



HAL
open science

Favoriser la visibilité du patrimoine cinématographique par les algorithmes. Une expérimentation de recommandation de contenu à contenu appliquée au catalogue de la plateforme Tënk

Ève Givois, Eric Kergosien, Laure Bolka-Tabary, Samuel Gantier, Bernard
Jacquemin

► To cite this version:

Ève Givois, Eric Kergosien, Laure Bolka-Tabary, Samuel Gantier, Bernard Jacquemin. Favoriser la visibilité du patrimoine cinématographique par les algorithmes. Une expérimentation de recommandation de contenu à contenu appliquée au catalogue de la plateforme Tënk. Guèvremont, Véronique; Brin, Colette. Intelligence artificielle, culture et médias, Presses de l'Université Laval, pp.213-241, 2024, 978-2-7637-5877-0. hal-04504627

HAL Id: hal-04504627

<https://hal.science/hal-04504627v1>

Submitted on 14 Jun 2024

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

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

Favoriser la visibilité du patrimoine cinématographique par les algorithmes. Une expérimentation de recommandation de contenu à contenu appliquée au catalogue de la plateforme Tënk*

Ève Givois¹, Éric Kergosien², Laure Bolka-Tabary², Samuel Gantier¹
et Bernard Jacquemin²

¹Université polytechnique des Hauts-de-France, LARSH, F-59300 Valenciennes, France.

{Prenom.Nom}@uphf.fr

²Univ. Lille, ULR 4073 – GERiCO, F-59000 Lille, France.

{Prenom.Nom}@univ-lille.fr

Résumé

Comment rendre visible le patrimoine cinématographique? Face à l'offre pléthorique et exponentielle des plateformes généralistes, les acteurs traditionnels de la valorisation du patrimoine cinématographique – cinémathèques, médiathèques, mais aussi distributeurs, programmeurs et même cinéphiles – sont confrontés à cet enjeu. Dans ce contexte, les acteurs de la médiation du cinéma patrimonial ne pourraient-ils pas bénéficier, à leur tour, de ces moyens pour toucher un public plus large? Les technologies

*Givois, È., Kergosien, É., Bolka-Tabary, L., Gantier, S., & Jacquemin, B. (2024). Favoriser la visibilité du patrimoine cinématographique par les algorithmes. Une expérimentation de recommandation de contenu à contenu appliquée au catalogue de la plateforme Tënk. In V. Guèvremont & C. Brin (Éds.), *Intelligence artificielle, culture et médias* (p. 213-241). Presses de l'Université Laval.

numériques dites intelligentes pourraient-elles améliorer la visibilité du patrimoine cinématographique ? En particulier, les dispositifs de recommandation algorithmique peuvent-ils aider à multiplier et optimiser les rencontres entre le public et les films de fonds, collections ou catalogues de films patrimoniaux ? Pour explorer l'hypothèse d'une rencontre fructueuse entre les intermédiaires de la culture cinématographique et la pratique de la recommandation algorithmique de contenu à contenu pour l'amélioration de la visibilité du patrimoine cinématographique, nous nous appuyons sur un cas particulier : celui de la plateforme de vidéo à la demande (VAD) et vidéo à la demande par abonnement (VADA) coopérative Tënk. La plateforme, qui propose une sélection variée de films documentaires patrimoniaux et plus récemment réalisés, a constitué le terrain d'une recherche expérimentale menée en 2020-2021, dans le cadre du projet AlgoDoc.

Mots-clefs : recommandation de contenu à contenu, patrimoine numérique, film documentaire, patrimoine cinématographique, algorithme, intelligence artificielle.

Abstract

How can film heritage be made visible? Faced with the plethora and exponential growth of generalist platforms, the traditional players in the promotion of film heritage – cinémathèques, media libraries, but also distributors, programmers and even film buffs – are faced with this challenge. Against this backdrop, couldn't those involved in the promotion of heritage cinema also benefit from these resources to reach a wider audience? Could so-called intelligent digital technologies improve the visibility of film heritage? In particular, can algorithmic recommendation systems help to increase and optimise encounters between audiences and films from heritage collections or catalogues? To explore the hypothesis of a fruitful encounter between film culture intermediaries and the practice of algorithmic content-to-content recommendation to improve the visibility of film heritage, we will use a specific case: that of the cooperative video-on-demand (VOD) and subscription video-on-demand (SVOD) platform Tënk. The platform, which offers a varied selection of heritage and more recently made documentary films, was the site of experimental research carried out in 2020-2021 as part of the AlgoDoc project.

Keywords: content-to-content recommendation, digital heritage, documentary film, film heritage, algorithm, artificial intelligence.

Comment rendre visible le patrimoine cinématographique ? Face à l'offre pléthorique et exponentielle des plateformes généralistes, les acteurs traditionnels

de la valorisation du patrimoine cinématographique – cinémathèques, médiathèques, mais aussi distributeurs, programmeurs et même cinéphiles – sont confrontés à cet enjeu. Dans ce contexte, les acteurs de la médiation du cinéma patrimonial ne pourraient-ils pas bénéficier, à leur tour, de ces moyens pour toucher un public plus large ? Les technologies numériques dites intelligentes pourraient-elles améliorer la visibilité du patrimoine cinématographique ? En particulier, les dispositifs de recommandation algorithmique peuvent-ils aider à multiplier et optimiser les rencontres entre le public et les films de fonds, collections ou catalogues de films patrimoniaux¹ ?

Pour explorer l'hypothèse d'une rencontre fructueuse entre les intermédiaires de la culture cinématographique et la pratique de la recommandation algorithmique de contenu à contenu pour l'amélioration de la visibilité du patrimoine cinématographique, nous nous appuyons sur un cas particulier : celui de la plateforme de vidéo à la demande (VAD) et vidéo à la demande par abonnement (VADA) coopérative Tënk². La plateforme, qui propose une sélection variée de films documentaires patrimoniaux et plus récemment réalisés, a constitué le terrain d'une recherche expérimentale menée en 2020-2021, dans le cadre du projet AlgoDoc³.

Nous commencerons par mettre en évidence les enjeux de visibilité du cinéma patrimonial sur le Web, pris entre une multiplicité d'acteurs et de pratiques de médiation, en nous concentrant sur l'espace français. La deuxième partie présentera les enjeux du déploiement d'un algorithme de recommandation dans un dispositif préexistant de médiation du cinéma documentaire, la plateforme Tënk. Il s'agira d'étudier le passage d'une recommandation éditoriale à une recommandation algorithmique, en relevant les contraintes de cette expérimentation. Dans une troisième partie, le processus d'expérimentation d'un système de recommandation algorithmique de films documentaires de contenu à contenu sera

1. Le Centre national de la cinématographie (CNC) français définit les films patrimoniaux par l'ancienneté de leur première date de sortie en salles, supérieure à dix ans. Nous entendons plus largement dans ce chapitre les films de patrimoine comme l'ensemble des films pris dans un processus de patrimonialisation impliquant l'affirmation de leur appartenance à l'histoire du cinéma, du fait de leurs qualités ou de la position de leurs auteurs dans le champ cinématographique.

