

Les podcasts natifs d'information en France : méthodologie pour un recensement de l'offre.

Article inédit, mis en ligne le 25 septembre 2023.

Marie-Éva Lesaunier

Marie-Eva Lesaunier est maîtresse de conférences en Sciences de l'Information et de la Communication à l'Université Sorbonne Paris Nord. Ses recherches portent sur la production et la circulation des biens médiatiques et culturels, notamment le documentaire, à travers une thèse soutenue en 2021, et le podcast dans le cadre du programme de recherche Obcast (CARISM).

marieevalesaunier@gmail.com

Plan de l'article

Introduction

Des initiatives de recensement préexistantes

Objectif et périmètre de la cartographie du podcast natif d'information de l'Observatoire

Le triple apport de la base de données

Conclusion

Notes

Références bibliographiques

RÉSUMÉ

Cet article présente la méthodologie employée pour le recensement des podcasts natifs d'information français dans le cadre de l'Observatoire du podcast (programme Obcast). Un tel recensement, dans un contexte de croissance rapide du secteur, contribue à une meilleure connaissance de l'offre, pour la recherche et pour le monde professionnel naissant, tant d'un point de vue quantitatif que qualitatif. Il constitue aussi un premier jalon, important pour les recherches en cours et à venir sur le secteur du podcast, puisqu'il permet de faire émerger des questions, des hypothèses et des terrains de recherche. Ainsi, nous montrons que les stratégies de valorisation des podcasts sur le site de leur producteur sont très hétérogènes et interrogeons le processus de normalisation technique et professionnelle à l'œuvre dans le secteur. Nous attirons enfin l'attention sur des corpus de podcasts revêtant un intérêt particulier du fait de leur thématique (coulisses du journalisme, notamment).

Mots clés

Podcast, audiovisuel, méthodologie, recensement de l'offre

FRENCH INFORMATIONAL PODCASTS. METHODOLOGY FOR A SUPPLY INVENTORY

Abstract

This paper introduces the methodology we used to achieve an inventory of the French informational podcasts. This work is part of the Obcast research program (Observing podcast). Such an inventory proves itself useful to both the researchers and the emerging professional sector, as it allows a better apprehension of the fast-growing field of podcast, both in quantitative and qualitative perspectives. We then argue that it constitutes an important landmark for further researches on podcasts, as it calls attention to research questions, hypothesis and objects all at the same time. Thus we highlight the various ways of valorizing podcasts on the editors' websites, and focus on what seems like a technical and professional standardization of the field. Finally, we draw attention to specific sets of podcasts, particularly relevant due to their subject (especially "journalism behind the scenes").

Keywords

Podcasts, audiovisual contents, methodology, supply inventory

PODCASTS INFORMATIVOS FRANCESES. METODOLOGÍA PARA UN INVENTARIO DE LA OFERTA

Resumen

Este artículo presenta la metodología que hemos utilizado para realizar un inventario de los podcasts informativos franceses. Este trabajo forma parte del programa de investigación Obcast (Observar los podcasts). Sostenemos que un inventario de este tipo resulta útil tanto para los investigadores como para el sector profesional emergente, ya que permite conocer mejor el campo en rápido crecimiento de los podcasts, tanto desde el punto de vista cuantitativo como cualitativo. Este inventario constituye un hito importante para futuras investigaciones sobre podcasts, centrando la atención simultáneamente en las preguntas de investigación, las hipótesis y los objetos. Así, mostramos que las estrategias de promoción de los podcasts en el sitio web de su productor son muy heterogéneas, y cuestionamos el proceso de normalización técnica y profesional en curso en el sector. Por último, llamamos la atención sobre podcasts de especial interés por su temática (periodismo entre bastidores, en particular).

Palabras clave

Podcast, audiovisual, metodología, inventario de la oferta

INTRODUCTION

Cet article propose un retour descriptif et analytique sur la méthodologie employée pour réaliser une base de données des podcasts natifs. Il s'inscrit dans le cadre de l'Observatoire du podcast (Obcast), un projet de recherche financé en 2021 par la Direction Générale des Médias et des Industries Culturelles (DGMIC) et mené au sein du Centre d'Analyse et de Recherche Interdisciplinaire sur les Médias (Carism, Université Paris-Panthéon-Assas)^[1]. Ce programme a vocation à analyser la situation naissante du podcast français

à la fois dans une perspective de recherche académique et dans un souci de faire état de nos résultats concernant le secteur lui-même. En effet, le podcast a connu un fort essor ces dernières années, entre un nombre de nouveaux titres publiés chaque année dans le monde en croissance exponentielle depuis 2016 (*Les Échos*, 2021), et en France, des auditeurs mensuels passant de 10,9 millions en 2019 à 17,6 millions en 2022, où 36,1% des internautes français déclarent écouter au moins un podcast par mois (*Médiamétrie*, 2022). Les travaux de l'équipe Obcast s'attachent à identifier les logiques de production et les conditions de leur réception (approches sociologiques, socio-économiques et juridiques), ainsi que les contenus eux-mêmes, dans une perspective socio-sémiotique. Plusieurs enquêtes sont ainsi en cours ou achevées et seront prochainement publiées (enquête par entretien auprès des média traditionnels, analyse des discours d'escorte, notamment).

Dans ce contexte, la première tâche de l'Observatoire a consisté en un recensement de l'offre de podcasts natifs français, un secteur extrêmement dynamique et mouvant depuis environ six ans. Nous avons ainsi constitué une base de données de 1 345 titres de podcasts et de 76 051 épisodes. L'objectif de cet article est donc d'expliquer de quelle manière la construction d'une base de données des podcasts natifs d'information français offre l'opportunité de mener des recherches empiriques sur ces productions. Pour ce faire, nous procéderons d'abord à un tour d'horizon d'initiatives de recensements préexistantes, aussi bien dans le monde professionnel qu'universitaire, afin de pouvoir les mettre en perspective avec le recensement Obcast. Nous préciserons ensuite le périmètre choisi pour notre étude et reviendrons sur quelques éléments de définition et d'histoire du podcast natif. Nous déploierons ensuite ladite méthodologie qui est l'objet de cet article et terminerons par mettre en évidence les trois principaux apports de la base de données : obtenir des statistiques précises sur l'offre de podcasts natifs d'information, faire émerger des hypothèses de travail et construire des corpus.

DES INITIATIVES DE RECENSEMENT PRÉEXISTANTES

D'autres recensements ont été réalisés depuis 2019, dans un contexte de développement rapide du secteur du podcast. En effet, c'est à partir de 2017 que la croissance du nombre de podcasts produits s'accélère, et ce de façon exponentielle sur les trois années suivantes, passant de 43 titres lancés en 2017 à 241 en 2019 et 417 en 2021 (tableau n°1). Notons que parmi les sept initiatives présentées ci-après, quatre proviennent directement des acteurs du secteur du podcast, relevant ainsi d'une logique de construction du groupe professionnel (Demazière, Gadéa, 2009) et de désignation de ses frontières, au même titre, par exemple, que la création du premier Syndicat des producteurs audio indépendants (PIA) en 2020.

Date du premier épisode	2015	2016	2017	2018	2019	2020	2021
Nombre de titres	7	7	43	119	241	325	417

Tableau 1. Nombre de titres de podcasts lancés par année recensés dans la base de données Obcast, 2015-2021

La première cartographie du secteur ayant émergé en France est à l'initiative de *CosaVostra*, un studio de production de podcasts réalisant à la fois des créations propres et pour des marques, prenant alors la fonction de prestataire. Elle a pour objectif de rendre compte de l'environnement du podcast, « composé de personnes et d'organisations de tous genres, allant

économique) des différents acteurs sur le marché.

La deuxième initiative est une œuvre individuelle, celle de Laura Pironnet, *podcast production manager* au sein du studio *Cosavostra*. Sur la plateforme de blog *Medium*, elle publiait en juin 2019 un premier recensement des podcasts natifs produits par les éditeurs de presse traditionnels. Ce dernier a été mis à jour en mai 2020 et fait état de 211 titres de podcasts. Ce recensement a l'avantage de tendre à une certaine exhaustivité sur un secteur donné et à une date donnée, et, déjà, de mettre en évidence des tendances relatives à l'appropriation du média podcast par les éditeurs de presse écrite, depuis mises à jour par l'équipe d'Obcast (Lesaunier, Di Scullo, Mercier, 2022).

