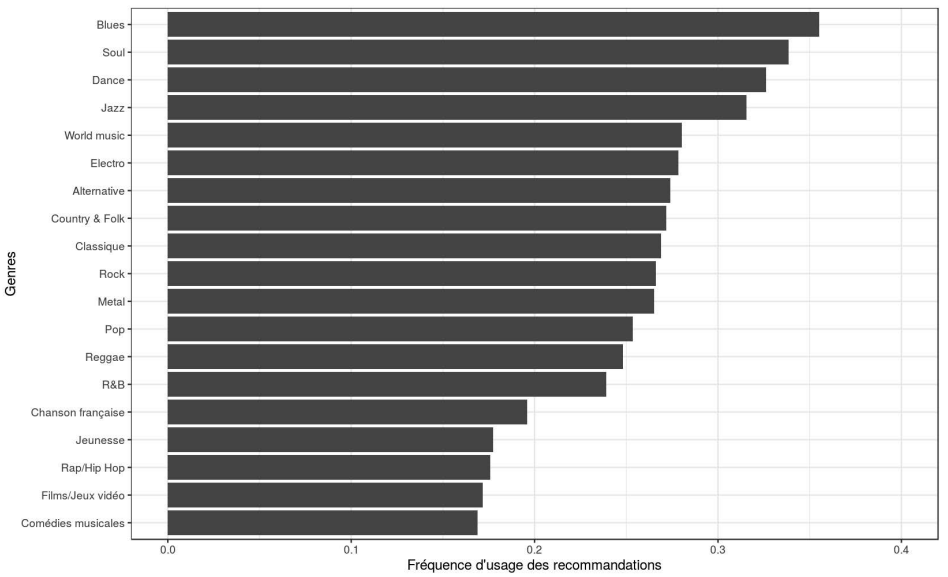
L'usage des recommandations est inégalement distribué par genres musicaux (figure 2) ; certains genres musicaux, comme le blues, le jazz ou la dance, semblent mieux se prêter à une écoute guidée. Ces liens peuvent être interprétés de deux manières. Une première hypothèse est que le guidage est biaisé en faveur de certains genres musicaux, c'est-à-dire qu'il orienterait préférentiellement les usagers vers certains genres. Cependant, une analyse de séquences d'écoutes successives permet de constater que, dans les séquences guidées comme dans les séquences libres, l'écoute tend à demeurer au sein d'un même genre musical. La probabilité de l'inertie – le genre du morceau écouté à t+1 est le même que le genre du morceau écouté à t – est toujours supérieure, et souvent largement supérieure, à la probabilité de changer de genre. Par ailleurs, lorsqu'il y a un changement, celui-ci se fait en priorité vers les genres

2. Il faut toutefois préciser que les dispositifs page de l'artiste sont d'interprétation encore plus difficile que les autres dispositifs, dans la mesure où l'usager a pu arriver sur cette page de nombreuses manières différentes.

Document téléchargé depuis www.cairn.info - Orange Labs - - 161.106.4.5 - 09/04/2019 10h55. © La Découverte

les plus fortement représentés dans la base, sans écart important entre les deux distributions. Or on trouve des genres très populaires (pop, rap/hip-hop) et des genres peu représentés (comédies musicales, jazz) à tous les degrés de la distribution des usages de recommandation. Il paraît donc raisonnable d'écarter cette hypothèse.

Figure 2. Fréquence d'usage des recommandations par genre musical



Champ : toutes les écoutes.

Source : auteurs.

Reste une explication : l'usage des recommandations dépend des contextes d'écoute, auxquels les genres musicaux sont liés. Si l'on suit cette hypothèse, c'est parce que certains genres musicaux se prêtent plus à la recommandation que d'autres que l'on constate des écarts entre eux. Dans la figure 2, on remarque ainsi deux groupes se distinguant clairement aux extrêmes de la distribution. D'une part, la musique pour enfant, le rap, les bandes originales de films et de jeux vidéo, et les comédies musicales font rarement l'objet d'écoutes recommandées ; d'autre part, le blues, la soul, la dance, et le jazz sont des genres où les écoutes recommandées sont importantes. Dans le premier cas, on constate qu'il s'agit soit de genres hétérogènes (comédies

Document téléchargé depuis www.cairn.info - Orange Labs - - 161.106.4.5 - 09/04/2019 10h55. © La Découverte

musicales, bandes originales) dont la mise en playlist ne fait pas sens : un auditeur ne veut pas écouter « de la musique de film », mais « la musique de *La la land* ». Dans le second, on remarque des genres qui se prêtent particulièrement bien à l'équipement d'autres activités sociales que l'écoute, et qui par conséquent sont plus susceptibles que d'autres d'être mis en fond sonore : la dance pour faire la fête, le blues pour une ambiance calme, etc.