2. Tënk, <https://on-tenk.com/>. Créée en 2016, dédiée au cinéma documentaire d'auteur, elle est implantée en Europe francophone (France, Belgique, Suisse, Luxembourg), au Canada et en cours de déploiement dans l'Union européenne et au Royaume-Uni.

3. AlgoDoc (Algorithme de recommandation de films documentaires) est un programme financé en 2020 par l'I-Site et la Satt Nord. Il réunit un consortium recherche-industrie composé des laboratoires GERiiCO de l'Université de Lille et DeVisu de l'Université Polytechnique Hauts-de-France, ainsi que Tënk et Spideo, société spécialisée dans la recommandation algorithmique de contenu audiovisuel. Spideo, <https://spideo.com/>.

développé. Les métadonnées rassemblées et fabriquées pour expérimenter ce système seront mobilisées. La dernière partie examinera les limites de la démarche de recommandation automatisée pour faciliter la rencontre entre le patrimoine cinématographique et le public.

1 Enjeux de la visibilité du cinéma patrimonial sur le Web

1.1 Mutations des acteurs et pratiques de médiation du cinéma patrimonial

L'activité de médiation du cinéma patrimonial dans l'industrie cinématographique a connu plusieurs mutations concomitantes ces dernières années en France.

« L'ivresse patrimoniale » du cinéma, qui s'amorce à la fin du xx^e siècle avec l'intégration du cinéma au sein du patrimoine national et international (Lindberg & Szczepaska, 2017, p. 13), connaît une nette accélération au début du xxi^e. L'incroyable visibilité actuelle du cinéma patrimonial (Laurent, 2020b), en France et dans le monde, a d'abord été portée par les acteurs traditionnels de la médiation des films sortis du circuit commercial – cinémathèques, ciné-clubs, centres d'archives, mais aussi festivals dédiés. La dynamique a été accentuée par l'émergence et l'affirmation d'un marché du cinéma de patrimoine auquel contribuent producteurs et distributeurs (qui trouvent là un moyen d'exploiter leur capital), exploitants et diffuseurs (salles de cinéma, télévision et éditeurs de vidéo notamment) ou nouveaux festivals lui étant spécifiquement consacrés (Laurent, 2020a).

En parallèle, la numérisation et, dans une plus large mesure, le développement des technologies de l'information et de la communication (TIC) ont renforcé ce phénomène sur la durée. En premier lieu, les TIC contribuent à la mutation des pratiques traditionnelles de médiation du cinéma de patrimoine. C'est le cas notamment des festivals dédiés au cinéma patrimonial (Taillibert, 2015), du secteur traditionnel de la médiation du cinéma que constitue l'éducation à l'image (Taillibert, 2021), comme de la médiation des œuvres audiovisuelles éditées dans les bibliothèques et en médiathèque. Par ailleurs, les TIC font également émerger de nouveaux acteurs redéfinissant le secteur de l'industrie cinématographique et les enjeux patrimoniaux du cinéma. La médiation numérique des œuvres cinématographiques a ainsi connu une expansion continue depuis l'orée du xx^e siècle

(Bullich, 2012). Les plateformes VAD/VADA, comprises à la fois comme un ensemble d'acteurs sociaux nouvellement constitués et comme outil technologique, composent désormais un intermédiaire central de distribution et diffusion du patrimoine cinématographique. Elles sont nombreuses dans ce nouveau paysage. Y dominent un petit nombre d'acteurs majoritaires internationaux généralistes, dont l'hégémonie est sortie renforcée de la crise sanitaire amorcée en 2020 et des confinements successifs. S'y maintiennent également plusieurs acteurs minoritaires plus spécialisés (Direction des études, des statistiques et de la prospective, 2021, p. 45). Dans l'ensemble de ces plateformes, le cinéma patrimonial constitue une part non négligeable des catalogues de films mis en ligne (Danard, 2020). Les plateformes cinéphiliques⁴ contribuent à cette dynamique, mais, plus encore, participent du processus de patrimonialisation des œuvres qu'elles mettent à l'affiche. En mettant en avant l'aspect diachronique de leur programmation, en choisissant de mettre en avant un ensemble rétrospectif d'œuvres (Taillibert, 2020), elles contribuent à les consacrer.

Ces plateformes cinéphiliques s'inscrivent dans la logique de l'économie de la longue traîne Benghozi et Paris (2014) en ce qu'elles proposent à leurs utilisateurs des œuvres *a priori* peu rentables, qui, en étant diffusées en ligne, trouvent l'occasion d'être vues par un public autrement dispersé dans le monde physique. En ce sens, la dématérialisation de la médiation du cinéma patrimonial a pu être vue comme l'occasion de toucher un public plus large et dispersé, voire de réaliser l'utopie de déconcentration des moyens de diffusion – ainsi, tout un chacun peut émettre, produire son propre média. Le phénomène charrie toutefois également son lot de menaces : ne soulève-t-il pas le risque d'une déshumanisation de la médiation, voire d'une perte de lien entre spectateurs ? Le risque est d'autant plus grand que la plupart de ces plateformes, du moins les plus populaires, dans le sillage de Netflix, mettent en avant leur utilisation d'algorithmes fondés sur l'expérience et les traces d'usages pour assurer à chaque utilisateur l'accès à une offre personnalisée.

4. Christel Taillibert a établi, en 2020, une liste de 17 plateformes cinéphiliques : Arte Créative et Boutique Arte, Benshi Studio, Bref Cinéma, Bretagne & diversité, CiNéMa Club, UPOPI et Ciel (Ciclic), Le Kinéscope, KuB, My French Film Festival, Nowave, Tènk, La Toile, Universciné, ainsi que Mémoire (Ciclic) et La Cinetek, dédiées spécifiquement au cinéma patrimonial (Taillibert, 2020, p. 140-141).

1.2 Usage des algorithmes à des fins de recommandation

Les organisations proposant des biens culturels sur des plateformes disposant des dispositifs de recommandation algorithmique ont tendance à accentuer la visibilité des contenus les plus populaires (Dessinges & Perticoz, 2019) et à maintenir leurs utilisateurs dans des « bulles de filtres » hermétiques (Pariser, 2011), ou à proposer des contenus moins convaincants que les explorations autonomes de catalogue ou les prescriptions éditoriales (Beuscart, Coavoux, & Maillard, 2019). De fait, ce type de recommandation est souvent opposé à une prescription éditoriale exigeante et éclairée : la première émanerait tendanciellement de plateformes généralistes et commerciales, tandis que la seconde serait plutôt mise en œuvre par des plateformes alternatives (Benghozi & Paris, 2016 ; Thuillas & Wiart, 2019a, 2019b). En outre, ce type de recommandation algorithmique prédictive personnalisée peut susciter la méfiance, par son opacité (Cardon, 2015) et sa normativité Doueihy (2013) ; Vayre (2018). Dans ce contexte, les acteurs de la valorisation du patrimoine cinématographique seraient plus enclins à mettre de côté les dispositifs de recommandation algorithmique, qui ne répondent pas à leur besoin de rendre visibles des films en fonction de leur valeur patrimoniale ou de leurs qualités, toutes deux mesurées à l'aune d'un regard expert, et non de leur popularité ou des usages de leurs publics.

