

Observatoire du traitement médiatique des enjeux écologiques

Remerciements	1
Pourquoi un observatoire des médias sur l'écologie ?	2
MÉTHODOLOGIE	
Le dictionnaire : une catégorisation des enjeux écologiques adaptée à l'espace médiatique	4
Les sources du dictionnaire	5
L'adaptation au discours médiatique	6
Les types de catégories	8
Les mises à jour du dictionnaire	10
Un algorithme identifiant les mots-clés du dictionnaire	11
Périmètre de l'Observatoire	11
Les médias audiovisuels	12
Source de données	12
L'exclusion des publicités	12
Le calcul du temps d'antenne	13
Comptabilisation des extraits grâce aux mots-clés	13
Traitement des mots-clés à haut risque de faux positif	14
Justifications, éléments d'attention et limites	15
Analyses comparées et tests de la méthodologie	16
Analyses comparées et tests de la méthodologie	17
Comparaison avec les résultats d'analyses à la main	19
Vision comparée des durées de 15 à 40 secondes	20
Des erreurs tirées à la hausse par les séquences longues	21
Des séquences longues ne faisant pas plus consensus chez les bénévoles	22
Analyse de la mise en contexte des conséquences de la crise climatique et de l'effondrement de la biodiversité	23
Conclusion	26
Références	27
Baromètres	28

Remerciements

L'Observatoire des médias sur l'écologie n'aurait pas pu voir le jour sans le financement de l'Agence de la transition écologique (Ademe). Merci à Valérie Martin dont l'accompagnement bienveillant et l'expertise ont enrichi le projet et ont facilité sa mise en place.

L'Autorité de régulation de la communication audiovisuelle et numérique (Arcom) a contribué tant financièrement que méthodologiquement à la réalisation de l'Observatoire. Merci à Alexine Conaut pour son apport dans la conception méthodologique et pour son engagement sans faille dans la communication avec cette institution.

Merci aux structures du consortium, associations et sociétés, d'avoir mis sur pied ce projet. DataForGood, Eleven Strategy, Expertises Climat, Mediatree, Pour plus de climat dans les médias, QuotaClimat ont travaillé de concert en apportant leurs connaissances, leurs compétences, leur force humaine et leur soutien financier.

L'Observatoire n'existerait pas sans l'engagement de ses collaborateurs permanents : Raphaël Di Piazza, Laura Fousson, Célia Gautier, Paul Leclercq, Eva Morel, Claire Morvan, Claire Peltier, Jean Sauvignon, Louna Wemaere, Veronika Zagyi.

Cette équipe a été également épaulée par Raphaël Demonchy, Kevin Gamblin, Benjamin ..., Fanny Poirier, Marc-Antoin Samson, Kevin Tessier, Albin Wagener dont la contribution et le soutien ont, lors des différentes étapes du projet, nourri et accéléré sa réalisation.

Merci également à Pietro Turati pour son investissement et son appui continus.

L'Observatoire a bénéficié de l'accompagnement d'un comité d'experts composé de Nataly Botero, Laurent Cordonier, Clare Dale, Arthur Grimonpont, Nicolas Hervé et Marie-Élodie Perga. Leur expertise et leurs critiques bienveillantes ont irrigué la méthodologie tout au long de son élaboration. Merci à elles et eux.

De nombreux et nombreuses bénévoles ont aidé lors des différentes étapes du projet. Venant des associations membres du consortium, mais aussi d'autres cercles, ils et elles ont investi de leur temps et de leur énergie pour que l'Observatoire puisse voir le jour. Merci.

Merci à La Jujotte et notamment à Amandine Legrand, Thomas Dupeyrat et Sullivan Lecomte pour leur soutien.

Pourquoi un observatoire des médias sur l'écologie ?

L'ampleur du changement climatique actuellement constatée ([+1,2°C sur la période 2018-2023](#) par rapport à l'époque préindustrielle) est [sans précédent dans l'histoire humaine](#). Les engagements climatiques actuellement pris par les États mèneraient, s'ils étaient respectés, à un réchauffement d'environ +2,5°C à la fin du siècle, un niveau constituant une [menace directe](#) pour les conditions d'habitabilité de la planète pour des centaines de millions d'humains. Limiter le réchauffement à +1,5°C, comme s'y est engagée la communauté internationale à Paris en 2015, n'est possible qu'à condition d'une réduction drastique, immédiate et coordonnée des émissions mondiales.

La perception collective des menaces climatiques et écologiques est une condition limitante importante au soutien démocratique de solutions proportionnées à leur gravité. La mise en œuvre de politiques efficaces est peu probable si le paysage informationnel nous divise sur les causes, suscite des doutes sur la gravité ou remet en question l'existence du changement climatique.

Le grand public se déclare inquiet mais démontre une connaissance faible et disparate des enjeux climatiques. D'après la plus [vaste enquête](#) sur la perception collective du changement climatique, conduite auprès de 1,2 million de personnes issues de 50 pays différents, près de deux tiers des enquêtés pensent que le changement climatique constitue une urgence mondiale (PNUD et Oxford University, 2021). Cependant, les connaissances du public sur les causes, les effets, et les solutions au changement climatique sont en fort décalage avec l'état des connaissances scientifiques, même élémentaires. Un [sondage](#) réalisé pour le World Economic Forum dans 30 pays révèle qu'un tiers de la population remet en question l'origine anthropique du changement climatique et que près de quatre personnes sur dix accordent une confiance nulle à modérée dans les sciences du climat. En France, selon une étude de 2022 de la Fondation Descartes [34% de la population](#) estime que le changement climatique est autant causé par l'activité humaine que par des phénomènes naturels.

Le rôle de l'information et des médias

Selon le GIEC (IPCC, 2022), les médias sont une source majeure d'information sur le changement climatique.

Face à ce constat, le traitement médiatique du climat et de l'environnement est généralement jugé insuffisant et insatisfaisant par les Français. Selon l'enquête Fractures françaises réalisée par la Fondation Jean Jaurès en 2023, [l'environnement est la deuxième préoccupation des Français](#), après le pouvoir d'achat, et pourtant seuls 11% des Français se sentent "tout à fait informés" sur la crise climatique (Baromètre des assises du journalisme, 2021), tandis qu'entre 42% et 53% des Français trouvent que la question du changement climatique et de l'environnement n'est pas assez présente dans les médias ([Baromètre Kantar, 2023](#); [Cordonier, 2022](#)). Plus de deux tiers des Français déclarent manquer d'informations scientifiques sur le réchauffement climatique et près de trois quarts d'entre eux ont le sentiment que, sur ce sujet, "on entend tout et son contraire" ([Fractures françaises, 2022](#)).

Alors, qu'en est-il vraiment ? Quel est le niveau, en quantité et en qualité, du traitement médiatique des enjeux environnementaux aujourd'hui en France ? C'est pour répondre à cette question d'intérêt général que l'Observatoire des médias sur les enjeux écologiques a été créé. Il n'existait jusque-là aucune mesure systématique, aux niveaux national, régional ou mondial, du traitement médiatique des enjeux écologiques. Ce

manque de données fiables, unifiées et validées du point de vue méthodologique ne permettait pas aux pouvoirs publics, aux usagers et consommateurs de médias, ni aux médias eux-mêmes, d'agir en faveur d'une information suffisante et de qualité sur les bouleversements environnementaux.

Disparates ou partielles, les données disponibles pointaient malgré tout des évolutions récentes dans le traitement médiatique de ces thématiques. [Une étude menée par des chercheurs de l'INA en 2020](#) montrait qu'en vingt ans, le nombre de nouvelles consacrées à l'environnement dans les journaux télévisés du soir des cinq chaînes historiques avait plus que doublé, et que leur place avait été revalorisée dans la hiérarchie des informations. Elle notait toutefois que l'augmentation du nombre de sujets restait fortement liée à l'actualité (catastrophes, négociations internationales sur le climat...). De même, [une étude menée en 2023 par l'institut Tagaday](#) montrait que le nombre de contenus consacrés au climat était passé de 22 000 par mois en moyenne en 2013 à plus de 150 000 en 2023. Cependant, seuls 21% de ces contenus faisaient le lien entre les conséquences comme les événements météorologiques extrêmes et le dérèglement climatique. Et, selon l'[étude Viavoice pour les Assises du journalisme de 2021](#)¹, les Français ont jugé l'actualité sur la crise climatique trop "anxiogène" (35%), "catastrophiste" (33%) et "moralisante" (25%).

Les outils de mesure développés par plusieurs acteurs de la société civile montraient eux aussi une augmentation de la couverture médiatique du dérèglement climatique, qui n'excédait cependant pas 4% du temps médiatique, avec de légères fluctuations en fonction de l'actualité ([Baromètre de l'Affaire du siècle en 2022](#) ; [Baromètre Data4Good x QuotaClimat](#) ; [Observatoire des JT sur les changements climatiques de Climat Médias](#)). Le travail de suivi régulier mené par ces associations révélait également un manque de transversalité dans ce traitement médiatique, tendant à masquer les liens entre les différentes crises environnementales, et les interdépendances entre secteurs. Quant à la biodiversité, elle était particulièrement sous-représentée.

La naissance de l'Observatoire des médias sur les enjeux écologiques

Pour répondre à ce besoin de données disponibles, quatre associations² œuvrant pour une meilleure couverture médiatique des enjeux écologiques ont décidé de fédérer leurs compétences. Avec le soutien de l'Ademe est né le projet d'un observatoire national unique, un outil de mesure fiable, robuste sur le plan méthodologique et pérenne. Dans la mise en place de cette méthodologie l'équipe de l'Observatoire a été accompagnée et conseillée par des scientifiques et experts de l'analyse de données, de l'information journalistique et des sciences environnementales.

Inscrit dans un appel à commons³, l'algorithme, la méthodologie et les résultats de cet Observatoire sont mis à la disposition de tous les citoyens⁴ comme une ressource destinée

¹ Rapport non publié. Voir un article de France Info sur ce sujet : Urgence climatique, crise sanitaire : face à un traitement médiatique jugé "anxiogène", les Français attendent un journalisme de "solutions".

https://www.francetvinfo.fr/sante/maladie/coronavirus/urgence-climatique-crise-sanitaire-face-a-un-traitement-mediatique-juge-anxiogene-les-francais-attendent-un-journalisme-de-solutions_4784129.html

² Les associations Data for Good, Expertises Climat, Pour plus de climat dans les médias, QuotaClimat, rejointes par Eleven strategy, spécialisé dans l'intelligence artificielle et la data et Mediatree, éditeur de logiciel spécialisé en données audiovisuelles. Le régulateur national de l'audiovisuel, l'Arcom, s'est également affilié au projet en participant à son financement et en apportant son expertise notamment dans l'élaboration du périmètre analysé pour la radio et la télévision.

³ Appel à commons 2023.

https://wiki.resilience-territoire.ademe.fr/wiki/Appel_%C3%A0_Commons_Sobri%C3%A9t%C3%A9_e_t_R%C3%A9silience_-_2023#Carte_des_Projets_de_communs_Relev.C3.A9_24_Nov_2023

⁴ Les données sources audiovisuelles appartiennent à Mediatree et ne sont pas publiques.

à faciliter la transition écologique.