Il est particulièrement intéressant de constater la place du rap et du jazz dans cette distribution. Alors que le jazz est un genre légitime, et que la figure dominante de l'auditeur de jazz, y compris dans la littérature sociologique, est celle de l'amateur (Roueff, 2002), son appartenance au groupe des musiques souvent écoutées par recommandation suggère qu'il existe également un autre usage, méconnu, mais répandu, du jazz comme genre d'accompagnement. À l'inverse, le rap est un genre populaire dont on aurait pu penser qu'il fait donc l'objet d'écoutes « distraites », d'accompagnement, plus souvent que des genres légitimes. Or il n'en est rien : il s'agit du genre homogène (par opposition aux musiques de films, catégorie qui dénote la fonction plutôt que le contenu musical) qui fait l'objet du moins d'écoutes recommandées, avec, dans une moindre mesure, la chanson française. On peut donc penser que les pratiques des auditeurs de rap ne sont pas si éloignées de celles des amateurs que l'on réserve habituellement aux musiques légitimes : ils prêtent une grande importance au choix direct des musiques écoutées, qui sont moins souvent déléguées aux algorithmes et aux experts de la plateforme. Cette observation de spécificités de l'écoute selon les genres musicaux rejoint le constat fait par Louail et Barthelemy (2017) : à partir de l'analyse d'une base de données similaire, ils observent que la variété de l'écoute varie fortement selon les genres musicaux, le rap étant ainsi caractérisé par une dispersion plus faible des personnes sur les titres que le jazz ou le classique. Il existe des spécificités de l'écoute selon les genres musicaux, qui ne sont sans doute pas complètement réductibles aux différences entre leurs auditeurs.

DISCUSSION : DES PROFILS OU DES CONTEXTES ?

Cette première analyse des données permet d'établir un certain nombre de résultats. Par rapport à l'inquiétude quant à l'émergence d'un auditeur passif et hétéronome, nous montrons que, dans les usages réels, les écoutes autonomes dominent : elles représentent les trois quarts des écoutes. Autrement dit, si les plateformes de streaming dessinent en effet la figure d'un amateur

Document téléchargé depuis www.cairn.info - Orange Labs - - 161.106.4.5 - 09/04/2019 10h55. © La Découverte

complètement pris en charge par les machines, cette figure est loin d'être habitée par les utilisateurs réels, qui choisissent la majorité de leurs écoutes, et usent préférentiellement de dispositifs autonomes d'exploration tels que les moteurs de recherche et les discographies d'artistes. Bien sûr, ces constats restent tributaires de la plateforme étudiée (toutes n'ont pas les mêmes outils de recommandation) et de la date de leur extraction (depuis 2014, les plateformes comme les usages ont évolué). De nouvelles enquêtes permettront sans doute d'affiner et nuancer ce constat, mais nos résultats attirent *a minima* l'attention sur la persistance importante d'une figure de l'auditrice autonome.

Le motif de l'écoute guidée représente un quart des usages mesurés, répartis équitablement entre écoute de flux et suivi ponctuel de recommandations algorithmiques. Dans tous ces cas, il y a effectivement délégation du choix à un dispositif technique : cette délégation est proche de celle du choix d'une station de radio dans le cas du flux, plus ponctuelle dans le cas d'un suivi de lien de recommandation. Ces outils semblent tenir leurs promesses d'exploration de la richesse des catalogues et d'enrichissement de la diversité consommée des utilisateurs, en les guidant préférentiellement vers des titres de longue traîne ou de notoriété intermédiaire.