Pourtant, si les algorithmes basés sur les traces d'usage sont les plus répandus à des fins de personnalisation de la recommandation, d'autres peuvent être mobilisés. Les algorithmes de contenu à contenu, notamment, utilisent de leur côté uniquement les métadonnées⁵ associées en amont aux objets de la recommandation. Ils s'appuient sur l'ensemble des données techniques (supports, format, etc.), juridiques (visa, titulaire des droits, etc.), descriptives (titre, auteur, mais aussi genre, résumé, etc.) (Delcroix, 2019), plus ou moins élaborées, qui leur sont soumises, pour calculer la similarité entre les contenus et proposer une recommandation. En cela, ils diffèrent des algorithmes nourris par des données d'usage (profil et traces du comportement des utilisateurs, que ces derniers soient conscients de leur production – avis, notation – ou non – nombre de clics, parcours de navigation, etc.) (Farchy & Denis, 2020, p. 42). En évitant de recourir à une personnalisation de la recommandation, l'usage d'algorithmes de recommandation de contenu à contenu tend à limiter l'entretien des utilisateurs dans des bulles de filtres forgées à l'image de l'antériorité de leur navigation. En outre, l'explicitation de la recommandation est facilitée par l'affichage simultané de la

5. Les métadonnées peuvent être définies comme les « informations décrivant un contenu [...] précis selon des caractéristiques objectives ou plus ou moins construites », selon Farchy & Denis, 2020, p. 42.

source du calcul algorithmique (le contenu de départ) et de l'objet de la recommandation (le contenu prescrit).

Or, les acteurs de la valorisation du patrimoine cinématographique, en tant qu'intermédiaires de la culture cinématographique (Jeanpierre & Roueff, 2014), disposent bien souvent des connaissances, compétences et infrastructures indispensables pour élaborer des métadonnées forgeant pertinemment la recommandation algorithmique. Ainsi les professionnels des médiathèques, archives ou cinémathèques peuvent-ils mettre leur culture professionnelle de la documentation et de la prescription des biens culturels au service de l'enrichissement des métadonnées, quand celles-ci ne sont pas déjà disponibles au sein de leurs établissements. De leur côté, les distributeurs ou programmeurs qui travaillent au sein de plateformes VAD/VADA cinéphiliques, forts de leur expérience d'éditorialisation, ont l'expertise nécessaire pour qualifier pertinemment les films patrimoniaux qu'ils mettent en ligne, et participer à leur indexation. La convergence de ces différentes pratiques et savoir-faire ne permettrait-elle pas d'élaborer des outils algorithmiques plus humains, utiles et accessibles à chacun de ces acteurs ?

2 Un algorithme de recommandation humain ?

Pour imaginer un tel dispositif de recommandation, revenons d'abord sur la fonction qui lui serait attribuée, profondément humaine : il s'agit de reproduire le mécanisme de recommandation tel qu'il est pratiqué par les professionnels de la médiation du cinéma, c'est-à-dire de mise en liens des films entre eux. La création de ces liens est l'enjeu de toute pratique éditoriale mise en œuvre par les différents acteurs de la médiation cinématographique. Ils peuvent se nouer dès le moment de la création des films ou de leur acquisition par les producteurs, distributeurs et diffuseurs. À l'autre bout de la chaîne, ils sont aussi tissés par les professionnels de la médiation audiovisuelle que sont les conservateurs, médiathécaires et documentalistes audiovisuels lors de la constitution de leurs collections ou de l'acquisition de leurs fonds. Ainsi, Jean-Yves de Lépinay, président de l'association Images en bibliothèque⁶, considère que la mission des médiathèques est « de

6. Créée en 1989, elle a pour mission « d'affiner la culture cinématographique des bibliothécaires, mais aussi de faciliter le repérage et l'accès aux films de tous les genres » (Du Val, 2017). Elle structure depuis un réseau de structures adhérentes pour promouvoir la médiation du cinéma en bibliothèque, ainsi que la formation des professionnels et l'éducation à l'image. Voir Images en bibliothèques, « À propos », <https://imagesenbibliotheques.fr/a-propos>.

créer non seulement avec les usagers, mais aussi des liens entre les œuvres⁷. Les plateformes VAD/VADA elles aussi, par leur programmation, contribuent à forger des liens entre les films.

2.1 De l'éditorialisation à la recommandation algorithmique : le cas de Tënk

Parmi l'ensemble des plateformes VAD/VADA actuellement en ligne, nous nous intéresserons tout particulièrement ici à Tënk, plateforme proposant à ses abonnés un catalogue de films documentaires d'auteurs, dont une part de films patrimoniaux : 36 % des 2200 titres programmés depuis les débuts de la plateforme en 2016 étaient sortis depuis plus de dix ans au moment de leur première mise en ligne. Le projet à l'origine de Tënk, avant que ses créateurs décident de la création d'une plateforme, était de constituer une télévision dont le flux serait entièrement dédié au cinéma documentaire (Lesauvier, 2019). Il s'agissait ainsi pour les documentaristes, programmeurs du festival des États généraux du film documentaire et autres passionnés du genre se rassemblant en coopérative, de donner au cinéma documentaire d'auteur son autonomie envers les chaînes de télévision pour programmer, c'est-à-dire sélectionner et faire des liens entre films documentaires. Aujourd'hui, la plateforme propose à ses abonnés une sélection de sept films par semaine en ligne pendant deux mois, ainsi qu'un large catalogue d'œuvres en VAD. Son activité de médiation passe par une programmation incarnée, préparée par un réseau de bénévoles sur la base de leur culture individuelle et subjective du cinéma documentaire, validée par l'équipe de direction artistique. Le cœur d'activité de la plateforme est ainsi l'activité d'éditorialisation, soit un processus de choix et valorisation de contenus provoqué par leur numérisation (Bachimont, 2007 ; Epron & Vitali-Rosati, 2018). Pour cela, l'équipe de Tënk mobilise un savoir-faire s'inscrivant dans la continuité des activités de programmation événementielle et de distribution traditionnelle du cinéma. Après cinq années d'existence, le nombre d'œuvres ayant déjà été diffusées par la plateforme a considérablement augmenté : « Le catalogue est passé de 7 films par

7. Jean-Yves de Lépinay (président d'Images en bibliothèque), « Quelle place et quelles missions pour les bibliothèques ? Appréhender les différentes offres et l'évolution des pratiques en médiathèque », *VàD et ressources numériques*, formation professionnelle en ligne organisée par l'association Images en bibliothèque à destination de bibliothécaires et médiathécaires, 29 avril 2021.

semaine à environ 700 films en [VAD]. Aujourd'hui, permettre aux utilisateurs de Tènk de s'y retrouver face à une offre de 700 films est primordial »⁸.