Une troisième initiative issue du secteur est celle de *Madame Black bow*, un studio de podcasts créé en juin 2019. Cette entreprise propose un « Répertoire des podcasts de marque » (Madamebb, 2019) qui, comme son nom l'indique, a vocation à répertorier un type particulier de podcasts : ceux produits par/pour des marques^[2]. Les 202 titres recensés sont ici classés par marque, puis par secteur.

Enfin, du côté des acteurs du secteur, la société d'hébergement de podcasts *Audiomeans* a lancé en mai 2021 « Podcast city », une carte au sens littéral, offrant une visualisation des principaux titres de podcasts existants^[3]. Cette dernière s'apparente à une contribution à la question de la catégorisation des podcasts en différentes thématiques, qui conditionne en partie leur accessibilité. La question de l'accessibilité des podcasts et de leur recommandation le plus souvent algorithmée a déjà fait l'objet de travaux (Morris, Patterson, 2015, Resler, 2017 ; Morris, 2021) et s'avère être source de tensions entre les acteurs du secteur en France. *Audiomeans* relate ainsi la genèse de leur cartographie : « Rapidement, une petite pousse germe : comment repousser les limites de cette ontologie officielle (puisque celle d'Apple est un standard de fait) ? Comment la rendre plus poreuse, moins algorithmique, plus visuelle, et pourquoi pas même artistique ? » (Audiomeans, 2021). Cette carte recense 494 titres de podcasts, organisés par thématique. Elle comporte une dimension interactive puisque l'utilisateur a la possibilité de filtrer podcasts natifs (n= 448) ou *replay* (n= 46), ou encore de n'afficher que les podcasts faisant l'objet de plus de 1 000 évaluations sur *Apple podcasts*.

Dans le monde universitaire et institutionnel, trois initiatives sont à relever. Tout d'abord, dans le cadre de l'extension du dépôt légal au web, l'INA collecte et documente les podcasts depuis 2020. L'Institut dispose ainsi d'une base de 4 609 podcasts, donnant accès aux fichiers audio et aux métadonnées liées (résumé, nombre d'épisodes, notamment). Par ailleurs, en 2020, l'entreprise *Spotify* a construit un jeu de données relatif au podcast nommé « *Spotify podcast data set* », constitué d'environ 100 000 épisodes de podcasts (le fichier audio, la retranscription textuelle de ce dernier, et les métadonnées associées à l'épisode) (Clifton *et alii.*, 2020). Les podcasts répertoriés sont en langue anglaise ou portugaise. Cette initiative a d'abord émergé dans le cadre d'un projet de recherche particulier, le *TREC podcast task*, avant d'avoir vocation à être partagée à une échelle plus large au monde académique... anglophone. Enfin, le *Department of Communication Arts and Libraries* de l'Université de Wisconsin-Madison a lancé en 2014 le projet *PodcastRE*. Dirigé par Jeremy Morris, il a pour objectif d'archiver le podcast, dont la présence sur le web est perçue comme éphémère. Les équipes ont collecté de façon automatisée et manuelle environ deux millions de fichiers audio de 16 000 titres de podcasts mis en ligne depuis 2007, d'une part, en « aspirant » chaque jour les flux RSS du classement des 100 podcasts les plus populaires sur *Apple podcast* aux États-Unis, en Grande-Bretagne, en France et en Australie, d'autre part en collaborant avec des chercheurs travaillant sur des corpus particuliers de podcasts indépendants, et sur des bases de données préexistantes portant sur des podcasts « marginalisés/moins visibles » (Hoyt *et alii.*, 2021, p. 3). Les chercheurs ont également créé un outil offrant la possibilité aux auteurs de podcasts de soumettre directement leur titre sur le site du projet. Ce site met à disposition un moteur de recherche en libre

accès, permettant de consulter la base de données, ainsi que des outils de visualisation, notamment concernant les temporalités des épisodes et les mots-clés qui leur sont associés, comme ils le résumant dans l'article de présentation du projet : « 1) *visualizing how topics and keywords trend over time*; 2) *visualizing networks of commonly associated keywords entered by podcasters*; and 3) *analyzing norms and common practices for the duration of podcasts (as a time-based media format, podcasting is unusual in that it is not bound by the programming schedules and technical limitations that provide strict parameters for most audiovisual forms)* » (Hoyt et alii., 2021, p. 3)^[4]. PodcastRE constitue ainsi un espace numérique pérenne d'archivage du podcast, y compris de langue française.

La base de données que nous présentons dans cet article possède ainsi des similitudes avec certaines des initiatives présentées ci-dessus. Néanmoins, par les objectifs qui lui sont affiliés, elle présente de nombreuses spécificités sur lesquelles nous reviendrons en fin d'article.

OBJECTIFS ET PÉRIMÈTRE DU RECENSEMENT DE L'OFFRE DE PODCASTS NATIFS D'INFORMATION DE L'OBSERVATOIRE

L'objectif premier du recensement de l'Obcast a été d'obtenir une vision relativement exhaustive de l'offre de podcasts natifs d'information et d'en identifier les principaux acteurs : qui fait du podcast ? Depuis quand ? Selon quels formats ? À propos de quelles thématiques ? Il s'est donc agi de constituer une base de données permettant, à terme, de saisir les caractéristiques du podcast natif d'information mais, avant, il convient de revenir sur quelques éléments de définition afin de préciser le périmètre du recensement.

Définition

Le terme « podcast » est formulé pour la première fois en 2004 par le journaliste Ben Hammersley dans un article sur le sujet, qui décrivait un « nouveau boom dans la radio amateur » (Berry, 2015, p. 3). Il est issu de la contraction entre les mots *iPod* (baladeur d'Apple) et *broadcast* (diffusion) (Cohen, 2019, p. 159), et s'est imposé pour désigner un fichier numérique audio pouvant être soit natif, c'est-à-dire n'ayant pas fait l'objet d'une diffusion linéaire sur une station de radio notamment, soit *replay*, désignant dans ce cas un programme audio numérique ayant fait l'objet d'une première diffusion, inscrite dans une grille de programmation. En 2022, les podcasts natifs représentaient 2% du marché de l'audio digital, quand la radio à la demande en représentait 4% (Mediamétrie, 2022). Ainsi, la définition du podcast est d'abord technique, comme l'explique le chercheur italien Tiziano Bonini dès 2015 : « *Technically, podcasting is a method for delivering audio – and/or video files via so-called RSS feed for download and later playback on various devices* (Markman/Sawyer 2014)^[5] » (Bonini, 2015, p. 21). Le flux RSS (*Real Simple Syndication*) est ainsi constitutif de la définition du podcast (Hoyt et alii., 2021).

De façon plus indigène, un podcast désigne également un programme audio ayant sa thématique et son identité propre, qui comporte un ou plusieurs fichiers (épisodes). Ainsi, le programme *Les couilles sur la table* (*Binge audio*, 2017) serait alors un podcast, comprenant plusieurs saisons et plusieurs épisodes. De même, en octobre 2022, le prix du « meilleur podcast de conversation » du Paris Podcast Festival a été remis à Guillaume Bonnet pour son programme *Ma dernière séance de psychanalyse*, et non pas pour l'un des 46 épisodes qui le composent. Au sens technique du terme, cependant, ces épisodes sont également des podcasts. Pour éviter de futures confusions, nous désignerons donc par « titre de podcast » les programmes numériques audio comportant un ou plusieurs épisodes et ayant leur identité propre, et « épisodes de podcast » les fichiers numériques audio, agglomérés sous

une même entité. Lorsque la distinction ne s'avère pas nécessaire pour la compréhension du propos, nous utiliserons simplement le terme de « podcast ». Notre observatoire et la base de données présentée ici s'intéressent donc aux podcasts natifs, c'est-à-dire aux programmes numériques audio spécifiquement produits pour une consommation délinéarisée, comportant un ou plusieurs épisodes et ayant leur identité propre.