La production de données, via cet Observatoire, permet d'**objectiver les constats sur l'information environnementale** :

- L'Observatoire **mesure l'évolution** du traitement médiatique des enjeux écologiques (d'abord audiovisuel au moment de la première mise en ligne de l'Observatoire, plus tard de la presse écrite).
- L'Observatoire propose ces données sur **un tableau de bord, accessible via une plateforme web ouverte au grand public**. À terme, l'Observatoire aspire à devenir une source de données précises et fiables sur le traitement médiatique des enjeux écologiques, pour tous les acteurs.
- La mise à disposition de ces analyses constitue un maillon essentiel de la mise en mouvement des médias **vers une meilleure information des Françaises et Français sur les enjeux écologiques**.

À partir des données des médias audiovisuels français, l'Observatoire fournit des analyses concernant la prévalence des thématiques traitées, dont les constats, les causes, les conséquences et les solutions à la crise écologique. Il mesure également si les médias créent des liens entre ces thématiques : abordent-ils les causes en constatant les problèmes ou en parlant des conséquences ? Proposent-ils des solutions quand ils évoquent ces problèmes ?

Ces analyses reposent sur une méthodologie ouverte, accessible et transparente pour que chacun – citoyen, acteur politique, acteur médiatique... – puisse s'en saisir.

MÉTHODOLOGIE

Le dictionnaire : une catégorisation des enjeux écologiques adaptée à l'espace médiatique

Les enjeux associés à la préservation de l'environnement sont extrêmement complexes, variés (en fonction des géographies, des priorités politiques, de l'actualité), et souvent intrinsèquement liés. Si établir de manière exhaustive l'ensemble de ces enjeux n'est pas l'objectif de l'Observatoire, il adopte une approche transversale et systémique adaptée aux problèmes écologiques.

Les enjeux écologiques, tels qu'ils sont étudiés dans la littérature scientifique et traités par la puissance publique et les acteurs économiques, peuvent schématiquement être regroupés en trois catégories : le changement climatique, l'érosion de la biodiversité et la surexploitation des ressources naturelles.

Le traitement médiatique de ces problématiques environnementales évolue en raison de différents facteurs. Leur importance dans le débat public, le degré de spécialisation des journalistes sur ces questions et la qualité de leurs liens avec les sources concernées (scientifiques, politiques, responsables associatifs, acteurs économiques...) influencent la

construction médiatique de ces problèmes.⁵ Aussi, selon leurs formats, leur ligne éditoriale et leurs priorités, les médias sélectionnent et hiérarchisent l'information valorisant certains aspects de l'actualité. De façon plus générale, les perceptions professionnelles de ces problèmes évoluent également, ce dont témoignent les changements internes des organisations : différents médias ont pris récemment des initiatives pour valoriser les informations sur les enjeux environnementaux, grâce à la création de services d'information spécifiques, à la formation de leurs journalistes et/ou à l'instauration de chroniques, programmes et rubriques dédiés.

Pour quantifier et qualifier le traitement médiatique des trois grandes catégories d'enjeux écologiques et ses évolutions, le principe de l'Observatoire consiste en une analyse exhaustive des programmes d'information audiovisuels français au prisme d'un dictionnaire de termes recouvrant de manière transversale ces problématiques.

Les sources du dictionnaire

L'Observatoire produit actuellement⁶ des données à partir des flux de contenus informationnels des principales chaînes de télévision et radios françaises et repère les séquences portant sur les enjeux écologiques grâce à une liste de mots-clés répartis selon les trois catégories d'enjeux. Le dictionnaire de l'Observatoire, qui pourra faire l'objet de révisions périodiques, contient actuellement plus de 1700 termes⁷ (voir le dictionnaire en annexe). Ces mots ont été identifiés et sélectionnés en plusieurs étapes.

Une première liste de mots et d'expressions ont été agrégés à partir d'un ensemble de documents officiels faisant référence pour définir les politiques publiques en matière environnementale. La première étape de la construction du dictionnaire consistait à passer en revue les grandes thématiques des documents de référence et à retenir⁸ les mots désignant l'un des aspects des trois enjeux faisant consensus selon ces sources. Ces documents sont :

- Le 6e rapport du GIEC : [Sixth Assessment Report — IPCC](#)
- Le [rapport de l'évaluation mondiale de la biodiversité et des services écosystémiques](#)
- “Mon parcours transition écologique” - formation de référence en écologie pour les cadres de l'Etat : « [Mon parcours transition écologique](#) », [une formation pour les cadres supérieurs de l'État | Le portail de la fonction publique \(fonction-publique.gouv.fr\)](#)
- Les ressources du Secrétariat Général à la Planification Écologique, et notamment le tableau de bord de suivi des indicateurs : [Baromètre de la Planification Écologique - Infogram](#)

Les documents portent sur l'origine des problèmes écologiques, leurs impacts sur

⁵ Voir par exemple : Comby, Romanet, 2012 ; Comby, 2015 ; Lefort, Poels, Quinton, 2020 ; Luedecke, Boykoff, 2017;

⁶ Le périmètre sera prochainement élargi à la presse écrite.

⁷ Le nombre des termes est susceptible d'évoluer suite aux révisions régulières.

⁸ Les membres de l'équipe, spécialistes des enjeux écologiques, ont établi à partir de ces documents une première liste – de “référence” – après un travail d'exploration, de concertation, par décision collective.

l'environnement et les leviers d'action. Ils servent à établir des consensus et des synergies au niveau des connaissances et des politiques publiques.

Lors de la constitution de cette liste de référence, chacun des trois enjeux écologiques a été divisé en quatre catégories, correspondant à leurs aspects majeurs renvoyant à la fois à la chaîne de causalité propre aux problèmes environnementaux dont font état les documents de référence. Ces aspects sont les constats, les causes, les conséquences et les solutions. Pour assurer un suivi plus centré sur les thématiques précises de la transition écologique, les termes du dictionnaire classés comme "causes" ou comme "solutions" ont été, de plus, répartis selon des sous-catégories englobant ces thématiques. La méthode de catégorisation est détaillée plus bas.

L'adaptation au discours médiatique

L'un des objectifs du dictionnaire est de s'adapter au registre médiatique en incluant les formulations les plus utilisées par les médias pour aborder les enjeux écologiques. L'ensemble des termes a été passé en revue pour apprécier leurs contextes et fréquences d'usage.⁹ Il s'agissait de repérer les synonymes et les mots fréquemment employés par les médias autour des mots-clés de la liste de référence. Quand ceux-ci relevaient du champ lexical de l'environnement, ils ont été ajoutés au dictionnaire et répartis dans les catégories correspondantes.

Cette vérification manuelle a également permis de filtrer les termes hors-sujet susceptibles de générer des "faux positifs" : quand plus de 10% des occurrences d'un mot-clé ne concernait pas les enjeux écologiques, ce mot-clé était qualifié de "haut risque de faux positif" et faisait l'objet d'un traitement particulier, détaillé plus bas.

Lors d'une troisième phase d'exploration, de nouveaux mots et expressions récurrentes repérés dans les contenus informatifs dédiés à ces sujets ont également été ajoutés aux catégories correspondantes du dictionnaire. Partant cette fois des contenus déjà répertoriés grâce à l'ensemble des mots-clés existants, il s'agissait d'appliquer une démarche qualitative et d'observer les discours informationnels au niveau des programmes. Entre le 2 et le 4 juin 2024 une trentaine d'émissions d'information de la radio et de la télévision contenant l'un des mots-clés du dictionnaire ont été écoutées ou regardées dans leur intégralité afin d'identifier d'autres termes fréquemment employés¹⁰ par les médias et propres au champ lexical des enjeux environnementaux.

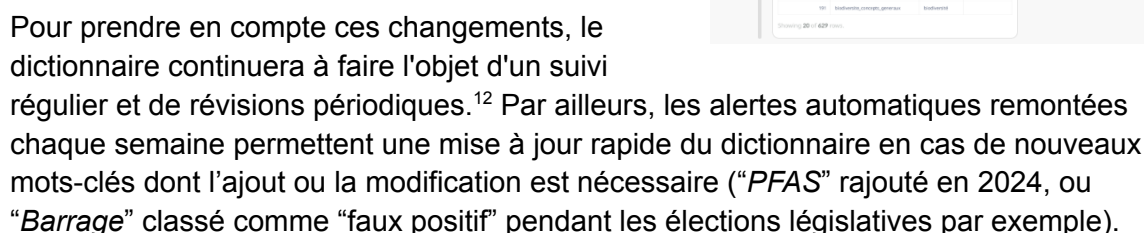
Enfin, des listes de mots ont été soumises par des rédactions des médias audiovisuels. Celles-ci sont issues de reportages que ces médias ont considérés comme représentatifs de leur traitement des enjeux écologiques. Leur pertinence a également été vérifiée.

⁹ Les recherches avec des mots-clés ont été réalisées sur la plateforme Keywords de Mediatree, société de veille médiatique et membre du consortium de l'Observatoire, sur l'ensemble des contenus d'information sur un an. Les mots relatifs aux enjeux écologiques ont été vérifiés dans leur contexte exact sur des dates et des contenus aléatoires.

¹⁰ Les nouveaux mots identifiés, ont ensuite été vérifiés via la plateforme Keywords de Mediatree, dans l'ensemble des contenus d'information sur un an. Les mots ayant au moins une cinquantaine d'occurrences ont été retenus dans le dictionnaire. Quand plus de 10% des occurrences d'un mot-clé retenu ne concernait pas les enjeux écologiques, ce mot-clé était catégorisé comme "haut risque de faux positif".

La liste de mots ainsi constituée a ensuite été l'objet d'un suivi permanent grâce aux données chiffrées issues des mesures sur les contenus audiovisuels du périmètre, c'est-à-dire dans plusieurs milliers d'heures de programmes d'information que l'Observatoire traite en continu. Les statistiques hebdomadaires programmées et indiquant le nombre d'occurrences pour les mots-clés les plus utilisés ont permis de vérifier les contextes d'usage pour ces mots. En effet, un nombre d'occurrence élevé peut être indicateur de mots hors-sujet, de la récurrence du mot en question dans des sujets environnementaux très divers ou encore de publicités (cf. méthode de filtrage plus bas). En fonction de ces contextes, la catégorisation du mot-clé pouvait alors être revue : il pouvait être classé comme "faux positif" ou repris dans une expression plus précise correspondant mieux à son contexte récurrent ou alors filtré quand utilisé par une campagne publicitaire.¹¹

Le registre lexical et idiomatique des médias se renouvelant en permanence, le dictionnaire doit permettre de suivre ces évolutions. Les associations thématiques et les risques de faux-positifs qui en découlent sont également à adapter à l'actualité. Enfin, l'évolution de l'état des connaissances scientifiques et celle des politiques publiques redéfinissent les cadres du débat public et des termes qu'il mobilise.



¹² En plus de ce suivi continu, des tests de visionnage sont prévus tous les 12 mois, afin de vérifier la pertinence de la classification opérée par le dictionnaire et compléter la liste avec de nouveaux mots clés, si nécessaire. Pour la description de ces tests, voir plus bas.

Les types de catégories

Le dictionnaire est structuré selon des axes thématiques. L'ensemble des mots-clés est catégorisé selon les aspects majeurs des enjeux environnementaux : leurs causes, constats, conséquences et solutions. De plus, les mots des catégories des causes et des solutions sont classés selon les grands terrains de la transition écologique et des types de ressources naturelles.