Il reste qu'il est difficile de qualifier systématiquement d'hétéronomes et de passives ces écoutes où l'agentivité algorithmique est plus marquée. Le suivi de lien de recommandation (« vous avez aimé tel artiste, vous devriez aimer tel autre ») peut aussi bien être vu comme une pratique de novice (comblant son incompétence par la délégation à l'algorithme) que comme une démarche d'amateur éclairé (utilisant ponctuellement un outil informatique pour compléter sa palette exploratoire), comme le montre une récente étude des gros utilisateurs de recommandation algorithmique qui met en évidence leur réflexivité (Karakayali, Kostem *et al*., 2018). Même dans le cas des flux, ceux-ci sont choisis, peuvent être interrompus, transformés par l'évaluation active des morceaux. Autrement dit, si le constat de l'usage de ces dispositifs rend crédible l'émergence d'une figure de l'écoute hétéronome, elle demande à être vérifiée et affinée par d'autres méthodes.

En particulier, une des questions qui restent ouvertes est de savoir si le recours à l'écoute guidée relève plutôt des dispositions ou des situations. Existe-t-il des individus plus ou moins autonomes dans leurs écoutes musicales, plus ou susceptibles de se laisser guider ? Ou, au contraire, tous les individus ont-ils ponctuellement recours au guidage en fonction du contexte attentionnel dans

Document téléchargé depuis www.cairn.info - Orange Labs - - 161.106.4.5 - 09/04/2019 10h55. © La Découverte

lequel ils se trouvent ? Pour traiter cette question, il faudrait disposer de plus d'informations que nous n'en avons ici sur les utilisateurs. Pour l'instant, la recherche d'associations entre les propriétés des individus et le recours au guidage algorithmique n'est guère convaincante ; la seule relation un peu tangible établit que les systèmes de guidage sont plutôt plus utilisés – de façon non exclusive – par les usagers intensifs.

Les travaux qualitatifs sont alors cruciaux pour comprendre dans quels contextes, avec quelle conscience, quelle compréhension, quelle forme d'attention, quel objectif, les utilisateurs acceptent de s'en remettre aux algorithmes. Le travail exploratoire de Melançon (2015) auprès de dix utilisateurs esquisse à la fois une variété des profils, entre les utilisateurs refusant de voir les recommandations (« je sais déjà ce que je veux écouter »), ceux qui les soupçonnent de relever de la publicité, et ceux qui acceptent certaines des suggestions. Il souligne aussi l'importance des contextes dans le recours aux recommandations : certaines sont acceptables dans un moment d'exploration, d'autres dans un contexte d'écoute distraite, pour être immédiatement débranchée si l'attention se refocalise. Gilliotte (2019) observe de même que les amateurs de musique mettent en œuvre différents régimes d'engagement attentionnel selon les moments et les contextes ; certains identifient ainsi explicitement des moments d'exploration, d'écoute distraite, d'approfondissement, tandis que pour d'autres l'idée même de « découverte » est hors de propos. Hagen (2015) décrit, dans un même registre, des manières hétérogènes de créer et faire circuler les playlists, montrant qu'un même artefact technique est le support de pratiques variées. Ces recherches convergent pour montrer que les modalités d'écoute plus ou moins guidées vont être mobilisées en fonction des situations, des régimes d'engagement des personnes et de leurs compétences. Comme les résultats de notre enquête quantitative, ils nous éloignent dans tous les cas de la figure d'un utilisateur gouverné par les algorithmes, nous immunisent contre la croyance au premier degré dans les discours techno-déterministes tenus par les entrepreneurs des plateformes comme par leurs critiques ; ils nous orientent au contraire vers la compréhension des cultures algorithmiques émergentes, au sein desquelles les auditeurs, selon leurs compétences et selon les contextes, vont refuser ou brancher ponctuellement les algorithmes, sans pour autant abandonner le contrôle de leur expérience musicale.

Document téléchargé depuis www.cairn.info - Orange Labs - - 161.106.4.5 - 09/04/2019 10h55. © La Découverte

RÉFÉRENCES

ASUR S., HUBERMAN B. A. (2010), « Predicting the Future with Social Media », in *Proceedings of the 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, vol. 1. WI-IAT '10. Washington, DC, USA: IEEE Computer Society, pp. 492-499.