Périmètre

Nous prenons ici en compte les podcasts natifs d'information, la notion d'information étant entendue dans son acception large, c'est-à-dire « la mise en forme d'un contenu porté à la connaissance d'un public » (Legavre, Rieffel, 2017, p. 55-56). En effet, le point de départ de notre base de données est centré autour des médias traditionnels de presse écrite, avec pour questionnement les modalités d'intégration du podcast aux médias d'information et le potentiel renouvellement des formats journalistiques. Le choix d'une acception large de la notion d'information est lié au fait que les médias s'étant lancés dans le podcast ne l'utilisent pas exclusivement pour traiter de l'actualité, mais aussi pour aborder des thèmes et des formats parfois à rebours d'un traitement strictement journalistique de l'information. C'est par exemple le cas de « S'aimer comme on se quitte », produit par *Le Monde* (2021), qui traite de « la vie des amoureux » en mettant en scène des chroniques incarnées par des comédiens. Le fait d'utiliser le podcast pour produire des contenus décalés par rapport aux formats et aux thèmes journalistiques traditionnels justifie précisément l'intérêt à accorder aux podcasts natifs français. En même temps, les podcasts de fiction sont quasi inexistantes parmi les médias traditionnels, nous éloignant ainsi des questions de départ. Afin de favoriser une analyse large, permettant de faire état de l'hypothétique diversité des podcasts, mais pertinente, en restant dans un questionnement relatif aux interactions entre podcasts et médias traditionnels, nous avons donc choisi de n'exclure de notre périmètre que les podcasts de fiction.

Le point de départ de notre base de données s'est donc centré sur les médias traditionnels de presse écrite. Pour ce faire, nous avons pris en compte les titres répertoriés à l'Alliance pour les Chiffres de la Presse et des Médias (ACPM) au 1^{er} mars 2022. Cela a permis d'avoir un recueil exhaustif des titres de presse quotidienne nationale (n=9), de presse quotidienne régionale (n=51) et de presse magazine (n=234) afin de savoir qui, parmi eux, produit des podcasts. Nous avons également pris en compte le classement ACPM des titres de podcasts les plus écoutés en février 2022 (n= 268) et recensé, pour un producteur figurant au classement, l'ensemble des podcasts natifs qu'il a produit par ailleurs. Cela a fait émerger, dans le recensement, les principaux acteurs du podcast natif en termes d'audience. De même, les productions originales des principaux studios de production de podcasts ont été systématiquement recensées (235 titres de podcast, 27 producteurs), ces derniers constituant l'un des acteurs centraux du déploiement du podcast natif. Nous avons également pris en compte les titres produits par les principaux *pure players*, stations de radios, chaînes de télévision privées et publiques, plateformes de diffusion. Nous avons enfin recensé une infime partie de la production amateur et indépendante, mais cette collecte est loin d'être exhaustive et a été effectuée à partir des titres soit les plus écoutés (figurant au classement ACPM), c'est-à-dire les plus connus par l'équipe de recherche (car très médiatisés, ou faisant l'objet d'un intérêt scientifique particulier, nous y reviendrons dans la dernière section de cet article). Il a enfin été nécessaire de borner temporellement notre étude, du fait de la fréquence importante à laquelle sont lancés de nouveaux titres de podcast. Nous prenons ainsi en compte les titres de podcasts produits entre le 1^{er} janvier 2002 (année du premier titre identifié) et le 30 juillet 2022. Ainsi, notre base de données est constituée de 1 345 titres de podcasts, pour 76 051 épisodes, réalisés par 397 producteurs.

Catégorie de producteurs	Nombre de producteurs	Titres de podcast produits	Nombre moyen de titres produits par producteur
Presse écrite	128	627	4,9
<i>dont Presse gratuite d'information (PGI)</i>	1	9	9
<i>dont Presse quotidienne nationale (PQN)</i>	9	59	6,6
<i>dont Presse quotidienne régionale (PQR)</i>	29	209	7,2
<i>dont Agence de presse</i>	1	7	7
<i>dont Pure player</i>	22	131	6
<i>dont Presse magazine</i>	66	212	3,2
INA	1	7	7
Radio	20	162	8,1
TV	12	127	10,6
Studio de production/ Label ^[6]	27	235	8,7
Individu/amateur	78	93	1,2
Plateforme de diffusion	2	94	47
TOTAL	397	1345	3,4

Tableau 1. Volume de production par type de producteur

Méthodologie du recensement

L'objectif de la base de données est non seulement d'avoir une vision claire de l'existant (quantité de podcasts, acteurs concernés), mais également de pouvoir générer des résultats statistiques au regard de différentes variables. Nous avons donc progressivement établi des catégories qui rendent compte des podcasts de la façon la plus exhaustive possible, afin de récolter des données susceptibles de répondre aux questions de recherche et d'en nourrir de futures. Ces catégories prennent en considération l'ensemble du processus de communication des podcasts (production, message, réception), et se sont progressivement portées au nombre de six : conditions de production ; identité du personnel créatif ; aspects éditoriaux ; espaces de la diffusion ; temporalités ; réception. Ces six catégories sont constituées de 41 items, décrits dans les tableaux suivants.

	Item	Description
1	Statut du producteur	Statut de l'acteur qui est à l'initiative du podcast : <i>Studio de production de podcast ; autoproduction ; groupe média privé ; groupe média public ; médias indépendants^[7] ; autres entreprises médias^[8].</i>
2	Type de producteur	Nature de l'acteur qui est à l'initiative du podcast : <i>Agence de presse ; association ; autoproduction ; entreprise ; INA ; Label ; Presse gratuite d'information (PGI) ; plateforme ; PQN ; PQR ; presse magazine ; pure player ; radio ; studio ; télécom ; TV</i>
3	Nom du producteur	Nom de l'acteur à l'initiative du podcast
4	Studio de production	L'entreprise a-t-elle fait appel à un studio de production (prestation technique et/ou créative) ?
5	Coproducteurs/ partenariats	Présence ou non de coproducteurs ou de partenaires dans la production du podcast

Tableau 2. Les 5 items de la catégorie « Conditions de production »

	Item	Description
6	Créateurs/auteurs/ équipe créatives	Nom des auteurs ou personnes associées au podcast, ainsi que leur occupation professionnelle
7	Catégorie professionnelle de l'auteur	Associer ladite occupation à une des catégories proposées : rédaction ; journaliste ; studio de podcasts ; artiste ; professeur/chercheur ; activiste/associatif ; étudiants ; professionnel du secteur dont traite le titre ; passionné/concerné ; Autres
8	Genre de l' <i>host</i> (animateur du podcast)	L'animateur du podcast se présente-t-il comme un homme, une femme, une personne non-binaire ? ^[9]
9	Nom de l' <i>host</i>	Sans objet

Tableau 3. Les 4 items de la catégorie « Identité du personnel créatif »

	Item	Description
10	Nom du titre	Titre du podcast
11	Collection	Le titre appartient-t-il à une collection ?
12	Résumé officiel	Résumé tel que présenté sur les plateformes d'écoute
13	Rubrique	Thématique générale (le domaine) abordée dans le podcast (n=17)
14	Mots-clefs	Items devant faciliter la constitution de corpus en précisant la rubrique précédemment explicitée (n=80)
15	Format	Type de format adopté par le podcast (n=9)
16	Étiquettes thématiques (tag) sur Spotify	Brefs syntagmes descriptifs du titre proposés par la plateforme
17	Étiquettes thématiques (tag) sur Apple podcast	Brefs syntagmes descriptifs du titre proposés par la plateforme

Tableau 4. Les 8 items de la catégorie « Aspects éditoriaux »

	Item	Description
18	Le titre est disponible sur le site du producteur	Sans objet
19	Le titre est disponible sur une plateforme de diffusion	Le titre est disponible une de ces plateformes de diffusion : <i>Apple podcast, Spotify, Deezer, Google podcast, Castbox, Podcast Addict, Youtube, Soundcloud, Stitcher, Amazon Music</i>
20	Le titre est disponible sur le site d'un partenaire	Sans objet
21	Le titre est disponible sur une autre plateforme d'écoute	Sans objet

Tableau 5. Les 4 items de la catégories « Espaces de la diffusion »^[10]

	Item	Description
22	Date de publication du premier épisode	Sans objet
23	Date de publication du dernier épisode au moment de la collecte	Sans objet
24	Nombre de mois d'écart entre les deux dates	Sans objet
25	Périodicité	Fréquence à laquelle sont publiés les nouveaux épisodes
26	Le titre est interrompu	De nouveaux épisodes sont-ils publiés ?
27	Nombre d'épisodes	Nombre d'épisodes disponibles au moment de la collecte
28	Nombre total d'épisodes pour le producteur concerné	Cumul du nombre d'épisodes de l'ensemble des titres produit par un même acteur
29	Durée moyenne des épisodes	Sans objet
30	La durée est variable	La durée varie-t-elle de plus de 50% d'un épisode à l'autre ?
31	Date de la collecte	Sans objet