Les catégories des liens causaux

Les mots-clés du dictionnaire ont été classés selon quatre aspects saillants :

- la catégorie des constats regroupe les concepts et mots généraux désignant les faits du dérèglement ou de dégradation de l'environnement¹³ et ceux permettant de les décrire¹⁴, ainsi que les principaux acteurs institutionnels concernés¹⁵ ;
- la catégorie des causes englobe les termes relatifs aux raisons de ces phénomènes de dérèglement et de dégradation de l'environnement¹⁶ ;
- la catégorie des conséquences inclut les mots relatifs aux impacts directs ou indirects des phénomènes de dérèglement et de dégradation¹⁷ ;
- la catégorie des solutions comprend les mots désignant les leviers d'action au niveau structurel, collectif et individuel aux problèmes liés aux phénomènes de dérèglement et de dégradation de l'environnement¹⁸, avec une attention particulière dans le cadre du réchauffement climatique à la distinction entre les solutions d'atténuation quand celles-ci permettent de lutter contre les causes du dérèglement climatique et les solutions d'adaptation quand celles-ci sont dirigées pour faire face aux nouveaux risques naturels.¹⁹

À bien noter que pour les mots-clés évoquant la crise des ressources naturelles, ces catégories ont été simplifiées afin de traiter leurs aspects les plus saillants :

¹³ Par exemple : changement climatique, habitabilité de la planète, bilan carbone, effet de serre, pollution émise localement, cycle du carbone ou diversité du vivant, fertilité des sols, faune sauvage.

¹⁴ Par exemple : glacier, océan, mer, arbre, eau, etc. Ce sont des mots du vocabulaire quotidien faisant partie du discours journalistique décrivant ou expliquant les phénomènes. Ils sont catégorisés en tant que "haut risque à faux positif".

¹⁵ Par exemple : GIEC, ministère de la transition écologique, IPBES, Union internationale pour la conservation de la nature, etc.

¹⁶ Par exemple : gaz à effet de serre, polluant, artificialisation des sols, engrais de synthèse, déforestation, extractivisme, etc. D'autres mots du vocabulaire quotidien et étant classés en tant que "haut risque à faux positif" sont également associés aux causes, entre autres : OGM, intrants, carburant, charbon, forage pétrolier, gâchis alimentaire, etc.

¹⁷ Par exemple : urgence écologique, acidification des océans, érosion côtière, fonte des glaciers, canicule, sécheresse, etc. D'autres mots du vocabulaire quotidien et étant classés en tant que "haut risque à faux positif" sont également associés aux conséquences, entre autres : cancer, sécurité alimentaire, accès à l'eau, feu, inondation, ouragan, etc.

¹⁸ Par exemple : conservation de la nature, forêt mosaïque, gestion durable des terres cultivables, indice de réparabilité, sobriété, etc. D'autres mots du vocabulaire quotidien et étant classés en tant que "haut risque à faux positif" sont également associés aux solutions, entre autres : forêt, végétation, pêche durable, moins consommer, etc.

¹⁹ Par exemple, les termes comme isolation thermique, recyclage ou sortie du charbon renvoient aux solutions d'atténuation tandis que ceux comme désalinisation des eaux, prévention des inondations ou résistant à la sécheresse indiquent des solutions d'adaptation.

- Les catégories “constats”, “causes” et “conséquences” ont été regroupées en un seul axe “problèmes” ;
- La catégorie “solutions” est maintenue, comme pour les deux autres crises.

Par ailleurs, puisque l’Observatoire des médias sur l’écologie n’a pas pour vocation de qualifier les différents leviers d’action, ceux-ci sont listés dans le dictionnaire parce qu’ils font partie du débat public, sans tenir compte du consensus à son égard.

Cependant, il se donne comme objectif d’adopter une vision systémique, considérant les interdépendances entre ces problèmes, pour prendre la mesure de l’urgence écologique. Compte tenu des corrélations, de nombreux mots ont été affectés à plusieurs catégories. La sobriété est, par exemple, une solution pour les trois crises. La déforestation est une cause pour le dérèglement climatique, l’érosion de la biodiversité et la dégradation des ressources naturelles. Enfin, certains mots renvoyant à l’une des conséquences d’une crise peut devenir la cause d’une autre. Ainsi, l’acidification des océans est une conséquence de l’élévation de la concentration en CO₂ dans l’atmosphère, et c’est une cause de l’érosion de la biodiversité et de la dégradation des ressources halieutiques.²⁰

Neuf catégories thématiques

Dans le cas des crises du changement climatique et de la biodiversité, des sous-catégories supplémentaires ont été créées pour les secteurs d’activité concernés par la transition écologique afin de suivre leur traitement dans les médias.

Ce classement thématique s’appuie sur la répartition, en neuf volets, des indicateurs du Secrétariat général à la planification écologique²¹ :

- transports
- énergie
- bâtiment
- industrie
- agriculture
- écosystèmes
- eau
- économie circulaire
- “général” pour regrouper les notions plus larges, relatives à plusieurs secteurs.²²

La typologie des ressources naturelles

L’Observatoire s’intéresse également aux problèmes d’épuisement et de dégradation des ressources naturelles. Lors des premières phases exploratoires, il a été constaté que les cadrages du débat public tendaient à se cristalliser autour des types de ressources comme les minéraux, l’eau, le sol, l’air, la forêt et l’énergie. Plutôt qu’une thématization par secteurs d’activité, une classification selon leurs caractéristiques semble alors mieux s’adapter aux discours médiatiques. Les types de ressource sont :

²⁰ Pour prendre la mesure de cette transversalité, voir le tableau de bord en ligne “[Quel recouvrement entre les différentes crises ?](#)”.

²¹ <https://e.infogram.com/95444dec-4126-496a-85ce-6a017217c0a4?src=embed>

²² Pour objectiver la thématization des mots désignant les causes et les solutions des deux crises, quatre personnes de l’Observatoire ont classé les termes du dictionnaire selon ces neuf thèmes, sans concertation préalable. Le secteur d’activité choisi par le plus grand nombre a été retenu pour chaque cause et chaque solution.

- Air
- Eau
- Économie circulaire
- Énergie
- Forêts
- Métaux et minerais
- Sols
- Concepts généraux

Cette typologie est une adaptation des typologies officielles (comme celle de l'Ademe)²³ aux besoins de notre analyse. La catégorie "eau" regroupe les mots renvoyant aux ressources en eau douce, mais aussi ceux des ressources halieutiques. La catégorie "énergie" englobe les hydrocarbures mais aussi les énergies renouvelables. L'économie circulaire concerne principalement le recyclage, le réemploi et la revalorisation des ressources utilisées. La dernière catégorie regroupe les termes généraux tels que "épuisement des ressources", "limites planétaires" ou "consommation responsable".

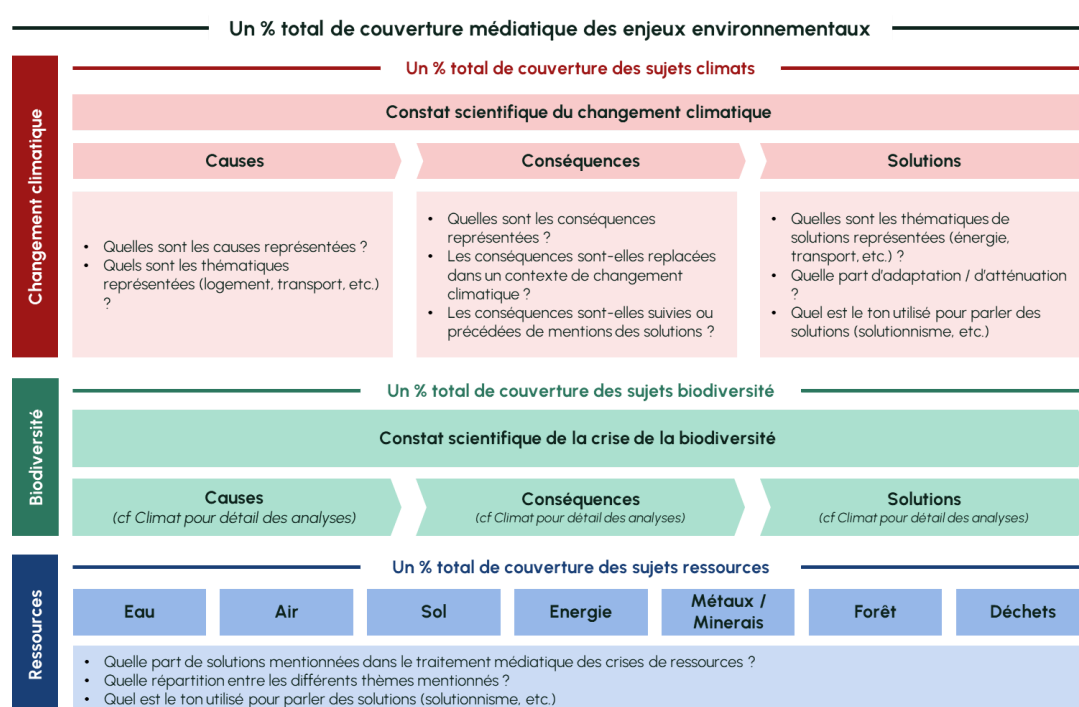


Fig.1 : Synthèse de l'architecture du dictionnaire et des principaux observables

Les mises à jour du dictionnaire

Les mises à jour du dictionnaire sont précédées par des tests de visionnage²⁴ des contenus médiatiques permettant d'identifier les sujets récurrents traités par les médias, ainsi que les mots et expressions pour en parler. Les mots nouveaux, représentatifs par leur fréquence aux discours des médias, sont intégrés au dictionnaire. Les mots, qui, par leur contexte dominant ne désignent plus exclusivement les enjeux écologiques, peuvent être reclassés

²³ <https://multimedia.ademe.fr/infographies/infographie-jour-depassement-ademe/>

²⁴ Des programmes choisis de façon représentative à l'ensemble du périmètre, sur des périodes aléatoires de l'année écoulée.

comme “haut risque à faux positif”. Par ailleurs, l’incertitude due à l’évolution lexicale sera intégrée explicitement au cours des années futures.

Un algorithme identifiant les mots-clés du dictionnaire

L’algorithme de l’Observatoire repère [les mots-clés du dictionnaire](#) dans les contenus médiatiques (pour l’heure les verbatims du flux audiovisuel et à terme les articles de la presse écrite).

Chaque mot-clé est cherché dans les retranscriptions avec sa variante régulière au pluriel (avec un “s” ou un “x”). L’outil n’est pas programmé pour les formes irrégulières, ni les formes conjuguées, ces dernières étant directement intégrées dans le dictionnaire.

Quand un mot-clé est employé dans des expressions, l’algorithme comptabilise la forme la plus longue. Ainsi, le terme “réfugié climatique” est reconnu comme tel et le mot “climatique”, également listé dans le dictionnaire, n’est pas compté séparément dans ce cas.

Par ailleurs, le code de l’Observatoire attribue une catégorie aux mots-clés suivant la catégorisation prévue par l’architecture du dictionnaire par enjeu environnemental, chaîne de causalité et secteur d’activité ou type de ressources.

Il prévoit [l’exclusion de certaines expressions du langage courant](#) contenant l’un des mots-clés, comme par exemple “en train de”, “parc d’attraction” ou “tombé à l’eau”.

Périmètre de l’Observatoire

L’Observatoire a fait le choix méthodologique de **mesurer les contenus d’information** (dans un premier temps au sein du traitement audiovisuel français, la presse écrite suivra dans une seconde phase) en délimitant les contenus analysés de façon précise, transparente et objective.