Sans que le catalogue soit démesuré, ce changement d'échelle encourage l'équipe à mettre en place trois types d'actions de recommandation : poursuivre la recommandation éditoriale, en continuant à solliciter des programmeurs bénévoles pour proposer des associations de films originales ; introduire de la recommandation sociale, en permettant aux abonnés de partager leurs coups de cœur auprès de leurs contacts ; et enfin développer une recommandation algorithmique (Gantier, 2020). Le déploiement d'un algorithme est ainsi pensé comme une façon de pallier les limites de la mémoire humaine des programmeurs de la plateforme et de concevoir une mémoire collective de la programmation. Il nécessite toutefois de mettre en œuvre un processus d'objectivation et de formalisation des critères de sélection éditoriale par l'indexation systématique du catalogue. Ceux-ci sont multiples, mais ne se limitent pas aux auteurs, aux périodes ou lieux de production, ou encore aux thématiques abordées. Ils s'appuient bien souvent sur les procédés sociotechniques mobilisés par les auteurs de ces films, leurs partis pris de réalisation, autrement dit leurs dispositifs cinématographiques.

2.2 Enjeux éthiques de la recommandation algorithmique

Cette expérimentation d'un dispositif de recommandation algorithmique n'est pas sans poser de questions éthiques, liées à la fois aux contraintes qui pèsent sur les différents partenaires concernés, et à la question plus large de l'usage d'algorithmes pour proposer des recommandations de biens culturels.

Tènk envisage de déployer un tel dispositif à la condition qu'il réponde à sa ligne éditoriale. La plateforme affiche ainsi la volonté de continuer à se positionner en opposition avec les acteurs majoritaires de la VAD, en adoptant une position conforme à ses valeurs coopératives. Elle tient à inscrire la recommandation algorithmique dans la continuité de ses pratiques de programmations collectives, subjectives et incarnées. En cohérence avec cette position, l'algorithme de contenu à contenu semble être adapté, puisqu'il permet à l'équipe de programmeurs de garder la main sur la fabrication des données nourrissant la recommandation.

8. Entretien avec Pierre Mathéus et Samuel Gantier, mené par Némésis Srouf (2021), « Can Algorithms be Ethical? Behind Machines, a Human-Centered Approach », *Rumo*, consultable en ligne : <https://rumo.co/can-algorithms-be-ethical-behind-machines-a-human-centered-approach/>, traduction des auteurs.

La technologie permet de prendre en considération des aspects subjectifs, humains, intelligents, de la recommandation, dès lors que ceux-ci sont décrits dans les métadonnées.

De son côté, la société Spideo, qui met à disposition du projet une API dédiée à la recommandation, Rumo, affiche une volonté d'améliorer significativement la transparence de son dispositif. L'enjeu pour l'entreprise, dont l'activité était jusqu'à peu concentrée sur l'analyse sémantique de contenus et la prestation de services documentaires (Farchy et Denis 2020, p. 168 ; Gantier, Givois, et Kergosien 2020), est d'investir le marché de la recommandation sans proposition de création de nouvelles métadonnées. Sa participation au projet lui permet à la fois de tester sa nouvelle API et d'accélérer son développement. Elle participe également plus largement d'une stratégie d'amélioration de la compréhension des mécanismes de recommandation algorithmique, dans l'idée de s'attirer la confiance de ses clients actuels et futurs dans ses outils.

Ce souci de transparence, aux implications commerciales dans le cas de l'entreprise Spideo, constitue toutefois, de manière plus générale, un enjeu de taille de l'éthique des algorithmes, car l'« opacité peut être vue comme un facteur favorisant les manquements ou les risques éthiques » (Adary & Domenget, 2017 ; Simonnot, 2018). À ce titre, l'utilisation d'un algorithme de recommandation de contenu à contenu semble encore s'imposer pour accroître la transparence dans la mesure où il facilite l'explicitation de l'opération computationnelle, en plus d'éviter l'intrusion dans les données personnelles des utilisateurs.

Préférer recourir à des algorithmes de recommandation centrés sur le contenu, plutôt qu'à d'autres mobilisant les traces d'usage des utilisateurs, présente l'avantage de ne pas s'affranchir d'une réflexion sur les modalités de fabrication de la donnée. Celle-ci n'est pas appréhendée comme la trace d'un « réel » qui lui préexiste (Bonenfant, Crémier, & Lafrance St-Martin, 2018), mais comme un produit construit socialement, en fonction de besoins préalablement identifiés dans l'optique de représenter une dimension spécifique des œuvres afin de produire un effet bien particulier. Les métadonnées associées aux films à recommander sont construites sciemment, à dessein de recommandation, et non pas récoltées, sorties de leur contexte de production, malgré les usagers qui les produisent – comme le sont les traces d'usage. Contrairement à ces dernières, dont la production n'est pas socialement considérée comme un travail (Tiffon, 2017), les métadonnées dédiées à la recommandation de contenu à contenu sont élaborées par des travailleurs rémunérés à cette fin. Les œuvres sont documentées par leurs médiateurs traditionnels (et non par des contributeurs bénévoles d'une indexation

folksonomique⁹) qui mettent ainsi à profit leurs compétences professionnelles – loin de l’activité de micro-travailleurs du clic (Casilli, 2019 ; Le Ludec, Wahal, Casilli, & Tubaro, 2020).

3 Expérimentation d’une recommandation automatisée

Pour éprouver la mise en œuvre d’un tel algorithme de recommandation de contenu à contenu et, en particulier, son humanité, nous avons donc eu recours à l’expérimentation. Il s’agissait notamment de déterminer à quelles conditions l’utilisation d’un algorithme de recommandation peut pertinemment faciliter la découvrabilité par le public du patrimoine cinématographique documentaire.

3.1 « Mettre en données » les formes des films documentaires pour mieux les recommander

La première étape de notre expérimentation a consisté à objectiver les critères de rapprochement éditorial entre films mis en œuvre par l’équipe de Tënk, afin de transposer la démarche de recommandation humaine dans un outil algorithmique. En d’autres termes, pour se voir recommandées au moyen d’un algorithme, les œuvres de cinéma documentaire doivent, à leur tour, être passées au crible d’une « mise en données » (Bullich & Clavier, 2018 ; Flichy, 2021 ; Menger, 2017 ; Vayre, 2017).

Chez Tënk, les rapprochements entre films sont établis par les programmeurs moins en fonction des thématiques ou du genre (catégorie sans grande signification lorsqu’on l’applique au seul cinéma documentaire, voir Kilborne, 2008) que sur des critères plus formels, davantage liés à la démarche d’auteurs et aux dispositifs choisis pour réaliser leurs films. La création d’un outil dédié à l’indexation de cet aspect des films documentaires a permis d’envisager la mise en place

9. Le terme a été forgé par la combinaison de *folk* (usagers) et *taxonomy* (règles de classification) par Thomas Vander Wal (2004) pour « désigner le phénomène d’indexation des documents numériques par l’usager » (Le Deuff, 2006).