Tableau 6. Les 10 items de la catégories « Temporalités »

	Item	Description
32	Données d'audience sur le site du producteur	Recueil des données d'écoute disponibles
33	Données d'audience sur <i>Apple podcast</i>	Recueil des données d'écoute disponibles (nombre d'évaluation et note)
34	Données d'audience sur <i>Castbox</i>	Recueil des données d'écoute disponibles (nombre d'abonnement au titre et nombre d'écoutes)
35	Données d'audience sur <i>Podcast Addict</i>	Recueil des données d'écoute disponibles (nombre d'abonnement au titre et nombre d'écoutes)
36	Données d'audience sur <i>Youtube</i>	Recueil des données d'écoute disponibles (nombre moyen de vues)
37	Données d'audience sur <i>Soundcloud</i>	Recueil des données d'écoute disponibles (nombre d'écoutes)
38	Nombre de téléchargements France, février 2022 (ACPM)	Sans objet
39	Nombre de téléchargements monde, février 2022 (ACPM)	Sans objet
40	Modalité d'accès	L'accès au titre est-il gratuit ou payant ?
41	Prix/récompenses	Le titre a-t-il obtenu des prix ou des récompenses ?

Tableau 7. Les 10 items de la catégorie « Réception »

La catégorie « Aspects éditoriaux » est construite autour de huit items (n°10 à 17). Si pour certains, l'intitulé même de l'item (ex. : « nom du titre ») ou une brève description suffisent, pour d'autres, quelques précisions doivent au contraire ici être apportées.

Formats (item n°15)

Tout d'abord, en nous appuyant sur les catégories journalistiques et médiatiques préexistantes (voir par exemple Lochard, 1996 ; Soulages, 2007), nous avons identifié huit formats endossés par les podcasts recensés : interview, documentaire/reportage, chronique, émission, récit fictionnalisant, portrait, débat, captation et retransmission. L'interview est le premier format représenté dans la base de données (26,6%). Dans ces épisodes, l'animateur-journaliste pose des questions à un unique invité. Ces questions sont souvent effacées au montage.

Le documentaire ou le reportage concernent les podcasts dont l'enregistrement a majoritairement été réalisé hors studio, sur un temps plus long et favorisant la mixité des points de vue. Le documentaire a, par rapport au reportage, une dimension narrative plus marquée et autorise une implication et une mise en scène de la subjectivité de son auteur plus fortes et plus visibles. Ce format occupe la deuxième place dans notre recensement (26,2%).

La chronique représente 13,8% de la base et désigne un format prenant la forme d'un commentaire dont le point de vue est plus ou moins personnel par l'animateur sur un sujet donné, sans interlocuteur.

Les émissions représentent 10% de l'ensemble. Nous classons dans cette catégorie les podcasts réunissant plusieurs personnes (animateurs, chroniqueurs, invités éventuels) au sein du studio, et étant clairement organisé en plusieurs séquences (chroniques ; débat ; reportages, *etc.*).

Le récit fictionnalisant (9%) désigne un format dans lequel l'animateur se fait narrateur. Seuls s'ajoutent à sa voix des sons d'ambiance, des bruitages, et autres artifices de dramatisation (absence d'interview ou d'archives au montage donc). Ce format est particulièrement utilisé dans les podcasts de faits divers (27,6% des podcasts de faits divers ont recours à ce format, et 52,6% au format du documentaire/reportage).

Le portrait (7% de l'ensemble) correspond à un format qui décrit une personnalité, connue ou non.

Les podcasts catégorisés comme relevant du format « débat » sont ceux qui impliquent une dynamique de groupe, la coprésence d'un ou plusieurs animateurs, et d'au moins deux invités incarnant plusieurs points de vue, ceux-ci n'étant pas nécessairement contradictoires. Ils représentent 4% de l'ensemble.

Les captations et retransmissions d'un événement (table ronde, conférence, *etc.*) publiées sous format podcast ne représentent qu'1,4% de l'ensemble.

Enfin, la catégorie « autres » (9 titres de podcasts) est appliquée aux formats hybrides tels que des docu-fictions, ou aux titres revendiquant une multiplicité de formats, comme le titre *Multimorphoses* produit par Fanny Cohen Moreau, dont la description sur les plateformes spécifie : « *Multimorphoses, c'est un podcast qui a pour particularité de changer de sujet, de forme, et d'équipes à chaque épisode. Les épisodes seront donc totalement indépendants les uns des autres. Notre défi : traiter des sujets d'une façon différente, se forcer à se renouveler, sortir de sa zone de confort pour vous proposer des formats innovants avec des thématiques peu traitées dans le monde du podcast* ».

Rubriques et mots-clefs (items n°13 et 14)

Nous avons par ailleurs regroupé les titres de podcasts en 17 rubriques thématiques, présentées dans l'encadré n°1.

Actualité générale (podcasts traitant de l'actualité, au sens large)
Culture
Écologie/nature
Économie/emploi
Faits divers/justice
Histoire
Journalisme (podcasts traitant de la pratique journalistique, montrant « l'envers du décor », occasion d'un discours réflexif sur le métier)
Politique (podcasts traitant de « l'enjeu » politique (Gerstlé, 1992) : féminisme, racisme, luttes sociales, guerre, <i>etc.</i>)
Religion
Santé
Sciences
Social et société (podcasts traitant d'enjeux sociétaux, comme le logement, le pouvoir d'achat, l'éducation, la famille, la sexualité, <i>etc.</i>)
Sport
Tranches de vie (podcasts donnant la parole à des « gens ordinaires » qui racontent leur métier, leur histoire, ce qui leur est arrivé, l'initiative qu'ils ont prise, le jour où leur vie a changé, <i>etc.</i>)
Vie politique (podcasts traitant du « jeu » politique (Gerstlé,1992) : élections, <i>etc.</i>)
Vie quotidienne et vie pratique (podcasts traitant de petits sujets et tracas du quotidien, prenant le plus souvent la forme de "tutos" ou un ton humoristique)
Autre (podcasts dont le sujet varie d'un épisode à l'autre. Ce sont souvent des titres valorisant l'expérimentation sonore, plutôt qu'une quelconque thématique.)

Encadré 1. Les 17 thématiques de la base de données

Les rubriques les plus représentées sont la culture (353 titres), le sport (111 titres) et la politique (87 titres). Les rubriques les moins représentées sont le journalisme (34 titres), la religion (15 titres) et la rubrique « Autre » (8 titres).

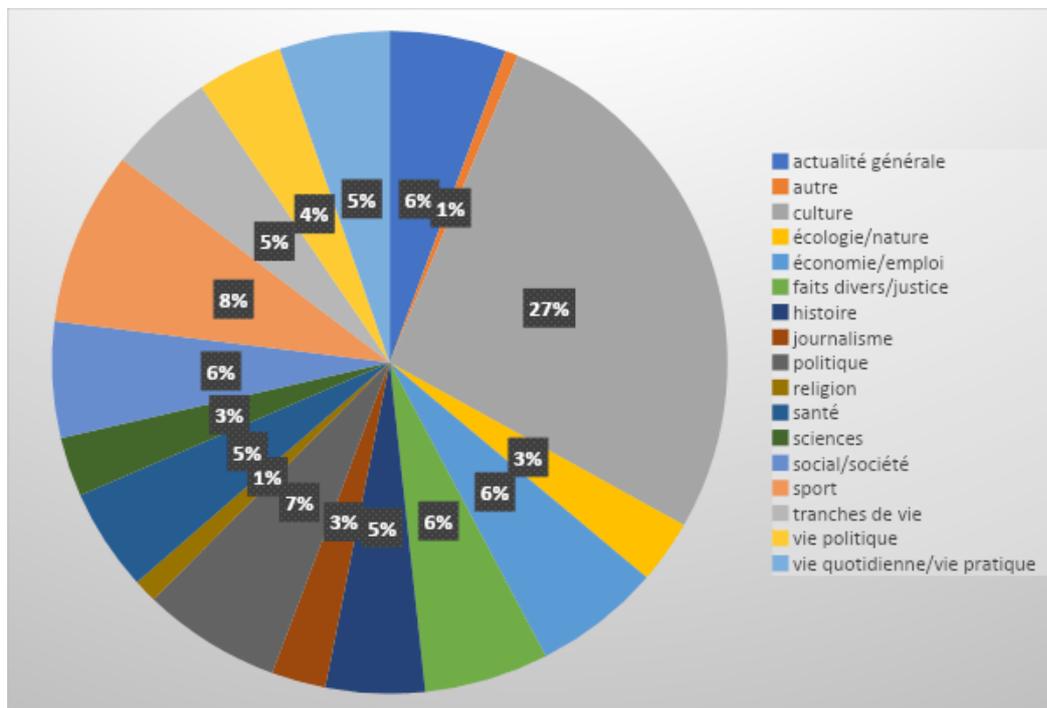


Figure 2. Répartition des thématiques des titres de podcast (en % de l'ensemble)