Ce choix de l’information est d’abord lié à l’objectif principal de l’Observatoire : suivre le traitement des enjeux écologiques dans l’actualité généraliste et politique.

Il s’explique aussi par des contraintes techniques et financières. Il fallait en effet que le périmètre soit circonscrit afin de respecter le budget dédié à l’achat de données par l’Observatoire. Il fallait aussi, techniquement, pouvoir délimiter les programmes suivis avec précision afin d’automatiser leur enregistrement dans la base de données de l’Observatoire. Ainsi, les programmes irréguliers, ponctuels, ne peuvent être inclus.

Concrètement, afin que l’analyse ne soit pas biaisée par l’importance relative accordée aux programmes d’information dans la grille des différents médias, l’Observatoire a choisi d’adopter la définition de “programmes d’information” issue de la classification de l’Arcom. Cela signifie que les documentaires, les programmes de fiction, les programmes de divertissement et d’*infotainment* en sont exclus.²⁵

²⁵ Pour découvrir l’ensemble des programmes du périmètre : <https://observatoiremediaecologie.fr/wp-content/uploads/2024/10/Doc-perimetre-site.pdf>

Les médias audiovisuels

Le flux audiovisuel analysé comprend les chaînes généralistes historiques et anciennes chaînes TNT, les chaînes d'information en continu ainsi que les radios généralistes publiques et privées à vocation nationale (catégorie E, selon la qualification des programmes audiovisuels définie par l'Arcom).²⁶ Plus précisément, l'Observatoire se concentre sur les contenus classés comme "journaux" et "magazines d'information" par l'Arcom et diffusés entre 6h et 23h. Il intègre les séquences météo directement adossées aux journaux télévisés du soir.

Pour tenir compte des différences entre chaînes généralistes et chaînes d'information en continu, l'Observatoire a mis en place deux méthodes distinctes pour ces deux catégories de chaînes télévisées. Seuls les temps d'information sont mesurés sur les chaînes généralistes quand, pour les chaînes d'information en continu, l'intégralité de la tranche 6h-23h est observée.

L'équipe de l'Observatoire est consciente que certaines émissions de télévision peuvent traiter d'environnement mais ne pas être incluses dans le périmètre analysé, dans la mesure où elles ne relèvent pas des catégories communes définies par l'Arcom. Ce choix méthodologique est motivé par la nécessité d'avoir des règles uniformes à tous les médias analysés.

Une mise à jour annuelle du périmètre aura lieu à chaque rentrée grâce à l'Arcom qui transmettra à l'Observatoire les modifications déclarées par les chaînes de radio et de télévision.²⁷

Source de données

La société Mediatree fournit les transcriptions textuelles du flux de contenu audiovisuel des programmes inclus dans le périmètre, via une API, accessible à l'Observatoire.

L'Observatoire dispose de ces données depuis mars 2023 et le volume de contenu augmente d'environ 1000 heures chaque semaine (dont plus de 720 h de télévision et 370 h de radio).

L'exclusion des publicités

Le flux de contenu cadré par ce périmètre inclut les publicités diffusées durant les programmes²⁸. Nombreuses sont celles qui contiennent l'un des mots-clés du dictionnaire : "covoiturage", "panneaux solaires", "énergie renouvelable", etc. sont des termes largement utilisés par les annonceurs. Pour éviter de les comptabiliser, le code de l'outil exclut [des formulations précises](#) des annonces où ces mots sont employés (par exemple : "covoiturage tous les moyens sont bons", "panneaux solaires groupe verlain", "l'huile de coude est aussi

²⁶ TF1, France 2, France 3 Île-de-France, Arte, M6, C8, BFMTV, Cnews, France 24, Franceinfo, LCI pour la télévision et France culture, France info, France inter, RFI, Europe 1, RMC, RTL, Sud radio pour la radio.

²⁷ Les changements majeurs éventuels parvenus au cours de l'année seront pris en compte.

²⁸ L'Observatoire exclura l'intégralité des séquences publicitaires à terme, à l'aide de l'intelligence artificielle.

une énergie renouvelable”).

Le contexte publicitaire de la majorité des mots-clés concernés a été repéré lors des tests du dictionnaire, ce qui a permis de les traiter en amont. Un suivi hebdomadaire automatisé des résultats permet de repérer les nouvelles campagnes en alertant de l’augmentation importante des occurrences d’un des mots-clés. Si l’augmentation repérée correspond effectivement aux publicités, les nouveaux énoncés sont exclus.

Le calcul du temps d’antenne

L’analyse des médias audiovisuels porte sur des **tranches horaires sélectionnées : celles de l’information** ([Voir les tranches horaires concernées](#)). C’est la somme de l’ensemble de ces tranches horaires qui sert de **période de référence**.

Les verbatims des programmes de cette période de référence sont subdivisés en blocs textuels correspondant à des séquences audio de 20 secondes, à l’intérieur desquels l’algorithme repère les mots-clés du dictionnaire.

Comptabilisation des extraits grâce aux mots-clés

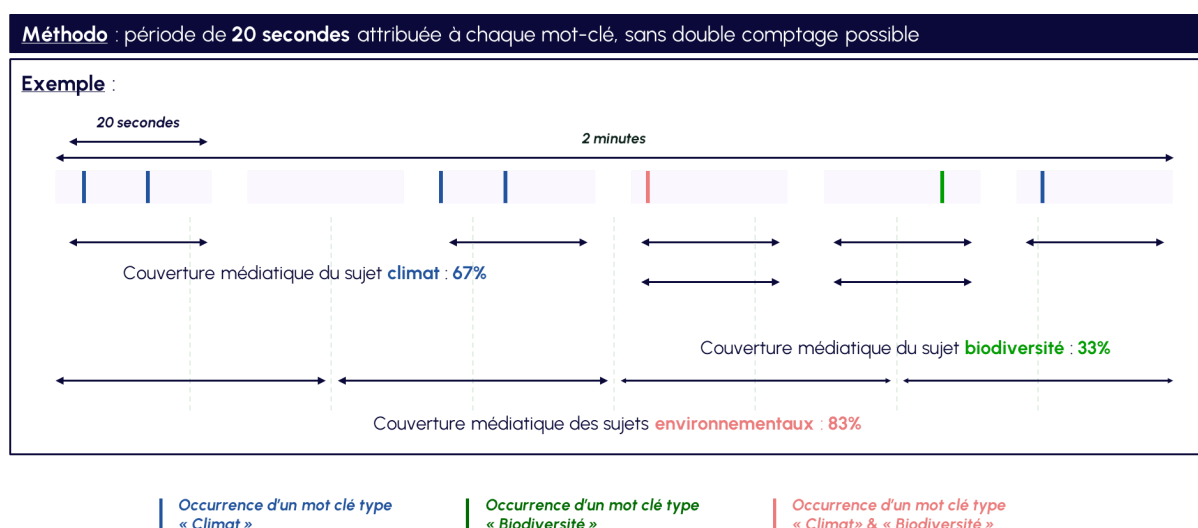


Fig.2 : Représentation schématique de la comptabilisation du temps d’antenne

En identifiant un ou des mots-clés à l’intérieur des blocs de 20 secondes, nous pouvons estimer la durée consacrée à **chaque thématique**. Les tronçons contenant plusieurs mots-clés ne sont comptés qu’une fois.

Nous comparons ensuite la durée consacrée à chaque thématique à **la durée totale de la période de référence pour obtenir un ratio** de temps d’antenne consacré au sujet, par chaîne et par radio.

*Par exemple - supposons que nous analysions 10 chaînes de radio, pendant les heures de grande écoute entre 6h30 et 9h30, sur une période d’une semaine. Si nous trouvons 5000 extraits de 20 secondes qui mentionnent un certain sujet, nous ferions le calcul : $5000 * 20s' = 1667$ minutes consacrées au sujet sur cette période.*

$3h * 7 \text{ jours} * 10 \text{ chaînes} = 12\,600 \text{ minutes de temps d'antenne total}$

$\text{Donc } 1667 / 12600 = 13 \% \text{ de temps d'antenne consacré à un sujet.}$

Représentativité des blocs de 20 secondes selon la distance entre mots-clés

La distance de 20 secondes entre deux mots-clés est une mesure applicable pour l'ensemble du contenu médiatique observé, quels que soient les types d'émissions.

Lors de l'exploration des données, il a été observé que la répartition des mots-clés dans les différents types de contenus médiatiques et sur différentes chaînes est restée relativement constante. En d'autres termes, les distributions des mots-clés ne varient pas beaucoup entre les différents types d'émissions et les différentes chaînes.

Distribution cumulée des distance entre deux mots-clés consécutifs (secondes) - 05/05 > 05/06

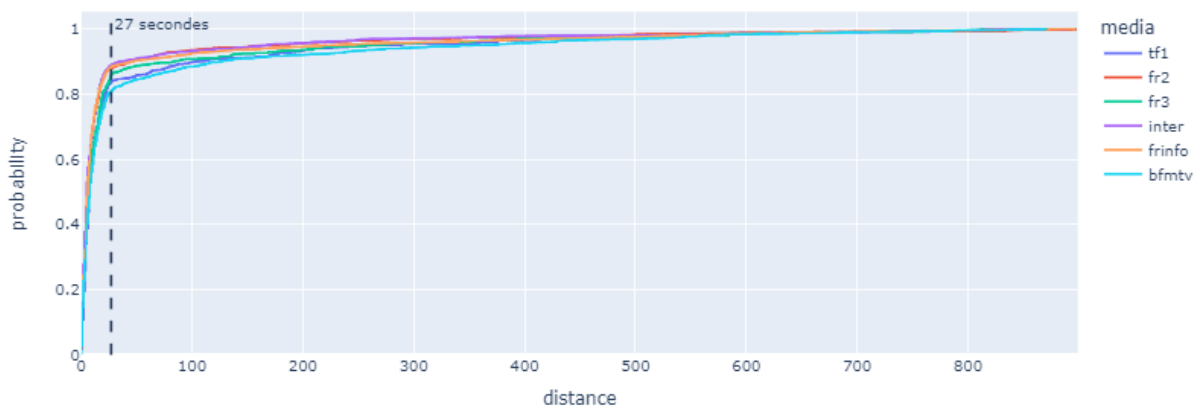


Fig. 3 : Dans le graphique ci-dessus, on observe la distribution cumulée des durées en secondes entre deux mots-clés successifs. On observe une rupture statistique nette autour de 27 secondes, correspondant à la distance temporelle entre deux mots clés successifs dans 80 à 90% des cas selon les chaînes.

Traitement des mots-clés à haut risque de faux positif

Lors des tests manuels, l'ensemble des mots dont plus de 10% des occurrences étaient hors sujet par rapport à la transition environnementale ont été qualifiés à "haut risque de faux positif". Les mots à haut risque de faux positif (par exemple "train") ne sont pris en compte **que s'ils sont énoncés autour d'un bloc de 20 secondes adjacent à des mots-clés sans risque de faux positif**, c'est-à-dire spécifiques au champ lexical des enjeux écologiques (par exemple "gaz à effet de serre") **ou à des mots-clés à "risque de faux positif" préalablement validés selon ce principe** (validation en cascade).