BASTARD I., BOURREAU M., MAILLARD S., MOREAU F. (2012), « De la visibilité à l'attention : les musiciens sur Internet », *Réseaux*, n° 175, pp. 19-42.

BEUSCART J.-S., MELLET K. (2012), *Promouvoir les œuvres culturelles. Usages et efficacité de la publicité dans les filières culturelles*. Ministère de la Culture – DEPS. https://www.cairn.info/promouvoir-les-oeuvres-culturelles--9782111281493.htm, consulté le 1er janvier 2019.

BEUSCART J.-S., MELLET K. (2016), « Shaping consumers' online voices. Algorithmic apparatus or evaluation culture? », in Robert Seyfert, Jonathan Roberge (eds.), *Algorithmic cultures. Essays on meaning, performance and new technologies*, London, Routledge, pp. 97-117.

BOURREAU M., MOREAU F., SENELLART P. (2011), « La diversité culturelle dans l'industrie de la musique enregistrée en France (2003-2008) », *Culture Études*, n° 5.

BRYNJOLFSSON E., YU J. H., SIMESTER D. (2011), « Goodbye Pareto principe, hello long tail. The effect of search costs on the concentration of product sales », *Management Science*, vol. 57, n° 8, pp. 1373-1386.

CELMA O., CANO P. (2008), « From hits to niches? Or how popular artists can bias music recommendation and discovery », in *ACM*, Las Vegas. http://mtg.upf.edu/files/publications/Celma-ACM-Netflix-KDD2008.pdf, consulté le 14 janvier 2019.

CHEVALIER J. A., MAYZLIN D. (2006), « The effect of word of mouth on sales. Online book reviews », *Journal of Marketing Research*, n° 43, pp. 345-54.

DONNAT O. (2009), *Les pratiques culturelles des Français à l'ère numérique. Enquête 2008*, Paris, La Découverte.

DONNAT O. (2018), « Évolution de la diversité consommée sur le marché de la musique enregistrée, 2007-2016 », *Culture Études*, n° 4.

ERIKSSON M., FLEISCHER R., JOHANSSON A., SNICKARS P., VONDERAU P. 2019. *Spotify teardown. Inside the black box of streaming music*. Cambridge (Ma), The MIT Press.

Document téléchargé depuis www.cairn.info - Orange Labs - - 161.106.4.5 - 09/04/2019 10h55. © La Découverte

FLEDER D., HOSANAGAR K. (2009), « Blockbuster Culture's Next Rise or Fall: The Impact of Recommender Systems on Sales Diversity », *Management Science*, vol. 55, n° 5, p. 697-712.

GALLOWAY A. R. (2012), *The interface effect*, Cambridge, Polity Press.

GILLESPIE T. (2016), « #trendingistrending: when algorithms become culture », in Jonathan Roberge, Robert Seyfert, *Algorithmic Cultures: Essays on Meaning, Performance and New Technologies*, London, Routledge, pp. 52-75.

GILLIOTTE Q. (2019), « L'expérience des biens culturels en régime numérique », thèse de doctorat, Paris, Télécom ParisTech.

HAGEN A. N. (2015), « The playlist experience. Personal playlists in music streaming services », *Popular Music and Society*, vol. 38, n° 5, pp. 625-645.

HAGEN A. N. (2016), « The metaphors we stream by. Making sense of music streaming », *First Monday*, vol. 21, n° 3.

HANRAHAN N. W. (2016), « La découverte musicale en régime numérique », in P. Le Guern (dir.), *Où va la musique ?*, Paris, Presses des mines.

HENNION A., MAISONNEUVE S., GOMART E. (2000), *Figures de l'amateur. Formes, objets, pratiques de l'amour de la musique aujourd'hui*, Paris, La Documentation Française.

KARAKAYALI N., KOSTEM B., GALIP I. (2018), « Recommendation systems as technologies of the self. Algorithmic control and the formation of music taste ». *Theory, Culture and Society*, vol. 35, n° 2, pp. 3-24.

KARPIK L. (2007), *L'économie des biens singuliers*, Paris, Gallimard.