Dans plus de la moitié des cas ($n=750$), en dépit d'une rubrique assez clairement identifiable, il s'avère nécessaire de préciser cette dernière à l'aide d'un ou deux mots-clés, et cela pour deux raisons principales. Premièrement, une seconde thématique peut également apparaître centrale dans la construction éditoriale du titre. C'est par exemple le cas du titre *Salut l'info !*, coproduit par *Franceinfo* et *Astrapi* dont la rubrique principale est l'actualité générale, mais qui s'adresse spécifiquement à un jeune public. Dans ce cas, le mots-clés « enfants » vient compléter la description thématique du titre. En second lieu, la rubrique principale est parfois trop large et appelle à être précisée. C'est particulièrement le cas des podcasts « sport » qui sont dès lors systématiquement assortis d'un mot-clé précisant le sport dont il s'agit, ainsi que des podcasts « culture ». En effet, cette dernière rubrique est très large, tout comme la définition complexe et polysémique du terme. Là aussi, des mots-clés ont ainsi systématiquement été adossés aux titres concernés et précisent à quelle définition de la culture renvoie le titre (culture générale ; patrimoine) et à quel type de production culturelle, le cas échéant (littérature, cinéma, théâtre, photographie, etc.). Ici, les mots-clés précisent et nuancent la première position occupée par la rubrique « culture » (voir figure n°3). Au total, 80 mots-clés sont utilisés (encadré n°2), avec 163 combinaisons (mot-clé seul ou combiné à un second mot-clé).

TYPE DE PRODUCTEUR	NOM DU PRODUCTEUR	NOM	Rubrique	Mots-clefs	Format
Presse magazine	Télé Star	Quoi voir ce soir By Télé Star	culture	télévision	Chronique
Radio	Nostalgie	100 ans de la radio	culture	médias	Documentaire/reportage
Presse magazine	Vanity Fair France	En séries	culture	Télévision	Émission
PQR	Midi Libre	L'accent des Causse et Cévennes	culture	patrimoine	Documentaire/reportage
Radio	Chérie FM	Horizon	culture	enfants	récit
Pureplayer	The Conversation France	Je ne veux pas travailler	culture	philosophie	Interview

Figure 3. Aperçu de la base de données mettant en évidence la fonction des mots-clefs pour la thématique « culture »

- Actualité culturelle
- Arts
- Astrologie
- Basket
- Bière
- Cinéma
- Complotisme
- Covid
- Culture générale
- Cyclisme
- Développement personnel
- E-sport
- Écologie
- Éducation
- Élections
- Enfants : podcasts destinés aux enfants
- États-Unis
- Fabrique des programmes
- Faits divers
- Féminisme : podcasts traitant frontalement du féminisme ou se revendiquant féministes
- Féminisme ? : podcasts « donnant la parole aux femmes », parlant de femmes au parcours remarquables, de femmes-entrepreneures ; les podcasts ont une approche « néolibérale » du féminisme.
- Football
- Formule 1
- Gastronomie
- Genre
- Golf
- Histoire
- Hockey sur glace
- Humour : les podcasts adoptent un ton humoristique
- Inceste
- International
- Jeu
- Jeunesse : podcasts donnant la parole à/traitant de la jeunesse
- Jeux vidéo
- Langues
- LGBTQI+
- Littérature
- Luttes sociales
- Médias
- Mer
- Météo
- Minorités ethnoraciales
- Mode
- Montagne
- Musique
- Numérique
- Parentalité
- Patrimoine
- Philosophie
- Photographie
- Podcast : les podcasts traitent du podcast, soit sous forme de curation, soit sous forme de conseils
- Psychologie
- Publicité
- Randonnée
- Relations
- Religion
- Rugby
- Running
- Russie
- Sexualité
- Ski
- Société
- Sociologie
- Sports de combat
- Sports de glisse
- Technologies
- Télévision
- Tennis
- Terrorisme
- Théâtre
- Tranches de vie : si le podcast ne repose pas exclusivement sur le témoignage de « gens ordinaires », il mobilise néanmoins cette parole. Ce mot-clef permet de repérer facilement l'attribution de la parole à des « anonymes », des « ordinaires ».
- Ukraine
- Vin
- Voile
- Voiture
- Voyage
- Whisky
- Non : absence de mots-clefs

Encadré 2. Liste des mots-clefs thématiques de la base de données

LE TRIPLE APPORT DE LA BASE DE DONNÉES

En faisant un retour sur les initiatives de recensement des podcasts préexistantes qui introduisaient cet article, nous proposons à présent de mettre en évidence les possibilités qu'offre la base d'Obcast, tant par les données qu'elle permet d'obtenir que par la façon dont ces dernières ont été extraites. Un des points forts et particulièrement précieux des recensements précédents concerne les outils de visualisation qui sont mobilisés, notamment dans le cas de *Cosa Vostra*, *Podacst City* et *PodcastRe*. Dans ce dernier cas, le recensement est également réalisé de manière automatisée grâce à la collecte des métadonnées associées aux podcasts. Ce faisant, les équipes récoltent un nombre très important de données que notre collecte ne nous a pas permis d'obtenir jusqu'à présent. Cependant, il nous semble qu'en dépit de ces visualisations « cartographiques » du podcast, le recensement d'Obcast est la première base faisant état de façon exhaustive de l'offre de podcasts par les médias traditionnels français, notamment de presse écrite. De plus, les nombreux items, renseignés dans une dynamique de recherche collective, donnent la possibilité de ne pas cantonner les points d'intérêt aux métadonnées disponibles et associées aux fichiers podcasts. D'une part, ces dernières fournissent des informations précieuses quant au podcast concerné, mais ne permettent pas nécessairement de répondre à toutes les questions qu'une recherche sur le podcast pose. D'autre part, ces métadonnées ont été renseignées par les producteurs des podcasts et les plateformes et peuvent donc être considérées comme des productions discursives. Les items de la présente base de données permettent ainsi de les interroger en tant que telles (identification et inventaire des étiquettes thématiques des plateformes, items n°16 et 17) et d'analyser les caractéristiques des podcasts de façon plus fine à travers un mode de collecte notamment fondé sur l'exploration et sur l'écoute des titres de podcasts, par opposition à une collecte automatisée « cantonnée » à l'aspiration des flux RSS. Nous faisons à présent état de trois principaux apports de cette base de données du podcast natif d'information français : des statistiques précises, des hypothèses de travail renouvelées et des corpus pertinents facilement repérables.

Des statistiques précises sur l'offre de podcasts natifs d'information

Comme nous avons signalé dans les développements précédents, le premier apport de notre base de données concerne les statistiques qu'elle autorise à produire sur l'offre de podcasts natifs d'information français. Nous complétons ici les chiffres précédemment mentionnés avec quelques statistiques saillantes, concernant la quantité de podcasts produite d'abord, et la temporalité ensuite (périodicité, régularité, durée).

L'investissement dans le podcast concerne une pluralité d'acteurs. En effet, les trois plus gros producteurs, en termes de titres de podcast, sont *Arte Radio* (60 titres), *Slate* (53 titres) et *L'Indépendant* (45 titres). En termes de nombre d'épisodes total produits, les plus gros producteurs sont le label *Choses à savoir* (10 387 épisodes, 21 titres), le studio *Bababam* (4 041 épisodes, 13 titres), et *Le Dauphiné Libéré* (2 876 épisodes, 3 titres). Dans le haut du classement, se trouvent donc des acteurs natifs du podcast (studio, label), tout comme des acteurs traditionnels (télévision, presse écrite).