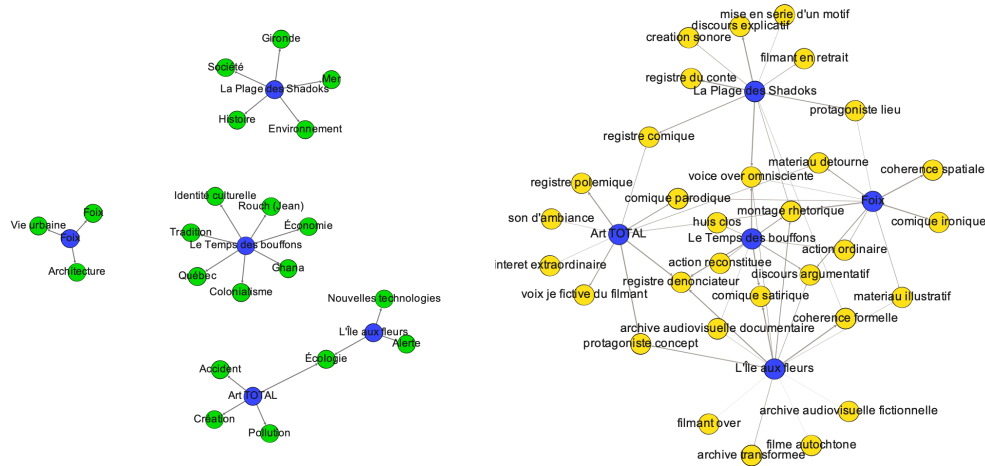
d'une recommandation automatisée se basant sur les aspects formels et socio-techniques, et non plus thématiques¹⁰. La collecte d'une connaissance experte de la variété de dispositifs mobilisés pour la réalisation de films documentaires, menée auprès des équipes éditoriales de Tènk, de chercheurs en études cinématographiques et au moyen de la littérature spécialisée, a donné naissance à un thésaurus des dispositifs de réalisation du cinéma documentaire (Givois, Bolka-Tabary, Kergosien, Gantier, & El Haoudi, 2021). Ce vocabulaire contrôlé et structuré, dont les termes peuvent être liés entre eux par des relations sémantiques d'équivalence (synonymie ou association) ou hiérarchiques (spécificité, généralité), a été mobilisé pour caractériser un ensemble de 177 films programmés par Tènk depuis sa création.

Au terme de cette démarche, il apparaît que les rapprochements entre films selon les dispositifs de réalisation du cinéma documentaire, indexés systématiquement sur un corpus restreint, s'avèrent plus riches que les rapprochements faits selon des métadonnées thématiques préexistantes (Gantier et al., 2020). La figure 1 permet de comparer les liens entre cinq films à la tonalité satirique et comique qui pourraient justifier de les voir recommandés entre eux : *Foix* (Luc Moullet, 1994), *La Plage des Shadoks* (Pierre Andrieux et Fabrice Marache, 2016), *Le Temps des bouffons* (Pierre Falardeau, 1985), *Art TOTAL* (Pierre Excoffier et Gwenn Pacotte, 2000) et *L'Île aux fleurs* (Jorge Furtado, 1989). Leur indexation thématique crée un seul lien entre deux d'entre eux : *Art TOTAL* et *L'Île aux fleurs*, qui traitent de sujets écologiques. Les trois autres sont complètement disjoints. À l'inverse, l'indexation des dispositifs de réalisation multiplie les liens entre ces cinq films : ainsi, quatre d'entre eux sont montés autour d'un discours (montage rhétorique), trois ont une voix hors champ omnisciente, trois autres relèvent d'un discours argumentatif, trois autres encore d'un registre dénonciateur. Ils ont en commun également un registre comique, que celui-ci soit parodique, ironique ou satirique.

Ces premiers résultats permettent de confirmer l'hypothèse de la pertinence d'une indexation alternative à l'indexation thématique, dès lors que celle-ci est menée à partir d'un vocabulaire construit en fonction des particularités du catalogue de Tènk. En établissant des liens inexistantes jusqu'à lors entre films, nous pouvons déduire que le jeu de métadonnées ainsi établi rend possible l'amélioration de la visibilité des films proposés dans le catalogue de Tènk.

10. Voir la VI du « Thésaurus des procédés filmiques du cinéma documentaires », publié sur OpenTheso, consultable en ligne <https://opentheso.huma-num.fr/opentheso/?idt=th228>.

Figure 1 – Comparaison des liens entre films produits par l’indexation thématique (à gauche) et l’indexation des dispositifs de réalisation (à droite). Graphes générés avec Gephi, ForceAtlas2 en mode LinLog avec déchevauchement et dissuadement des concentrateurs (*hubs*) puis ajustement des labels

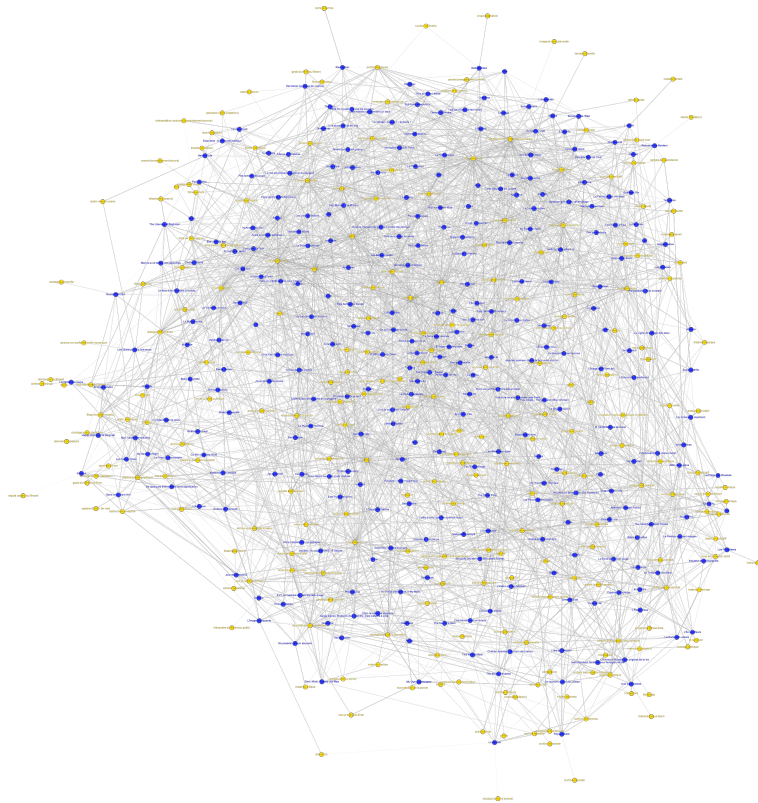


Toutefois, la démultiplication des liens entre films ne constitue pas une condition suffisante pour proposer aux utilisateurs de la plateforme des recommandations satisfaisantes. Encore faut-il que, de l’ensemble de ces nouveaux liens, ressortent bien les plus pertinents aux yeux des programmeurs dont le dispositif algorithmique cherche à imiter la recommandation. La figure 2, dans laquelle figurent les 177 films indexés et les termes décrivant leurs dispositifs de réalisation associés, rend prégnant l’enjeu de la hiérarchisation des liens entre films selon leur pertinence. La multitude de liens créés par l’indexation des dispositifs y apparaît fortement entremêlée.