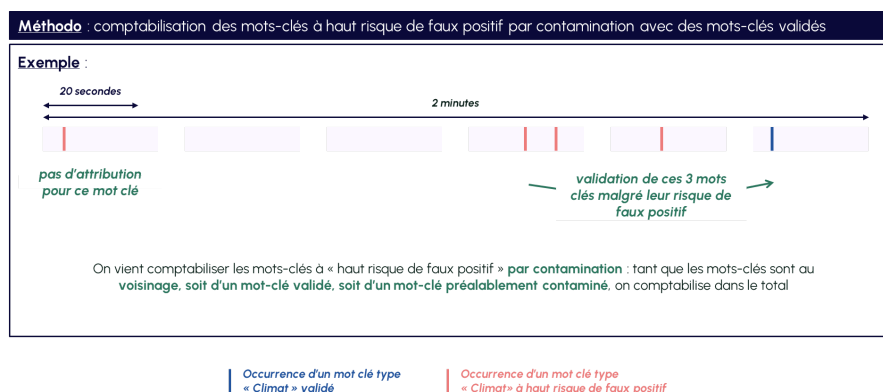


Fig 4 : Représentation schématique de la comptabilisation des mots-clés dit à “haut risque de faux positif”

En identifiant des extraits de **20 secondes** à partir de mots clés, il existe un risque de surestimer ou de sous-estimer le temps d'antenne réellement alloué au sujet.

Justifications, éléments d'attention et limites

Estimation préliminaire du risque de surestimation

À la télévision ou à la radio, une mention peut être faite sur une durée plus courte que la durée de 20 secondes que nous avons retenue (par exemple, un extrait de 5 secondes est comptabilisé comme un extrait de 20 secondes). L'Observatoire a examiné cette question en étudiant au hasard 100 extraits de 2 minutes issus d'une sélection aléatoire de médias, sur la période du 12 janvier au 19 janvier 2024 et contenant le mot-clé “réchauffement climatique”. **Résultat (voir Fig. 2) : cette analyse a montré que sur les 100 extraits, la durée potentielle de surestimation liée à ce créneau de 20 secondes était de l'ordre de 8%.**

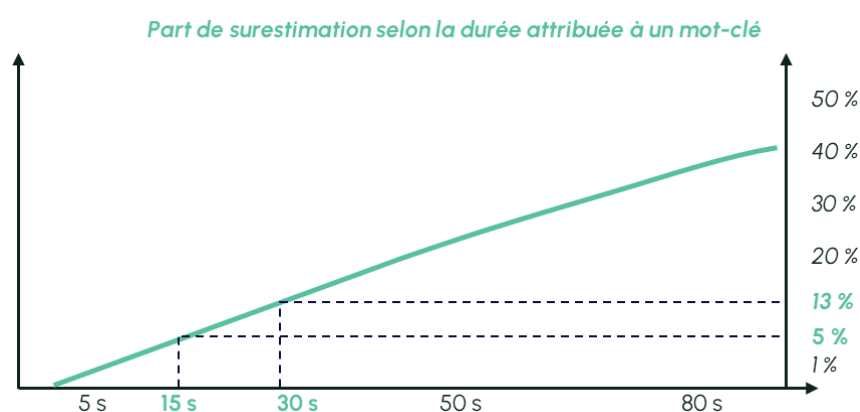


Fig. 5 : Part de sur-estimation selon la durée attribuée à un mot-clé, données mesurées et analysées sur la période du 12 janvier 2024 au 19 janvier 2024. Axe vertical : % de surestimation ; axe horizontal : durée attribuée autour d'un mot-clé.

Le traitement spécifique dédié aux mots-clés à haut risque de faux positif permet par ailleurs de limiter le risque de surestimation.

Enfin, les mots-clés du dictionnaire peuvent être employés par les médias dans des contextes négatifs ou de désinformation. Ce temps d'antenne est pourtant comptabilisé en tant que temps alloué aux enjeux environnementaux.²⁹

Risque de sous-estimation

La méthodologie d'analyse par mots clés est associée au risque de ne pas identifier certaines séquences parlant des enjeux écologiques mais ne contenant pas les mots clés du dictionnaire. Pour réduire ce risque, l'Observatoire repose sur un dictionnaire détaillé de plus de **1700** mots clés (voir pp. 3-7).

Certaines expressions du dictionnaire pourraient se situer à la limite de deux blocs de 20 secondes (les deux mots de "diversité biologique" pourraient se trouver ainsi séparés). À moins que l'un des mots de cette expression soit aussi un mot-clé du dictionnaire (comme c'est le cas de "écologique" dans "transition écologique"), celle-ci ne sera pas décomptée.

Tant pour le risque de surestimation que de sous-estimation, une analyse plus approfondie a été menée, notamment pour préciser la durée optimale des séquences contenant les mots-clés, voir : *Vision comparée des durées de 15 à 40 secondes*

Analyses comparées et tests de la méthodologie

Risque d'erreurs de transcription

Il existe également une marge d'erreur liée aux formes spécifiques des mots résultant de l'algorithme de transcription.

Celui-ci utilisant la reconnaissance vocale automatique convertit un contenu audio en texte. Cependant, les transcriptions par cette technologie ne sont pas parfaites. Le taux d'erreur sur les mots – ou Word Error Rate (WER) – mesure le pourcentage de transcriptions incorrectes de mots dans l'ensemble. Dans le cas des transcriptions utilisées par l'Observatoire le WER est de 10 %.³⁰

Lors des tests du dictionnaire les variations des mots-clés ont été listées et intégrées au dictionnaire quand celles-ci étaient régulières (par exemple "CO2" transcrit comme "co deux" ou "TGV" comme "tégévé"). Néanmoins, certains mots-clés complexes (les noms propres notamment) sont mal retranscrits et avec des variantes d'une occurrence à l'autre. Il peut donc manquer quelques séquences si le mot clé recherché est complexe. Cependant, ces mots-clés complexes ne représentent qu'une faible part de la couverture médiatique totale, induisant un risque de sous-estimation ou de surestimation très limité dans la part totale des contenus attribués aux sujets environnementaux, comme l'illustre le graphique ci-dessous (Fig. 3) représentant les mots-clés les plus importants.

²⁹ L'élaboration d'une méthode de filtrage de ces contenus est prévue pour la seconde phase du projet.

³⁰ Mediatree communique à ses clients toute modification importante dont la modification de l'algorithme de transcription.

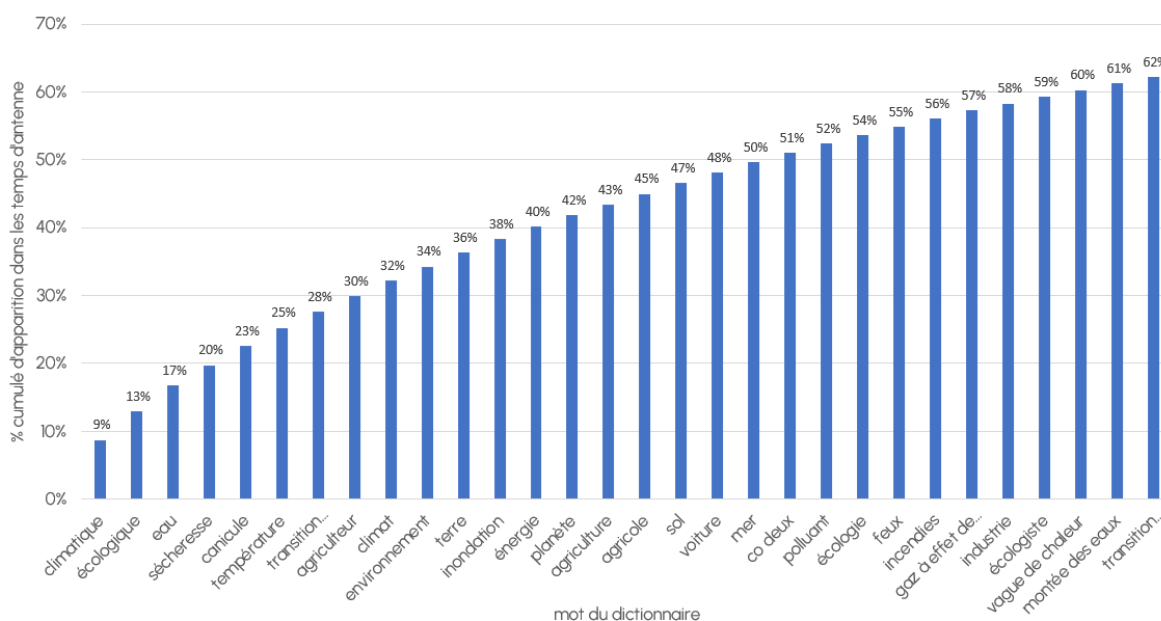


Fig. 6 : Le graphique ci-dessus, calculé sur la période 1 janvier 2024 - 4 novembre 2024 illustre que parmi les mots clés du dictionnaire, une trentaine correspondent à 62% du temps d'antenne comptabilisé pour les enjeux de changement climatique, rendant les calculs particulièrement robustes à d'éventuelles omissions. Note : certains de ces mots sont à "haut risque de faux positif" mais leur occurrence dans ce classement a bien été validée par la méthodologie correspondante.

Analyses comparées et tests de la méthodologie

Compte tenu des vérifications statistiques exploratoires (8 % de surestimation), la méthode de mesure par blocs de 20 secondes est efficace pour cibler avec précision les passages sur le sujet des crises écologiques. Toutefois, d'après les observations répétées, elle rencontre ses limites dans le cas des reportages longs, surtout quand ceux-ci emploient un vocabulaire peu spécifique aux sujets environnementaux.

Si l'extension des blocs pourrait paraître une solution pour contourner cet écueil, elle risque d'induire des biais dans les mesures.

Quelques illustrations des enjeux de la prolongation des blocs

Un allongement éventuel de la durée d'un bloc validé par la présence d'un mot clé augmente la durée attribuée, mais il affecte également la validation en cascade des mots-clés à "risque de faux positif". Dans le cas d'un reportage ou d'une interview où la description et l'explication vulgarisée mobilisent un type de vocabulaire courant, l'extension de la durée d'un bloc réduit le temps non mesuré entre deux mots-clés éloignés et permet d'étendre le champ d'attraction entre un mot-clé spécifique et les mots-clés plus courants mais attachés au sujet en contexte (les mots du dictionnaire à "risque de faux positif").

Pour mieux représenter l'un de ces cas de figure, nous pouvons prendre un exemple de la base de données. Le texte du reportage est divisé en blocs de 20 secondes. Les mots en gras en vert foncé sont des mots-clés spécifiques aux enjeux écologiques. Les mots en gras de couleur vert clair sont des mots-clés à “haut risque de faux positif”.