Du côté de la temporalité, la périodicité la plus fréquente est celle de l'hebdomadaire (356 titres), suivie de la série (318 titres) et du bimensuel (131 titres). La périodicité quotidienne concerne 57 titres. Il est à noter que 228 titres de podcast ont une périodicité qu'on peut qualifier d'irrégulière, c'est-à-dire dont la sortie des épisodes ne permet pas d'identifier une quelconque forme de récurrence, en dépit de ce qui est parfois annoncé dans les discours d'escorte. De plus, 207 titres ont eu une durée de vie de moins de trois mois et 372 de moins d'un an. 496 titres sont interrompus, soit 36,8% des 1345 titres recensés. Ce chiffre monte à 51% si l'on exclut les titres éditorialisés autour d'une périodicité close (les unitaires et les séries, construits autour d'un nombre défini d'épisodes). En d'autres

termes, plus de la moitié des titres recensés ne sont plus mis à jour, posant ainsi la question des coûts (temporels, matériels et économiques) sous-tendus par une production régulière de podcasts et, par voie de conséquence, celle de leur monétisation.

La durée moyenne des épisodes est de 22 minutes, durée qui correspond à la recommandation majoritairement établie par les acteurs du secteur et par les manuels (voir notamment le chapitre « *Determining the length for your show* », Morris, Tomasi, 2021, p. 94-95). Néanmoins, toutes les durées sont représentées dans la base de données : 140 titres ont une durée moyenne inférieure ou égale à 3 minutes, et 114 durent une heure ou plus (dont 6 plus de deux heures).

Faire émerger des hypothèses de travail

Le deuxième apport de cette base de données est que sa mise en œuvre fait émerger des questions de recherche, directement issues des informations et des obstacles rencontrés lors de la collecte. Retenons trois exemples.

Tout d'abord, une difficulté est apparue quant à l'identification de l'offre de podcasts des éditeurs de presse. En effet, l'accès et la valorisation des podcasts sont très différenciés d'un média à l'autre. Concernant les titres de presse quotidienne nationale, on observe ainsi que le site du *Monde* est doté d'une sous-rubrique « podcasts » au sein de la rubrique « Actualités ». Un clic sur cette sous-rubrique amène directement à l'espace dédié aux podcasts du *Monde* dont les différents titres sont clairement identifiables (figure n°4). En cliquant sur les différents titres de podcast, le résumé s'affiche, suivi des différents épisodes, accompagnés d'un bref article de mise en contexte et d'un lecteur audio (*player*). Il existe également sur cette page une mise en valeur des épisodes les plus écoutés, qui font l'objet d'un encart dédié.

The screenshot shows the top navigation bar of the Le Monde website. The main logo 'Le Monde' is centered. To the right, there are links for 'Se connecter' and 'S'abonner'. Below the navigation bar, there is a horizontal menu with categories: ACTUALITÉS, PRÉSIDENTIELLE 2022, ÉCONOMIE, VIDÉOS, DÉBATS, CULTURE, M LE MAG, and SERVICES. The main content area is divided into two columns. The left column, titled 'EN CE MOMENT', features several highlighted topics: Guerre en Ukraine, Élection présidentielle 2022, Corse, Coronavirus et pandémie de Covid-19, Passe vaccinal, Procès des attentats du 13-Novembre, and Mali. The right column, titled 'ACTUALITÉS', lists various sub-topics: International, Planète, Santé, Politique, Sciences, Big Browser, Société, M Campus, Disparitions, Les Décodeurs, Le Monde Afrique, Podcasts, Sport, Pixels, and Le Monde & Vous.

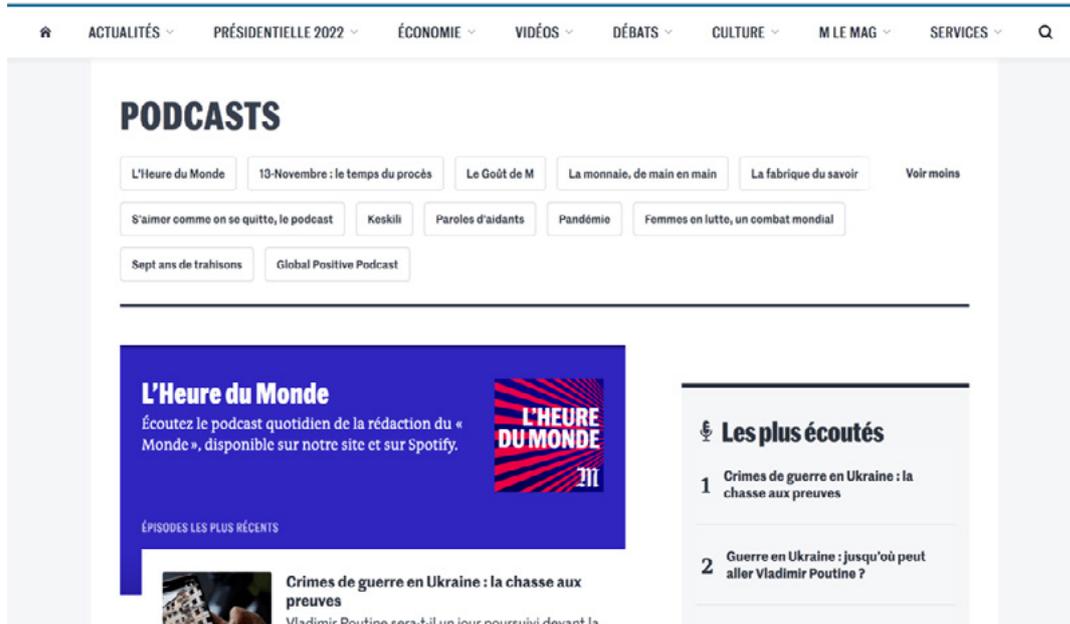


Figure 4. Les podcasts sur Lemonde.fr, mars 2022

Pour *Libération*, le site ne propose pas de rubrique dédiée aux podcasts dans son architecture, et il est compliqué de se repérer sur le site du fait de l'absence d'un espace faisant l'inventaire des productions audio du journal. Ce sont plutôt les épisodes qui font l'objet d'une présentation, sous forme de courts articles écrits, ne permettant pas de repérer une éventuelle cohérence éditoriale au sein d'un même titre de podcast. Cette dernière n'est ainsi pas mise en valeur en tant que telle. Le podcast en tant qu'objet médiatique ne fait donc pas l'objet d'une valorisation intégrée à l'architecture du site. Nous faisons des constats similaires concernant les sites de *l'Humanité* et du *Figaro* : il n'y a là non plus pas de rubrique « Podcast ». En revanche, des traces de cette catégorie apparaissent en entrant manuellement l'adresse URL <http://www.titredusite.fr/podcasts>. Ce faisant, on trouve alors, dans le cas du *Figaro*, une offre clairement déclinée et régulièrement mise à jour.

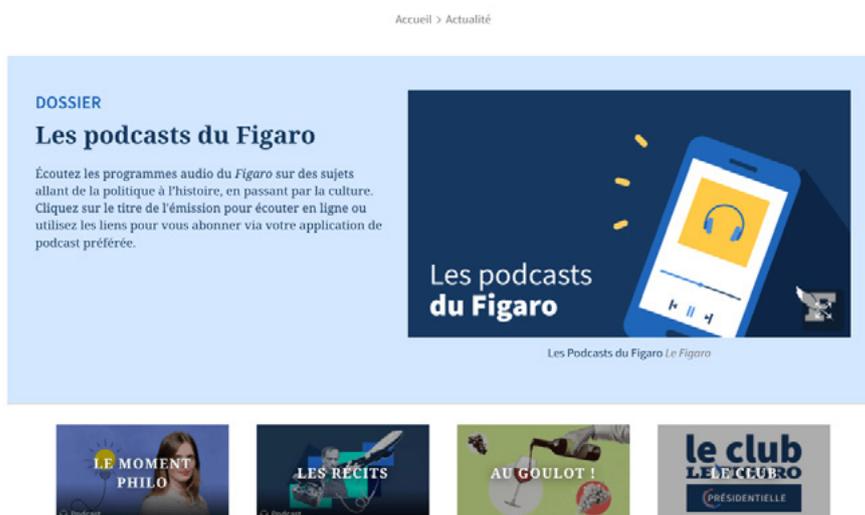


Figure 5 : les podcasts sur Lefigaro.fr, mars 2022

Les podcasts sont ainsi hébergés et valorisés de façon très hétérogène sur les différents sites de presse, ce qui complique le travail de cartographie. Surtout, cela incite à interroger ce travail de valorisation des podcasts par leurs producteurs. Si cette hétérogénéité est en partie corrélée à celle de la quantité de titres de podcasts produits par les éditeurs de presse, il conviendra également d'interroger le type de stratégies et leur degré d'élaboration relatifs à l'investissement de ce nouveau territoire médiatique par des acteurs traditionnels.