Pourtant, si l’on zoome sur ce graphique à la recherche des cinq films au ton comique et satirique présentés dans la figure 1, on les retrouve bien à proximité (figure 3), au milieu d’autres avec lesquels ils partagent une forme de discours argumentative et un montage rhétorique (*Aujourd’hui* d’Esther Choub), une caractéristique formelle telle que l’intervention du réalisateur ou de la réalisatrice en voix hors champ (*La Femme aux cent visages* de Jean-Daniel Pollet) ou bien le détournement de matériaux de leur sens initial (*The Centrifuge brain project* de Till Novak).

La principale fonction de l’outil algorithmique mis à disposition par Spideo, Rumo, est de faire le tri des liens entre films formés par l’indexation. Il propose,

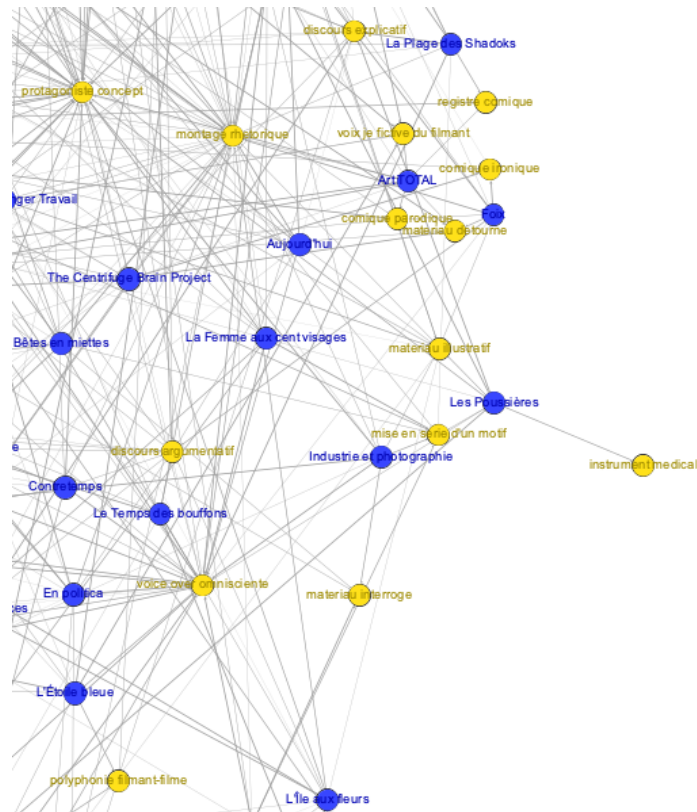
Figure 2 - Liens entre films formés par l'indexation des dispositifs de réalisation. Les nœuds correspondant à des films sont colorés en bleu, les descripteurs en jaune. Graphe généré avec Gephi, ForceAtlas2 en mode LinLog avec déchevauchement et dissuadement des concentrateurs (*hubs*), avec l'indexation de 177 films du catalogue de Tënk.



pour chaque film, une liste de 30 autres ordonnés selon un score calculé au moyen d'une mesure de similarité cosinus.

Le tableau présente la liste hiérarchisée des films associés au *Temps des bouffons* (à gauche) et le score du premier film recommandé, *L'île aux fleurs* (à droite).

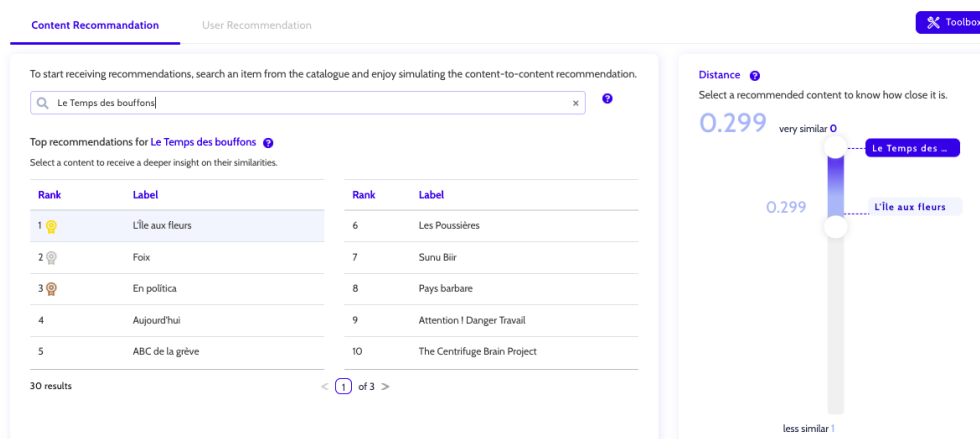
Figure 3 – Liens entre films formés par l’indexation des dispositifs de réalisation (vue de détail). Les nœuds correspondant à des films sont colorés en bleu, les descripteurs en jaune. Graphe est généré avec Gephi, ForceAtlas2 en mode LinLog avec déchevauchement et dissuadement des concentrateurs (*hubs*), avec l’indexation de 177 films du catalogue de Tènk



3.2 Évaluer la pertinence de la recommandation algorithmique

Le fruit de cette recommandation soulève toutefois beaucoup d’interrogations. Ces liens sont-ils susceptibles d’amener les utilisateurs à découvrir des contenus à la fois nouveaux pour eux et similaires au film de départ ? Plus spécifiquement, sont-ils à la hauteur d’une recommandation éditoriale proposée par un programmeur ? Ces questions soulèvent celle de l’évaluation de la recommandation, de ses buts et de ses moyens.

Figure 4 – Capture d'écran du tableau de bord de Rumo.



Les buts, d'abord. L'évaluation de notre expérimentation a été menée dans deux directions différentes. Premièrement, nous avons cherché à distinguer les convergences et divergences entre recommandation éditoriale et algorithmique en les comparant. Deuxièmement, pour nous assurer que les métadonnées soumises à l'algorithme soient optimisées à des fins de recommandation, plusieurs formats ont été testés.

Les moyens, ensuite. Pour suivre ces pistes d'évaluation de la recommandation, plusieurs outils d'appréhension, visuelle et quantitative, ont été mobilisés. Le tableau de bord de Rumo n'apportant pas de vision globale de la recommandation et imposant un examen film à film, des solutions de rechange ont été trouvées dans les outils de représentation de réseaux.