19:47:00 il gagne mille cinq cents euros par mois sur un marché qui pèse sept milliards en France 19:47:07 (début du reportage) Le futur Yellowstone européen pourrait bien se trouver en Roumanie au cœur des Carpates. Une fondation de **préservation de l'environnement** souhaite réintroduire dans les montagnes bison, ours ou encore le lynx. Un défi 19:47:20 pour protéger cette **forêt** parmi les dernières **forêts** primaires du continent. Mais ce retour du monde sauvage ne se fait pas sans heurts avec la population locale. Dans les Monts Fogarash, les plus hautes montagnes des Carpates du Sud, vivent des animaux disparus d'Europe à 19:47:40 l'état sauvage depuis plus d'un siècle. Certains animaux ont des colliers gps. Ces colliers émettent un signal toutes les seize heures environ. 19:48:00 Cet après-midi par exemple, le dernier message a indiqué un lieu tout proche d'ici. Ces animaux, Adrian Iida les connaît bien. Lui et son équipe les ont réintroduit progressivement ces dernières années. “Regardez. Là il y a des poils et 19:48:20 ça c'est le signe qu'ils sont passés par là.” Les bêtes semblent éviter le petit groupe signe de leur réadaptation à l'état sauvage, (19:48:30) mais l'équipe dispose de dizaines de caméras pour suivre l'évolution de la **faune** de la zone. Et lorsque les équipes de télévision ne sont 19:48:40 pas là pour les déranger, c'est un spectacle extraordinaire qui se joue dans cette clairière. Voici des bisons d'Europe à l'état sauvage. Très répandues sur le continent jusqu'au Moyen-Âge, l'espèce n'a survécu qu'en captivité. Adrien et son équipe 19:49:00 en ont déjà relâché cinquante-six. Une première génération de bisons arpente désormais ces **forêts**. Les bisons rejoignent ainsi les autres espèces qui peuplent le massif. 19:49:20 L'ours brun, lynx, loups et cervidés constituent l'un des réservoirs de **faune sauvage** les plus riches d'Europe. 19:49:40 L'objectif du projet Life est de créer une zone sauvage et celui de notre fondation est de créer le **parc** national des montagnes de Fagaras. Il s'agit alors d'un **parc** national qui 19:50:00 s'étend sur une très grande superficie près de deux cent mille hectares, donc quelque chose de très grande pour l'Europe”. Adrian travaille pour un organisme privé, la Fondation Conservation Carpathia. Depuis 2009 elle s'est donnée pour mission la création d'un **parc** national en plein cœur 19:50:20 du pays. Carpathia souhaite faire de ces montagnes situées à deux heures au nord de Bucarest un Yellowstone européen, inspiré du grand **parc** américain et pour cela elle **rachète** des parcelles de **forêt** à des particuliers. Déjà plus de vingt mille hectares qu'elle restaure et préserve. 19:50:40 Pourtant le travail de la fondation ne fait pas l'unanimité, car contrairement à son lointain cousin américain, le futur **parc** est entouré de zones d'habitation, des villages habités depuis des siècles. 19:51:00

Extrait d'un reportage du JT du soir de France 3 Île-de-France, diffusé le 3 décembre 2023³¹

Le surlignage en bleu indique les blocs de 20 secondes décomptés par l'Observatoire. Conformément à la méthode, les blocs de 20 secondes en beige ne sont pas comptabilisés car le bloc contenant le mot à “risque de faux positif” (“faune”) ne suit pas directement le bloc contenant le mot-clé déjà validé (“forêt”). Or, une subdivision du texte en blocs de 30 secondes aurait permis de prendre ce passage. Des blocs de 40 secondes auraient permis de prendre l'extrait.

Pour illustrer les répercussions éventuelles de l'extension avec des chiffres, prenons l'exemple du 19 juin 2024 : 2,51% du temps d'antenne global a été consacré aux sujets environnementaux à ce jour, avec des blocs d'une durée de 15 secondes. En élargissant la durée des blocs à 20 secondes ce pourcentage passe à 3,49 %, à 30 secondes il est à 5,43 % et à 40 secondes à 6,76 %. Cette augmentation mécanique ne signifie pas pour autant d'avoir mieux mesuré les contenus réellement consacrés à l'environnement.

En effet, l'extension de la durée des blocs augmente le risque de surestimation et de comptabilisation des extraits hors sujet. L'application de **la méthode à 30 ou à 40 secondes peut avoir comme répercussion de doubler le temps d'antenne considéré comme dédié aux enjeux écologiques sans contenu effectif.**

Ce serait concrètement le cas lors des simples mentions d'un des mots-clés spécifiques aux enjeux sans développement du sujet, de l'emploi d'un des mots-clés hors-sujet (comme l'illustre l'utilisation de “recyclage” dans le langage courant), des publicités non repérées ou des brèves ayant une durée de quelques secondes à une minute.

Pour concrétiser ces types de discours par un exemple illustratif : dans l'extrait suivant, issu

³¹ Afin de rendre la compréhension de ces extraits plus facile, leurs transcriptions ont été “nettoyées” : les majuscules et la ponctuation ont été rajoutées, les mots non reconnus par l'algorithme, dont essentiellement les noms propres, corrigés.

de la base de données, le texte du débat est divisé en blocs de 20 secondes. Les mots en gras en vert foncé sont des mots-clés spécifiques aux enjeux écologiques. Les mots en gras de couleur vert clair sont des mots-clés à “haut risque de faux positif”.

19:56:40 Ce sera le premier mai apparemment, je plaisante. Ce sera bien pour la Cgt. Oui pour le Rassemblement National aussi qui organise de grands événements chaque premier mai. Je plaisantais. Alors on peut remarquer que c' est assez comique et que ce gouvernement n'a pas d'autres idées que de **recycler** des vieilles lunes qui ont montré on l'a vu avec le lundi de Pentecôte, a eu un effet zéro. D'ailleurs le lundi 22:57:00 de Pentecôte est redevenu quelque part un jour férié. Nicolas il nous reste trois quatre minutes ensemble. Ce qui fait mal au cœur c'est que depuis trois mois euh qu'on nous vend ce budget qui était donc maintenant arrivé à l'Assemblée. Justice fiscale, justice fiscale, justice fiscale, justice fiscale mon oeil ce sont les classes moyennes 22:57:20 comme prévu qui vont passer à la caisse c'est comme ça et c'est pas autrement. Et le meilleur moyen de ne pas avoir ce genre de débat, ce serait peut-être déjà de penser à faire de la croissance dans ce pays. Ah mais (22:57:30) je suis très heureux de vous entendre le dire, parce qu'en fait quand on parle de croissance, on a l'impression maintenant d'employer un terme interdit notamment parce que certains considèrent que la croissance est quelque chose qui 22:57:40 aujourd'hui est contraire à l'environnement, à la **décarbonation**. Est-ce que quelque chose est fait pour la croissance selon vous ? Pas du tout, parce que je pense même que le choc fiscal qui est dans ce budget qui est de l'ordre de trente milliards d'euros c'est à dire l'équivalent de ce qui avait fait François Hollande et Jean-Marc Ayrault en deux mille treize est, au contraire, de **nature** à faire ralentir la croissance je pense qu'au bout du compte en fait on aura moins de recette fiscale de ce qu'on escompte justement parce que 22:58:00

Extrait d'un débat sur le plateau de CNews, diffusé le 28 octobre 2024³²

Les blocs de 20 secondes en bleu sont décomptés comme temps d'antenne parlant des enjeux environnementaux. Or, dans le premier passage le mot-clé “recycler” est hors contexte et dans le second les mots “environnement” et “décarbonation” sont des mentions. Dans ce type de cas, les blocs à 30 ou à 40 secondes augmenteraient le risque de surestimation du temps d'antenne consacré à ces sujets, tandis que la méthode à 20 secondes limite la marge d'erreur.

Pour mieux estimer quelle serait la récurrence de ces types de contexte, en plus des estimations statistiques sur l'ensemble de la base de données, nous avons également évalué les différentes méthodes de mesure automatique sur un échantillon, en les comparant aux décomptes effectués par un ensemble de bénévoles.

Comparaison avec les résultats d'analyses à la main

Afin d'estimer la durée la plus optimale des blocs, quatre méthodes – à 15, 20, 30 et 40 secondes – ont été comparées avec les résultats issus des mesures effectuées par les bénévoles³³.

Plusieurs centaines de programmes ont été sélectionnés entre le 10 juillet 2023 et le 26 juin 2024 de sorte de couvrir différentes périodes de l'année (différentes saisons, hors et durant des grands événements liés aux enjeux écologiques, hors et en période de campagne électorale) en incluant chaque chaîne et chaque émission du périmètre.³⁴ À partir de cette liste, une nouvelle sélection – visant un échantillonnage incluant des exemples de l'ensemble des contenus audiovisuels de notre base de données – a permis de constituer

³² Afin de rendre la compréhension de ces extraits plus facile, leurs transcriptions ont été “nettoyées” : les majuscules et la ponctuation ont été rajoutées, les mots non reconnus par l'algorithme, dont essentiellement les noms propres, corrigés.

³³ Les tests se sont déroulés entre le 15 juillet et le 17 septembre 2024.

³⁴ Pour des raisons de faisabilité, nous avons exclu les chaînes d'information en continu de l'échantillon.

une liste de programme d'une durée de 32 heures environ.³⁵ Les 31 émissions sélectionnées ont été regroupées en 8 bouquets de 4h approximativement, contenant différents types de contenus, de différentes dates. Chaque bouquet de 4h a été visionné par deux bénévoles sans concertation afin de mesurer les différences d'interprétation entre deux individus.

Le test a été réalisé via la plateforme Keywords de Mediatree.³⁶ Les bénévoles ont eu comme consigne de décompter à la seconde près les passages dans un programme donné parlant "d'environnement ou de la transition écologique au sens large".³⁷ Les durées ont été additionnées afin d'être comparées aux résultats de l'outil pour ces mêmes programmes.

69 % des répondants (13 personnes sur les 16 participants au test ont rempli le questionnaire diffusé) ont entre 24-29 ans, 92 % ont un diplôme du supérieur long. 46 % sont adhérents d'une association. Les répondants sont enclin à investir de leur temps pour une cause : 92 % indiquent avoir déjà fait du bénévolat (sans précision sur le domaine). Près de 31 % des répondants déclarent l'avoir fait plusieurs fois.

Les participants ont une importante familiarité avec les médias : 23 % y travaillent et 38 % les analysent (dans le cadre de leurs études, de leurs activités d'enseignant ou de leurs engagements citoyens). Ils consultent quotidiennement les médias (92%) et 36 % estiment même y consacrer plus de deux heures par jour. Ces pratiques suggèrent qu'ils connaissent les codes et le langage des médias.

Les bénévoles sont également familiers des sujets environnementaux, car 61 % des répondants consultent au moins hebdomadairement les informations sur ces enjeux à travers les médias spécialisés et pour 31 % cette pratique est quotidienne. Quant à leur niveau de connaissances (sur une échelle de 1 à 10), ils estiment maîtriser ces sujets sans en être experts (38 % s'attribue la note de 8, 31 % la note de 7 et 23 % la note de 6). Enfin, ils ne sont pas familiers du type d'exercice réalisé dans le cadre du test, car pour 11 personnes sur 13, c'était leur première expérience de veille médiatique sur les sujets environnementaux.

Encadré : Le profil des participants au test

Vision comparée des durées de 15 à 40 secondes

La moyenne des durées relevées par les deux bénévoles pour les 31 émissions a été comparée aux décomptes automatiques à 15, 20, 30 et 40 secondes.

³⁵ L'échantillon par dates : 3 août 2023, 20, 25 et 29 septembre 2023, 3 et 9 décembre 2023, 19 et 27 février 2024, 9, 14, 15, 25 et 30 avril 2024, 7, 8, 9, 24 et 26 juin 2024.

³⁶ Les participants ont eu leur propre accès, avec l'indication de la date et de l'heure du début des programmes qui leur ont été attribués.