KONSTAN J. A., RIEDL J. (2012), « Recommender systems. From algorithms to user experience », *User Modeling and User-Adapted Interaction*, vol. 22, n° 1-2, pp. 101-123.

LASH S. (2007), « Power after hegemony. Cultural studies in mutation? », *Theory, Culture and Society*, vol. 24, n° 3, pp. 55-78.

LAZARSFELD P. L., KATZ E. (2008), *Influence personnelle. Ce que les gens font des médias*, Paris, Armand Colin.

LOUAIL T., BARTHELEMY M. (2017), « Headphones on the wire ». *arXiv:1704.05815 [physics]*, avril. http://arxiv.org/abs/1704.05815, consulté le 14 janvier 2019.

MAILLARD S. (2015), *Information et consommation de biens culturels à lre numérique. Contributions empiriques sur le marché de la musique*, thèse de doctorat, Paris, Télécom ParisTech.

MELANÇON L. (2015), « La Nubémorphose des pratiques de consommation musicale. Le cas des plateformes de diffusion musicale en continu ». Rapport de recherche.

Document téléchargé depuis www.cairn.info - Orange Labs - - 161.106.4.5 - 09/04/2019 10h55. © La Découverte

SENSE, Orange Labs. https://lmelancon.net/2016/03/19/la-nubemorphose/, consulté le 14 janvier 2019.

MORRIS J. W. (2015), « Curation by code. Infomediaries and the data mining of taste », *European Journal of Cultural Studies*, vol. 18, n° 4-5, pp. 446-463.

OBSERVATOIRE DE LA MUSIQUE (2015), « Indicateurs de la diversité musicale dans le paysage radiophonique. Rapport 2015 ». https://philharmoniedeparis.fr/fr/ressources-numeriques/observatoire-de-la-musique/etudes/la-diversite-musicale-dans-le-paysage, consulté le 14 janvier 2019.

PASQUALE F. (2016), *Black box society. The Secret Algorithms That Control Money and Information*, Cambridge, Harvard University Press.

PRAG J., CASAVANT J. (1994), « An Empirical Study of the Determinants of Revenues and Marketing Expenditures in the Motion Picture Industry ». *Journal of Cultural Economics*, vol. 18, n° 3, pp. 217-235.

PREY R. (2018), « Nothing personal: algorithmic individuation on music streaming platforms ». *Media, Culture & Society*, vol. 40, n° 7, pp. 1089-1100.

ROUEFF O. (2002), « Faire le jazz : la coproduction de l'expérience esthétique dans un jazz-club », *Revue de musicologie*, vol. 88, n° 1, pp. 67-93.

SALGANIK M., DODDS P. S., WATTS D. J. (2006), « Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market », *Science*, n° 311 (5762), pp. 854-856.

SEAVER N. (2019 à paraître), « Captivating Algorithms. Recommender Systems as Trap », *Journal of Material Culture*, http://nickseaver.net/s/seaver-captivating-algorithms.pdf, consulté le 14 janvier 2019.

SENECAL S., NANTEL J. (2004), « The influence of online product recommendations on consumers' online choices », *Journal of Retailing*, vol. 80, n° 2, pp. 159-169.

SEYFERT R., ROBERGE J. (eds.) (2016), *Algorithmic cultures. Essays on meaning, performance and new technologies*, London, Routledge.

SHARMA A., HOFMAN J. M., DUNCAN J. W. (2015), « Estimating the Causal Impact of Recommendation Systems from Observational Data », *Proceedings of the Sixteenth ACM Conference on Economics and Computation - EC '15*, pp. 453-470.

SNEP (2018), « Économie de la production musicale – Édition 2018 », http://www.snepmusique.com/actualites-du-snep/economie-de-la-production-musicale-edition-2018/

STERNE J. (2006), « The mp3 as a cultural artifact », *New Media & Society*, vol. 8, n° 5, pp. 825-842.

STIRLING A. (2007), « A general framework for analysing diversity in science, technology and society », *Journal of the Royal Society Interface*, vol. 4, n° 15, pp. 707-719.

Document téléchargé depuis www.cairn.info - Orange Labs - - 161.106.4.5 - 09/04/2019 10h55. © La Découverte