3.2.1 Buts de l'évaluation

Un des enjeux de l'implémentation d'un algorithme pour recommander des films entre eux sur la plateforme Tënk étant d'automatiser une tâche éditoriale humaine, son évaluation ne peut faire l'impasse d'une comparaison entre recommandation algorithmique et recommandation éditoriale. Cette comparaison implique plusieurs défis. Elle nécessite d'abord la récolte d'une recommandation éditoriale directe auprès d'un panel à la culture du cinéma documentaire assez riche et homogène pour que ses membres aient vu les mêmes films et qu'il soit possible de croiser leurs suggestions de recommandations. Dans cette perspective, huit spécialistes du cinéma documentaire (programmeurs de Tënk, cher-

cheurs et formateurs en études cinématographiques, tous également auteurs de films documentaires) ont été consultés. Un premier corpus de 112 films leur a été soumis, en leur demandant les liens qu'ils peuvent faire entre chacun de ces films selon leurs caractéristiques formelles et discursives (voir Gantier et al., 2020). Ce recueil a permis de relever quelques liens forts entre films, des liens confirmés par au moins trois experts. Parmi eux, cinq familles de films fréquemment associés ont pu être recensées. Au terme de l'exercice, les limites de la démarche ressortent clairement : quatre films sur 112 avaient été vus par les huit experts sollicités. Ce dispositif d'évaluation de la recommandation se trouve ainsi confronté à une limite d'échelle, celle-là même qui pousse à recourir à un algorithme pour générer des recommandations.

Parallèlement, la seconde piste d'évaluation a consisté à faire varier légèrement les métadonnées soumises à l'algorithme (nombre de descripteurs par film, avec ou sans pondération, avec ou sans prise en considération des liens sémantiques entre termes descripteurs du thésaurus) pour observer leurs effets sur la recommandation et déterminer quel format de métadonnées amenait les résultats les plus intéressants. Après la computation algorithmique, nous nous sommes également demandé s'il était possible d'établir un seuil de pertinence du score de la mesure cosinus calculée par Rumo¹¹, ou si l'on pouvait considérer qu'au-delà d'un certain ordre, dans la liste des recommandations proposées, la pertinence des films proposés n'était pas moindre.

3.2.2 Moyens de l'évaluation

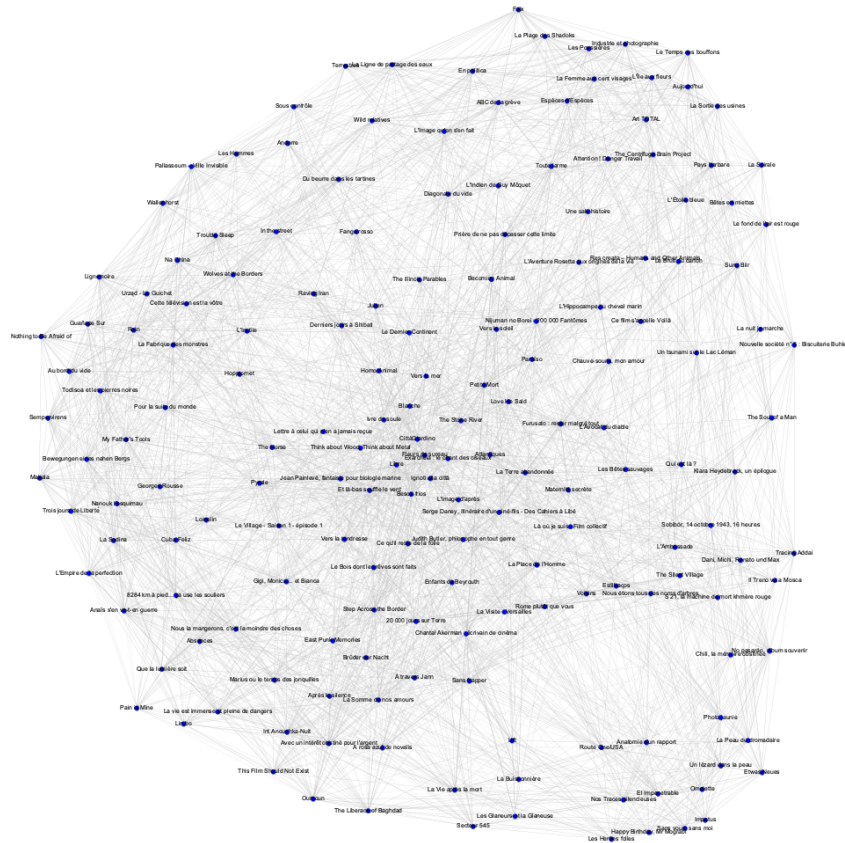
Pour mettre en œuvre ces comparaisons, les outils de visualisation et d'analyse de réseaux présentent l'avantage de donner une vision globale des résultats du calcul algorithmique.

La représentation s'appuie sur une indexation des dispositifs cinématographiques pondérés selon une échelle par paliers (1, 3, 5, 8, 10).

La visualisation du réseau formé par la recommandation sous forme d'un graphique nœuds-liens (figure 5), au même titre que la représentation des liens entre films et descripteurs (figure 2), n'est pas des plus lisibles. Tout au plus pouvons-nous repérer le positionnement des films dans le plan, repérer ceux qui semblent faire le lien entre plusieurs ensembles de films recommandés entre eux (au centre du graphique) et d'autres qui semblent former des sous-ensembles plus isolés (à

11. Sur les modalités du recueil de la recommandation éditoriale, mais aussi sur les enjeux de l'utilisation de cette mesure cosinus, dite cosinus de Salton, voir Gantier et al. (2020).

Figure 5 – Représentation en réseau des liens proposés par Rumo entre 177 films du catalogue de Tènk.