Le second exemple de pistes de recherche ouvertes par la base de données concerne le rythme de lancement de nouveaux podcasts. Nous constatons un fort engouement de la part de certains titres de presse, mais qui semble rapidement retombé. En effet, depuis le mois de septembre 2017, de nouveaux podcasts ont été créés chaque mois, pour un nombre de 20 titres par mois en moyenne entre septembre 2017 et juillet 2022. Nous l'avons vu plus haut, le nombre de nouveaux titres croît de manière exponentielle à partir de 2017 (tableau n°1). Lors de ces années d'engouement, certains acteurs se démarquent par leur très forte productivité. C'est le cas du titre de presse quotidienne régionale *L'Indépendant*, qui est à l'initiative de 45 titres de podcast, dont 30 lancés entre le 26 août 2019 et le 21 décembre 2020. Parmi ces 30 titres, 21 sont aujourd'hui interrompus, et 18 ont eu une durée de vie de moins de deux mois, quand la promesse éditoriale semblait annoncer une production régulière ou pérenne dans le temps. Comment interpréter cet engouement, rapidement suivi d'un essoufflement ? Peut-on y voir le signe de stratégies de lancement dans le podcast mal élaborées de la part des titres concernés ? D'un engouement ayant mal pris en compte les enjeux économiques sous-tendus par le format (ratio coûts/monétisation) ? D'un foisonnement d'initiatives individuelles dispersées au sein des rédactions ? Ces observations en ligne d'ordre statistique conduisent à interroger les stratégies mises en place par les entreprises de presse et cela à l'aide d'entretiens semi-directifs.

Le dernier exemple que nous souhaitons aborder ici nous incite à interroger la diffusion du podcast dans une perspective socio-technique. En effet, sur les 1 345 titres de podcasts recensés, chaque plateforme de diffusion que nous avons indexée en héberge environ 80% (81% pour *Deezer*, 84% pour *Castbox* et *Spotify*, 87% pour *Apple podcast* et 90% pour *Google podcast*). En revanche, seulement 15% sont mis en ligne sur *Youtube* et *Soundcloud*. En effet, la proportion de podcasts disponibles sur *Youtube* ou *Soundcloud* est plus importante pour les titres moins récents. Ainsi, en 2017, respectivement 32,6% et 58,1% des nouveaux titres sont disponibles sur *Youtube* et *Soundcloud*. Ce chiffre chute progressivement, pour atteindre respectivement 15,6% et 6,7% des nouveaux titres en 2021. Alors que des chercheurs ont mis en évidence qu'elles étaient des plateformes « historiques » de diffusion des podcasts (Berry, 2015), cette migration vers d'autres plateformes semble être le signe d'un processus de structuration du secteur, ainsi que celui de sa professionnalisation. Nous avons en effet constaté que des podcasts, indépendants en général, étaient disponibles sur *Soundcloud*, puis à partir d'un moment, ne l'étaient plus. C'est le cas des titres de podcast *Course épique*, et *Histoire de Darons*. Dans les deux cas, un court épisode explique que désormais, le podcast ne sera plus diffusé sur *Soundcloud*. Pour *Histoire de Darons*, son créateur Fabrice Florent explique : « Aujourd'hui, *Soundcloud* est à part de mon hébergeur, et je voudrais tout rassembler dessus. » (Florent, 2019). L'auteur veut gérer l'ensemble des espaces de diffusion à partir d'un même hébergeur (ici, *Acast*) et ainsi homogénéiser et simplifier la gestion de son flux. Cela indique une forme de rationalisation technique de l'hébergement et de la diffusion des podcasts, où l'on passerait de manières de faire relativement bricolées (*Soundcloud* est d'ailleurs une plateforme historique de la diffusion sonore amateur dans son ensemble) à un environnement technique structuré autour d'acteurs et de pratiques professionnalisés. Ces observations doivent être prolongées et la question de la rationalisation technique doit accompagner celle de la standardisation du média podcast et de ses formats.

Construire des corpus

Enfin, le troisième apport de la base de données réside en ce qu'elle permet d'identifier et de construire des corpus d'analyse pertinents. En ce sens, elle constitue un outil précieux d'ingénierie de la recherche, notamment dans sa dimension collective.

Au sein du programme Obcast, une équipe de travail est par exemple dédiée à l'étude des podcasts féministes, c'est-à-dire traitant frontalement du féminisme ou revendiquant une approche féministe (quelle qu'elle soit). D'une part, la base de données a bénéficié de la connaissance de l'équipe de ce terrain avec l'apport de titres et, d'autre part, elle permet de rapidement repérer les titres de podcasts concernés par leur axe de travail (grâce aux rubriques et aux mots-clefs), ainsi que l'ensemble des métadonnées et caractéristiques qui y sont associées et ont été présentées dans cet article.

De la même façon, la construction de la base de données met en évidence deux corpus particulièrement intéressants : d'une part, les podcasts sur les coulisses du journalisme (voir Dollé, 2020), dont l'existence s'est avérée significative au fur et à mesure de la collecte, puisque nous en recensons 34 titres. Cette rubrique singulière pose la question de la façon dont le podcast peut incarner pour les journalistes et les rédactions un moyen de pallier la défiance des citoyens, et notamment des jeunes, à l'égard des médias dits *mainstream*, tout autant qu'un moyen de faire du podcast à moindre coût, en faisant intervenir des collègues qui racontent les coulisses d'une enquête, à l'image du titre *Making O.F. de Ouest France*. D'autre part, nous avons également relevé un plus petit nombre de titres ayant pour thématique principale le média podcast. Les « méta-podcasts » (Berry, 2015, p. 171) forment ainsi un corpus de 9 titres, produits par des studios (n=4), des plateformes d'hébergement (n=2), des amateurs (n=2) et un *pure player*, lancés entre 2017 et 2022. Ces titres sont intéressants en termes d'insertion du podcast dans le champ élargi de la critique culturelle, voire comme agents de légitimation de la pratique et de l'objet, mais aussi pour interroger l'éventuelle normalisation du podcast et la professionnalisation de son secteur (de nombreux épisodes traitent ainsi de questions comme : comment faire un bon podcast ? Comment bien choisir sa thématique et son format ? Comment monétiser son podcast ? *etc.*)

CONCLUSION

Cet article avait pour objectif d'expliquer pourquoi et comment réaliser une base de données du podcast natif d'information français. Nous reviendrons donc ici d'abord sur le *pourquoi*, puis sur le *comment*.

Pourquoi ? D'abord, un tel travail a vocation à participer au développement récent et à la consolidation en cours de la situation française du podcast et ce, en lui tendant un miroir, rendant visible aux acteurs l'offre qu'ils constituent, tant dans sa dimension quantitative (nombre de producteurs, nombres de titres et d'épisodes) que qualitative (thématiques, formats, durées, *etc.*). C'est bien là le but premier d'un observatoire : observer, certes, mais pour faire voir. Ensuite, ce recensement constitue finalement un premier jalon important pour les recherches en cours et à venir sur le secteur du podcast permettant, comme nous l'avons montré, de faire émerger tout à la fois questions, hypothèses et objets de recherche.

Comment ? La méthodologie employée ici revêt incontestablement un caractère artisanal, particulièrement saillant dans le mode de collecte des données liées aux podcasts. Cela implique de fait des difficultés quant à l'exhaustivité recherchée dans un tel recensement, l'offre de podcasts étant caractérisée par un mouvement continu : de nouveaux titres apparaissent et disparaissent chaque jour. Si ces difficultés sont en partie surmontées pour les

titres de podcasts produits par des médias traditionnels et des studios de podcast, elles persistent vis-à-vis des podcasts amateurs et indépendants. Cela pose la question de la nécessité d'une automatisation du recensement grâce à des outils issus de l'informatique. L'automatisation de la base fait, à ce jour, l'objet de réflexions, à la fois dans des logiques de partenariats nationaux et internationaux et à partir de corpus et de questions de recherche préalablement déterminées. À ce titre, Maëlle Bazin, membre du programme Obcast, a d'ores et déjà commencé à travailler avec le laboratoire numérique de l'INA à l'automatisation d'une partie de la base de données sur la question des représentations de genre dans les podcasts. Néanmoins, nous avons montré que l'artisanat d'une collecte manuelle constitue également une ressource pour la recherche, définie alors non seulement comme une méthode, mais également comme une *exploration*, participant pleinement à la richesse des enquêtes menées à partir d'elle^[11].