³⁷ Les instructions, telles que formulées dans le tableur en ligne, rempli par les participants : "À l'aide du compteur de Mediatree, notez dans le tableur, à la seconde près, le début et la fin des passages du programme où vous pensez que l'on parle d'environnement ou de transition écologique au sens large. Relevez les minutes et secondes de manière précise. Il peut s'agir par exemple : des impacts du changement climatique, de l'érosion de la biodiversité et de la dégradation des ressources naturelles, ainsi que des moyens de l'enrayer, de l'importance d'espèces animales ou végétales pour l'humanité ou pour la biodiversité, des solutions économiques, industrielles, technologies, agricoles etc. pour protéger le climat, la biodiversité, et les ressources de la planète... Bref, tout ce qui vous fait penser aux défis environnementaux et à l'avenir de notre planète. Les publicités insérées dans certains programmes ne sont pas à prendre en compte. Si vous identifiez plusieurs passages (même d'une petite durée), notez le début et la fin de chaque passage. Si vous trouvez que l'ensemble d'une émission parle de ces sujets, vous pouvez la comptabiliser en entier".

Sur 116 220 secondes de contenus au total, les bénévoles ont relevé 9004 secondes, selon la somme des moyennes des secondes décomptées par deux individus, pour une émission. Si l'outil applique la méthode des blocs de 15 secondes, il décompte 5280 secondes. Avec des blocs de 20 secondes, ce relevé est à 6880 secondes. En allongeant les blocs à 30, puis à 40 secondes, les mesures passent respectivement à 9360 et 12280 secondes.

Par rapport à la moyenne des bénévoles, les méthodes à 15 et 40 secondes mesurent avec les décalages les plus importants : à la sous-estimation et à la surestimation.

Il convient donc de **se pencher sur les durées de 20 et 30 secondes plus en détail**, pour comprendre pour chaque séquence analysée les sources d'erreurs, et affiner le choix.

Des erreurs tirées à la hausse par les séquences longues

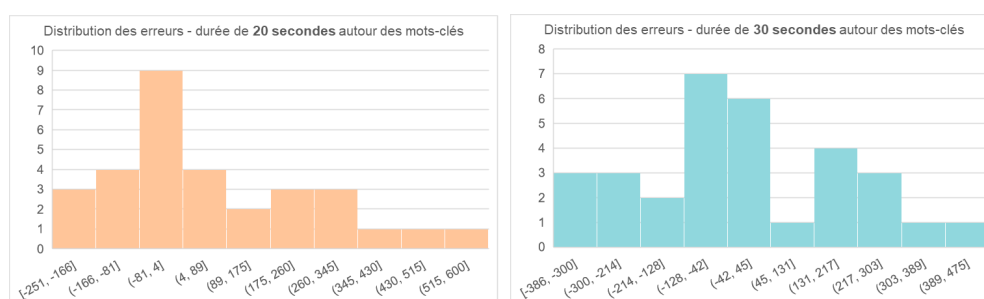


Fig. 7 : distribution comparée des erreurs [moyenne bénévole - machine] selon les différents choix de durées pour 20 et 30 secondes. Une valeur faible (segment [-386 : -128] par exemple) indique des cas de surestimation supérieur à 2 minutes par l'algorithme.

L'étude de ces distributions (Figure 7) montre notamment que les deux distributions suivent toutes deux, en application du test de Shapiro-Wilk, une loi normale.³⁸ En revanche, l'estimation réalisée avec 30 secondes possède une variance 12% supérieure à celle avec 20 secondes. Lorsque l'on retire les 10% des séquences avec le plus d'erreurs, pour réduire l'effet des très longues séquences non captées par l'algorithme, cet écart de variance monte à 36%.

Variance	Paramètre à 20 secondes	Paramètre à 30 secondes
100% des échantillons	1.3 e6	1.4 e6 (+12%)
90% des échantillons sans les valeurs extrêmes	6.2 e5	8.9 e5 (+36%)

Tableau 1 : comparaison entre paramétrage à 20 et 30 secondes, pour respectivement 100% et 90% des échantillons

L'estimation avec une durée de 30 secondes semble avoir une **plus grande tendance à la compensation** : n'estimant mieux les résultats que dans leur ensemble, mais en

³⁸ seuil de rejet de H_0 fixé à 0.05

compensant les sous-estimations des faux positifs par une plus grande surestimation des autres programmes.

Notamment, en mettant de côté les **10% des séquences avec le plus d'erreurs** (en pratique seulement des faux négatifs), l'analyse de moyennes des erreurs, donne les résultats suivants :

- Moyenne avec une durée fixée à 20 secondes : **23 secondes de sous-estimation moyenne**
- Moyenne avec une durée fixée à 30 secondes : **54 secondes de surestimation moyenne**

Par ailleurs, ces 10% de séquences avec le plus d'erreurs s'avèrent être, pour 50% d'entre elles, les séquences avec le plus d'erreurs, quelle que soit la durée choisie : **modifier la durée attribuée n'est à ce stade pas un élément de résolution pour mieux capter ces séquences.**

Compte tenu de ces éléments, choisir une durée algorithmique de 30 secondes ne servirait qu'à compenser une éventuelle sous-estimation par ailleurs, mais collerait moins à la réalité.

Il convient donc de **choisir une durée de 20 secondes**, puis de s'intéresser au traitement fait par l'algorithme des séquences plus longues.

Des séquences longues ne faisant pas plus consensus chez les bénévoles

Pour aller plus loin, regardons les mesures réalisées par plusieurs bénévoles pour la même séquence. Si l'on déploie les comparaisons non pas par rapport à la moyenne des durées retenues par deux bénévoles pour une émission, mais **avec les valeurs prises individuellement** – en doublant donc chaque émission et par conséquent la durée totale – en moyenne l'écart auquel on peut s'attendre est de **5,8 % à 20 secondes et 6,4 % à 30 secondes.**

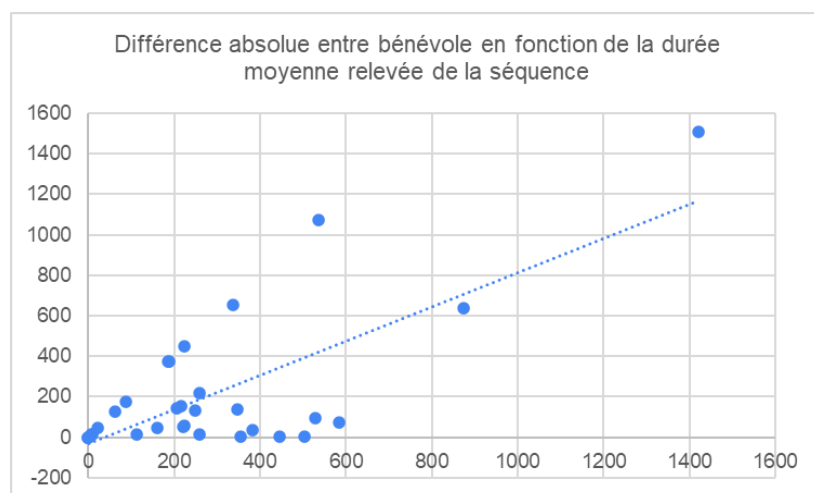


Fig. 8 : représentation de la différence entre bénévoles entre eux (en secondes) en fonction de la durée moyenne relevée par les bénévoles (en secondes).

Par ailleurs, l'analyse des erreurs montre que **la durée des émissions de l'échantillon affecte les écarts entre les relevés des bénévoles et ceux de l'outil** (voir Figure 8). L'augmentation de la durée d'un programme augmente mécaniquement la variance entre les bénévoles eux-mêmes.

Enfin, la différence absolue entre deux bénévoles sur l'ensemble des émissions est plus importante que celle entre les moyennes des bénévoles et l'algorithme paramétré à 20 secondes, et la variance l'est également (voir infra). Ceci indique que **les différences d'interprétation pourraient être plus importantes entre deux personnes que les différences de mesure entre un ensemble d'individus et l'outil**.

	Comparaison bénévoles vs. algorithme	Comparaison entre bénévoles
Ecart moyen	157 secondes	214 secondes (+36%)
Ecart type	148 secondes	342 secondes (+132%)

Tableau 2 : comparaison entre les bénévoles entre eux et les bénévoles par rapport à l'algorithme calibré avec une durée de 20 secondes.

Compte tenu de ces éléments, la **durée retenue reste bien de 20 secondes, avec une attention particulière portée à l'intégration de reportages aujourd'hui non décomptés par l'outil**, et par ailleurs décomptés de façon non uniforme entre les bénévoles.

Analyse de la mise en contexte des conséquences de la crise climatique et de l'effondrement de la biodiversité

De plus en plus intenses et de plus en plus fréquents, les effets du dérèglement climatique et, dans une moindre mesure³⁹, ceux de l'érosion de la biodiversité rythment l'actualité. Les explications sur l'origine de ces phénomènes – les lier aux faits et/ou aux causes des dérèglements – favorisent la compréhension des enjeux par le public. L'Observatoire mesure donc le taux de "mise en contexte" des conséquences de ces deux crises.

Dans le cadre des analyses, la "mise en contexte" s'observe lorsqu'un mot-clé relatif aux conséquences (de la crise climatique ou de biodiversité) est mentionné en co-présence d'un mot clé relatif aux constats ou aux causes de ces crises, dans les 5 minutes autour de la conséquence. Par exemple, dans un reportage sur les inondations récentes, s'il est également précisé que ces phénomènes sont amenés à être de plus en plus fréquents en raison du changement climatique, alors la conséquence est bien mise en contexte.

³⁹ [Voir les données de l'Observatoire sur la comparaison de ces deux enjeux.](#)

Choix méthodologiques de l'analyse de mise en contexte

En analysant les données, nous observons que 50% des liens (entre une conséquence et un constat / une cause) sont établis en 5 minutes ou moins (Fig. 9). Donc, un mot clé de constat ou de cause a été prononcé 5 minutes avant ou après un mot clé de conséquence. Par ailleurs, cette durée semble correspondre à la durée moyenne d'un reportage de journal télévisé.