NOTES

[1] Ce programme est dirigé par Arnaud Mercier, que nous remercions ici pour ses éclairages et apports relatifs à ce travail de recensement. Nous remercions également Flore Di Sciullo, chercheuse post-doctorante dans le cadre du projet Obcast, pour ses apports relatifs à la construction de la base de données ici présentée. Enfin, nous remercions l'ensemble des membres du programme Obcast pour les échanges précieux qui ont enrichi ce travail de recensement.

[2] Un type de podcasts qui semble gagner en légitimité, puisqu'il a pour la première fois fait l'objet d'un prix spécial lors du dernier Paris Podcast Festival en octobre 2022.

[3] Cette cartographie est disponible à l'URL <https://map.audiomeans.fr/>, consulté le 7 décembre 2022.

[4] La distinction entre studio de production et label est récente et la définition de ce dernier n'est pas stabilisée au sein du secteur du podcast. À ce stade, nous pouvons définir le label comme un regroupement d'auteurs de podcast indépendants et préexistants permettant une meilleure communication et monétisation des titres, quand un studio participe directement à la conception et à la production des podcasts.

[5] Médias associatifs ou n'appartenant pas à un groupe de média.

[6] Cette dénomination désigne principalement les plateformes de diffusion (qui se font alors productrices de podcasts (Spotify, Deezer, etc.) ou les entreprises au statut mixte telles que l'AFP.

[7] Nous remercions Maëlle Bazin, doctorante et assistante de recherche au Carism, qui a élaboré cet item, nous permettant d'objectiver des données relatives au genre dans les podcasts natifs, notamment à la représentation des femmes.

[8] Les stratégies de diffusion et de valorisation des podcasts par les médias traditionnels sur leur propre site font l'objet d'une publication en cours d'évaluation.

[9] À ce jour, la base de données de l'Obcast est uniquement accessible aux membres du programme de recherche.

[10] Les stratégies de diffusion et de valorisation des podcasts par les médias traditionnels sur leur propre site font l'objet d'une publication scientifique en cours d'évaluation.

[11] À ce jour, la base de données de l'Obcast est uniquement accessible aux membres du programme de recherche.

RÉFÉRENCES BIBLIOGRAPHIQUES

- Audiomeans (2021), « Podcast city, une ville foisonnante au creux de l'oreille », *Audiomeans*, [en ligne], consulté le 2 décembre 2022, <https://www.audiomeans.fr/post/podcast-city-une-ville-foisonnante-au-creux-de-l-oreille>
- Berry, Richard (2015), « A Golden Age of Podcasting? Evaluating Serial in the Context of Podcast Histories », *Journal of Radio & Audio Media*, vol. 22, n°2, p. 170-178.
- Bonini, Tiziano (2015), « The Second Age of Podcasting: Reframing Podcasting as a New Digital Mass Medium », *Quaderns del CAC*, vol. 41, p. 21-30.
- Clifton, Ann ; Reddy, Sravana ; Yu, Yongze ; Pappu, Aasish ; Rezapour, Rezvaneh ; Bonab, Hamed ; Eskevich, Maria ; Jones, Gareth ; Karlgren, Jussi ; Carterette, Ben ; Jones, Rosie (2020), « 100,000 Podcasts: A Spoken English Document Corpus », *COLING*, [en ligne], consulté le 7 décembre 2022, <https://www.aclweb.org/anthology/2020.coling-main.519>
- Cohen, Évelyne (2019), « La baladodiffusion : de la réécoute à la création sonore de podcasts », *Sociétés & Représentations*, vol. 48, n°2, p. 159-167.
- Demazière, Didier ; Gadéa, Charles (dir.) (2009), *Sociologie des groupes professionnels : Acquis récents et nouveaux défis*, Paris : La Découverte (collection « Recherches »).
- Dollé, Ella (2020), « Le podcast «L'envers du récit» : les coulisses de la fabrication de l'information montrées par *La Croix* », *Effeuilage*, vol. 9, n°1, p. 70-74.
- Florent, Fabrice (2019), « IMPORTANT ! Changez de flux pour *Histoires de Darons* », *Histoires de Darons*, décembre 2019, [en ligne], consulté le 7 décembre 2020, <https://soundcloud.com/histoires-darons>
- Gameiro, Aurélie (2021), « La cartographie «ultime des podcasts» mise à jour en 2022 », *CosaVostra*, 3 mars 2021, [en ligne], consulté le 7 décembre 2022, <https://www.cosavostra.com/guides/la-cartographie-ultime-des-podcasts-mise-a-jour-2021/>
- Gerstlé, Jacques ; Duhamel, Olivier ; K. Davis, Dennis (1992), « La couverture télévisée des campagnes présidentielles. L'élection de 1988 aux États-Unis et en France », *Pouvoirs*, n° 63, p. 53-69.
- Hoyt, Eric ; Bersch, J.J. ; Noh, Susan ; Hansen, Samuel ; Mertens, Jacob ; Morris, Wade Jeremy (2021), « PodcastRE Analytics : Using RSS to Study the Cultures and Norms of Podcasting », *DHQ : Digital Humanities Quarterly*, vol. 15, n°1, [en ligne], consulté le 7 décembre 2022, <http://www.digitalhumanities.org/dhq/vol/15/1/000519/000519.html>
- Legavre, Jean-Baptiste ; Rieffel, Rémy (dir.) (2017), *Les 100 mots des sciences de l'information et de la communication*, Presses Universitaires de France (collection « Que sais-je ? »).
- Lesauvier, Marie-Eva ; Di Sciuolo, Flore ; Mercier, Arnaud (2022), « L'irrésistible essor des podcasts d'information », *The Conversation*, 10 novembre. En ligne : <https://theconversation.com/lirresistible-essor-des-podcasts-dinformation-193961>
- Les Échos (2021), « 2020, année de la consécration pour les podcasts », *Les Échos*, 3 février 2021, [en ligne], consulté le 24 avril 2023, <https://www.lesechos.fr/tech-medias/medias/2020-annee-de-la-consecration-pour-les-podcasts-1287215>
- Lochard, Guy (1996), « Genres rédactionnels et appréhension de l'événement médiatique. Vers un déclin des « modes configurants » ? », *Réseaux*, vol. 76, no. 2, p. 83-102.
- MadameBB (2020), « État des lieux des podcasts de marque en France », [en ligne], consulté le 7 décembre 2022, <https://www.madamebb-prod.com/podcasts-de-marque>

Médiamétrie, « Global Audio. La radio, pilier de la consommation audio », communiqué de presse, 3 mai 2022

Morris, Wade Jeremy ; Patterson, Eleanor (2015), « Podcasting and its Apps : software, sound, and the interface of Digital Audio », *Journal of Radio & Audio Media*, vol.22, n°2, p. 220-230.

Morris, Wade Jeremy (2021), « Infrastructures of discovery: examining podcast rating and ranking », *Cultural Studies*, vol.35, n°4-5, p. 728-749.

Morris, Tee ; Tomasi, Chuck (2021), *Podcasting for Dummies*, 4th edition, John Wiley&Sons Inc.

Pironnet, Laura (2020), « Mise à jour 2020 du panorama des podcasts de médias de presse écrite », *Medium*, 18 mai 2020, [en ligne], consulté le 7 décembre 2020, <https://medium.com/@pironnetlaura/mise-%C3%A0-jour-2020-du-panorama-des-podcasts-de-m%C3%A9dias-de-presse-%C3%A9crite-44a84c0c2bb2>

Resler, Seth (2017), « Do podcasts have a discoverability problem? », *Jacobs Media*, [en ligne], consulté le 2 décembre 2022, <https://jacobsmedia.com/do-podcasts-have-a-discoverability-problem-we-asked-leaders-in-the-space/>

Soulages, Jean-Claude (2007), *Les rhétoriques télévisuelles. Le formatage du regard*.