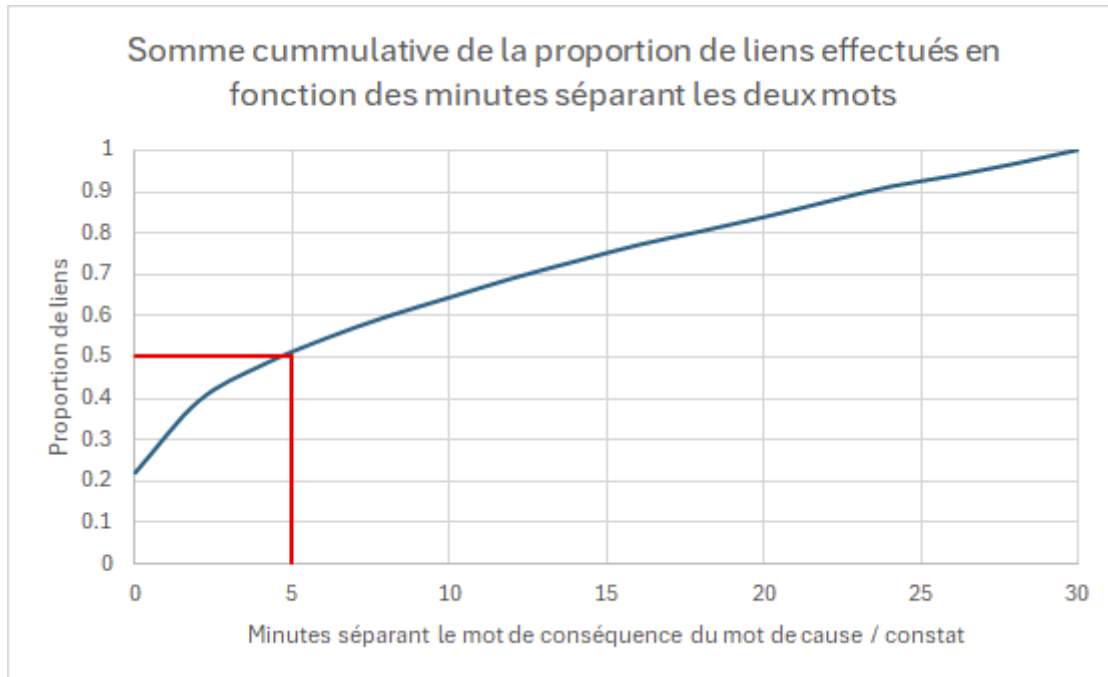


Fig. 9 : Distribution cummulative du temps séparant un lien entre une conséquence et une cause / constat (Analyse sur tout le mois de Sept. 2024)

De plus, des analyses précises de mentions de cooccurrences ont permis d'identifier que certains mots-clés à "hauts risques de faux-positif" de la catégorie "constats" accompagnaient le plus souvent les mots relatifs aux conséquences. Or, ces mots du quotidien facilitant la description des phénomènes, ne les indiquent pas avec précision. Par exemple, dans le cas des conséquences de la crise climatique, la co-occurrence avec le mot "eau" (mot de constat à "haut risque de faux-positif") ne permet pas d'indiquer avec précision une mise en contexte de cette conséquence avec la crise climatique (Fig. 10). Les mots à "hauts risques de faux-positif" ne sont donc pas pris en compte dans l'analyse de la mise en contexte.

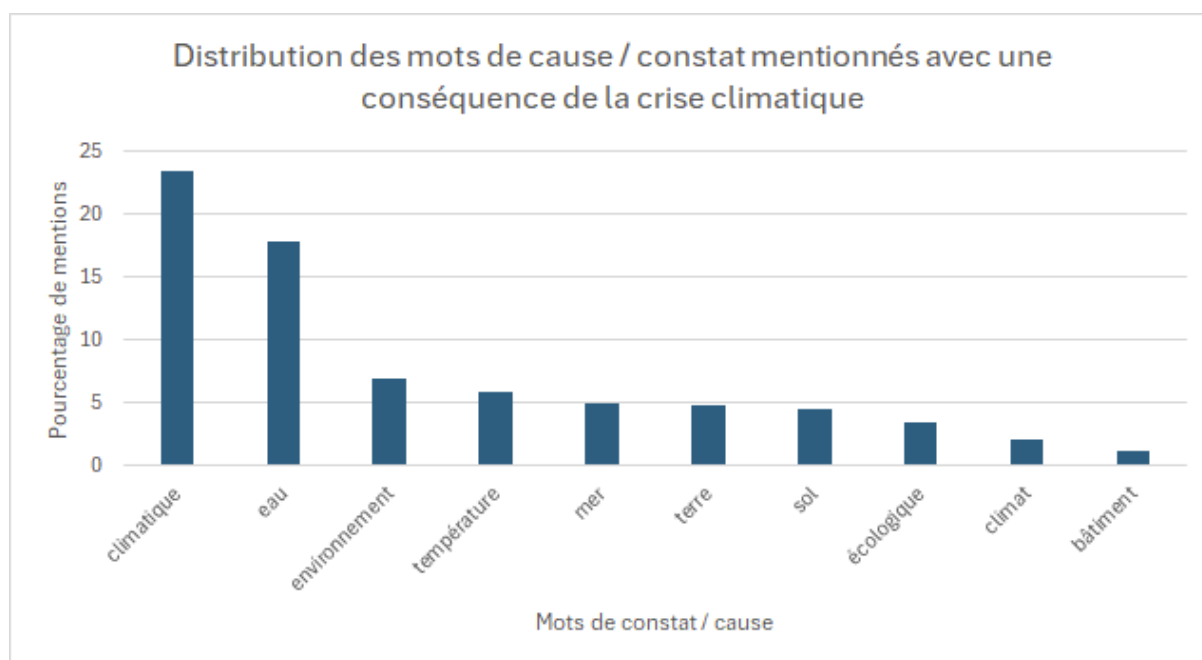
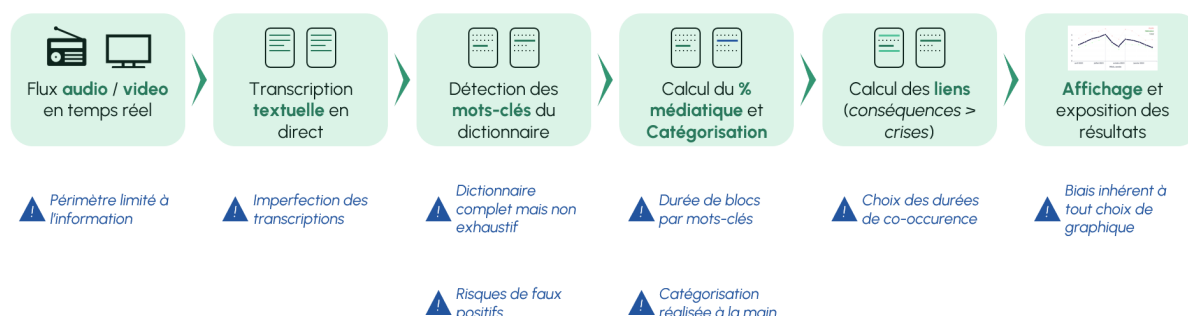


Fig. 10 : Co-occurrences des mots de constat / cause lorsqu'un lien avec une conséquence de la crise climatique est établie (Analyse sur le périmètre de France 2 en sept. 2024)

Conclusion

L'Observatoire des médias sur l'écologie est un outil de partage de connaissances conçu pour les citoyens, les médias, les politiques publiques et plus généralement tous les acteurs intéressés par la transition écologique. Il est créé grâce à des ressources institutionnelles, entrepreneuriales et associatives, ainsi que des compétences multiples émanant des différents partenaires participants afin de répondre à l'urgence écologique et dynamiser la résilience.

Ses résultats, issus de données en masse et des analyses quantitatives, montrent des tendances, au niveau général. Toutefois, sa méthodologie, bien que testée, comme toute méthodologie présente évidemment des limites.



Récapitulatif des points méthodologiques avec leurs limites

Grâce aux soutiens institutionnels déjà confirmés, une seconde phase se peaufine. Celle-ci permettra en premier lieu de continuer les tests et d'affiner les méthodes.

Durant cette seconde phase, et pour répondre à la multiplicité des contenus abordant les enjeux écologiques, le périmètre de l'Observatoire s'élargira. De nouveaux genres d'émission audiovisuelles s'ajouteront aux programmes d'information déjà traités. Par ailleurs, une quantification plus avancée de l'incertitude sur les résultats pourra être menée.

La presse écrite, en papier et en ligne, fera également partie des médias suivis.

Toujours dans le signe de la coopération avec la société civile et les médias, l'Observatoire pourra mieux appréhender le traitement médiatique des enjeux écologiques.

Références

Aday (2023). 10 ans de traitement médiatique du climat et des événements météo extrêmes, publié le 17 octobre 2023.

[https://www.aday.fr/news/10-ans-de-traitement-mediatique-du-climat-et-des-evenements-met-
eo-extremes](https://www.aday.fr/news/10-ans-de-traitement-mediatique-du-climat-et-des-evenements-met-
eo-extremes)

Boykoff M. & Luedecke G. (2016). *Environment and the Media*, Oxford Research Encyclopedia of Climate Science, Oxford University Press USA.

Browning P., Ficher S., Flinn C. et al. (2021). *People's Climate Vote*, Rapport PNUD & Oxford University.

<https://www.undp.org/sites/g/files/zskgke326/files/publications/UNDP-Oxford-Peoples-Climate-Vote-Results.pdf>

Comby J-B. (2015). *La question climatique. Genèse et dépolitisation d'un problème public*, Paris, Raisons d'Agir, 250 p.

Comby J-B. & Romanet V. (2012). Un problème réchauffé ? Les changements climatiques dans les médias généralistes 2007-2011, Rapport de recherche, Institut Français de Presse de l'Université Paris 2 & GIS « Climat, environnement et société », Paris.

http://www.gisclimat.fr/sites/default/files/Rapport_GISClimat_ifp_2012_0.pdf

Cevipof, Fondation Jean Jaurès, Ipsos/Sopra Steria pour Le Monde (2022). *Fractures françaises 2022*, 10ème édition.

https://www.jean-jaures.org/wp-content/uploads/2022/10/Fractures_francaises_vague10.pdf

Cevipof, Fondation Jean Jaurès, Institut Montaigne & Ipsos/Sopra Steria pour Le Monde (2023). *Fractures françaises 2023*, 11ème édition.

https://www.jean-jaures.org/wp-content/uploads/2023/09/Fractures_francaises_2023.pdf

Cordonier L. (2022). *Information et engagement climatique*. Étude de la Fondation Descartes.

https://www.fondationdescartes.org/wp-content/uploads/2022/11/Information_et_engagement_climatique.pdf

Lefort V., Poels G., Quinton F. (2020). En 20 ans, l'environnement est devenu un sujet d'information générale, Étude INA, *La revue des médias*, novembre 2020.

<https://larevuedesmedias.ina.fr/series/etude-environnement-information-medias-tv-radio-climat-pollution-biodiversite-energie>

IPBES, 2019 : Summary for policymakers of the global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. E. S. Brondizio, S. Díaz, J. Settele et al. (eds.), IPBES secretariat, Bonn, Allemagne, 56 p.

https://files.ipbes.net/ipbes-web-prod-public-files/2020-02/ipbes_global_assessment_report_summary_for_policymakers_fr.pdf

IPCC, 2021: Summary for Policymakers. In: *Climate Change 2021: The Physical Science Basis*. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, Masson-Delmotte, V., Zhai P., Pirani A. et al. (eds.), Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 3–32.

IPCC, 2023: Summary for Policymakers. In: Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, H. Lee and J. Romero (eds.), IPCC, Geneva, Switzerland, pp. 1-34.

https://www.ipcc.ch/report/ar6/syr/downloads/report/IPCC_AR6_SYR_SPM.pdf

Kantar Public pour La Croix (2023). La confiance des français dans les médias, Résultats de la 37ème édition du baromètre La Croix / Kantar Public, novembre 2023.

<https://fr.themedialeader.com/wp-content/uploads/2023/11/Barometre-Medias-La-Croix-Kantar-Public-novembre-2023.pdf>

SAP & Qualtrics XM (2020), *Toward a More Sustainable World. A global study of public opinion*, Étude pour le World Economic Forum présentée en 2020.

https://www3.weforum.org/docs/WEF_More_Sustainable_World.pdf

Baromètres

Baromètre de l’Affaire du siècle :

<https://laffairedu siecle.net/wp-content/uploads/2022/04/20220406-CP-Barom%C3%A8tre-climat.docx.pdf>

Baromètre de la Planification écologique :

<https://e.infogram.com/95444dec-4126-496a-85ce-6a017217c0a4?src=embed>

Baromètre Data for Good & Quotaclimat :

<https://x.com/QuotaClimat/status/1588495361475170305>

Observatoire des JT de Climat Médias :

<https://observatoire.climatmedias.org/>