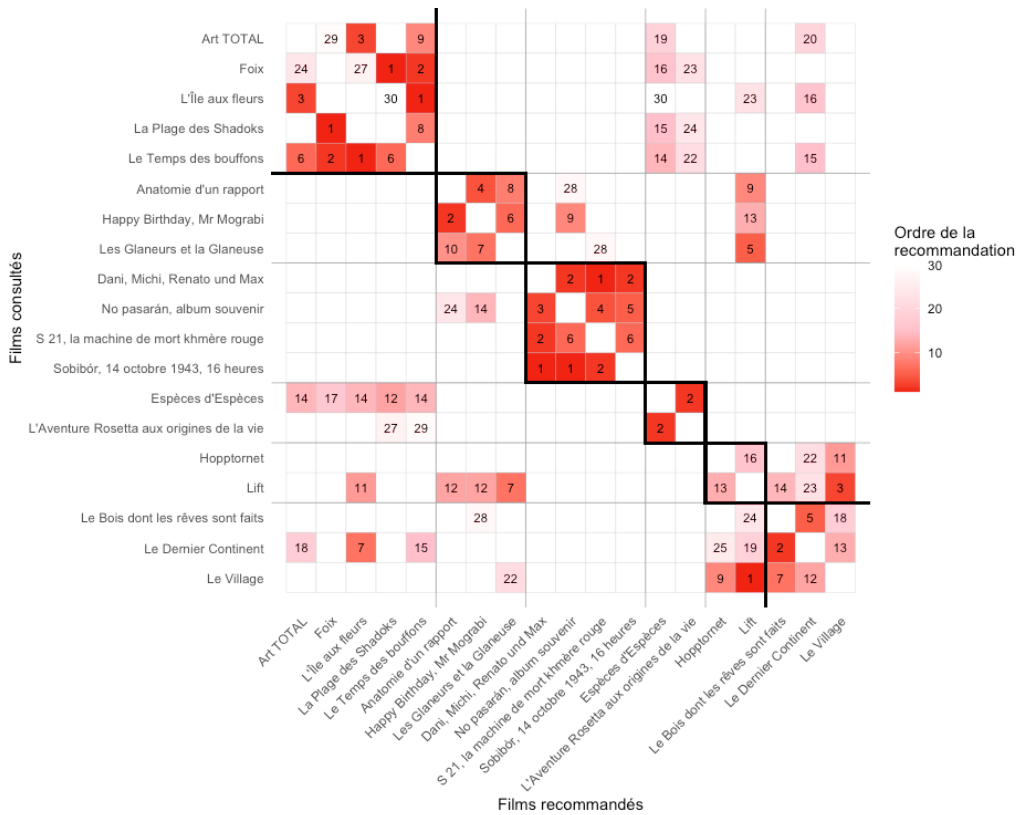


la périphérie). Mais les liens, bien qu'ils soient proportionnels au score de similarité cosinus, se confondent et se superposent.

Pour dépasser ces limites et mettre en œuvre la comparaison entre recommandations humaine et algorithmique, nous nous sommes donc tournés vers une autre forme classique de représentation des réseaux, la matrice d'adjacence. Celle-ci place chaque sommet à la fois en ligne et en colonne d'un tableau à double entrée. Si deux sommets sont reliés, la case à l'intersection de leur ligne et de leur colonne est remplie d'une valeur numérique continue caractérisant le lien. Malgré les difficultés d'appréhension de l'outil, moins intuitif que la représentation en graphique nœud-lien, la matrice de contingence présente notamment l'avantage de rendre plus visibles les absences de connexion et, du même coup, de faciliter l'identification de communautés de sommets, dès lors qu'un réordon-

nancement des lignes et colonnes en fonction des liens est opéré (Henry & Fekete, 2008). Ainsi, pour comparer recommandation éditoriale et algorithmique, les films qui avaient été recommandés entre eux par trois membres du panel ont été ordonnés en cinq groupes, reconstituant les familles de films préalablement identifiées.

Figure 6 – Matrice d’adjacence présentant les liaisons algorithmiques entre films du corpus. Films associés entre eux par au moins trois des membres du panel de spécialistes, s’appuyant sur une indexation des dispositifs cinématographiques pondérés selon une échelle logarithmique de base 2



Les liens sont caractérisés par l’ordre du film dans la recommandation proposée par Rumo.

Cette matrice d’adjacence laisse apparaître une forme d’équivalence entre recommandation algorithmique et recommandation éditoriale : les films reliés entre eux par les experts se trouvent bien recommandés ensemble après un calcul de similarité cosinus, les blocs centraux de la matrice étant bien remplis de liens,

plus ou moins forts selon les groupes. Des liens extérieurs aux blocs centraux sont toutefois à noter.

La visualisation en matrice d'adjacence a permis également de comparer les formes prises par la recommandation après avoir fait varier le nombre de descripteurs pris en considération ou leur pondération notamment. Les figures 6 et 7 montrent ainsi l'évolution de la recommandation en fonction de la pondération : la figure 6 a été générée à partir de métadonnées pondérées suivant une échelle logarithmique de base 2, tandis que la figure 7 avait pour base des données pondérées suivant une échelle par paliers (détaillés dans le tableau 1).

Les liens sont caractérisés par l'ordre du film dans la recommandation proposée par Rumo.

Figure 7 - Matrice d'adjacence présentant les liaisons algorithmiques entre films du corpus. Films associés entre eux par au moins trois des membres du panel de spécialistes, s'appuyant sur une indexation des dispositifs cinématographiques pondérés selon une échelle par paliers (1, 3, 5, 8, 10)

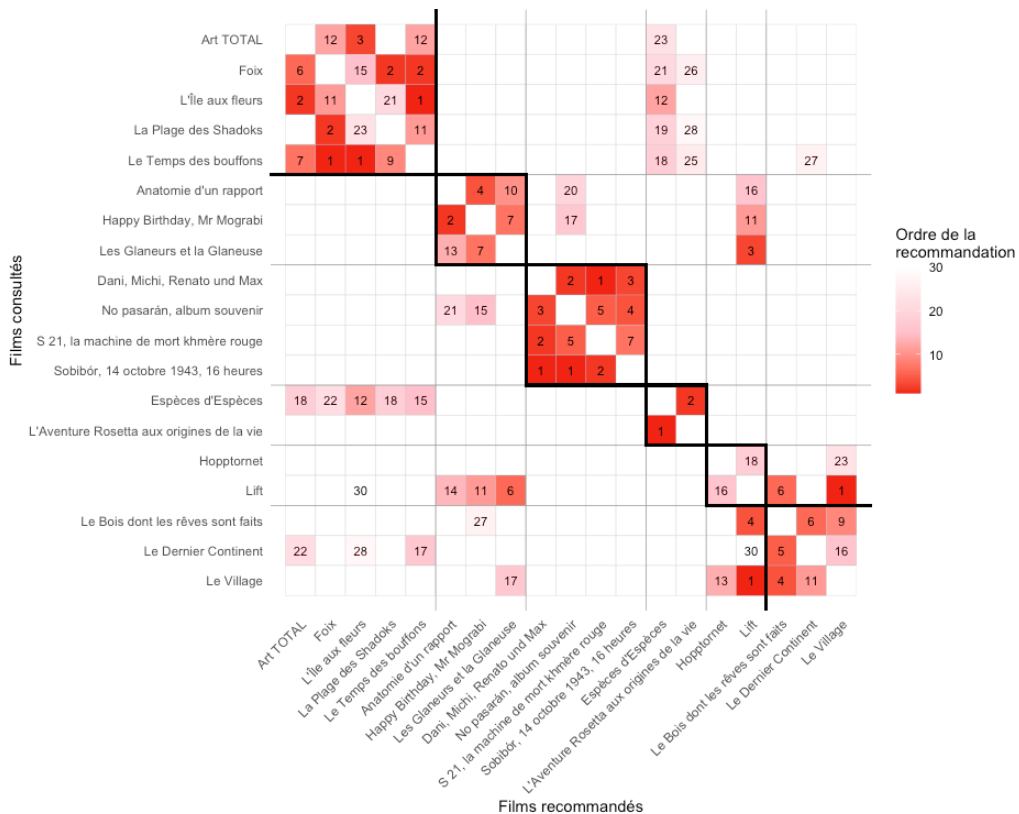


Table 1 – Deux échelles de pondération testées et évaluées dans les figures 5 et 6.

Catégorie des descripteurs	Pondération selon une échelle \log_2	Pondération par paliers
Descripteurs n° 1 à 3	16	10
Descripteurs n° 4 et 5	8	8
Descripteurs n° 6 à 8	4	5
Descripteurs n° 9 et 10	2	3
Descripteurs n° 11 et suivants	1	1

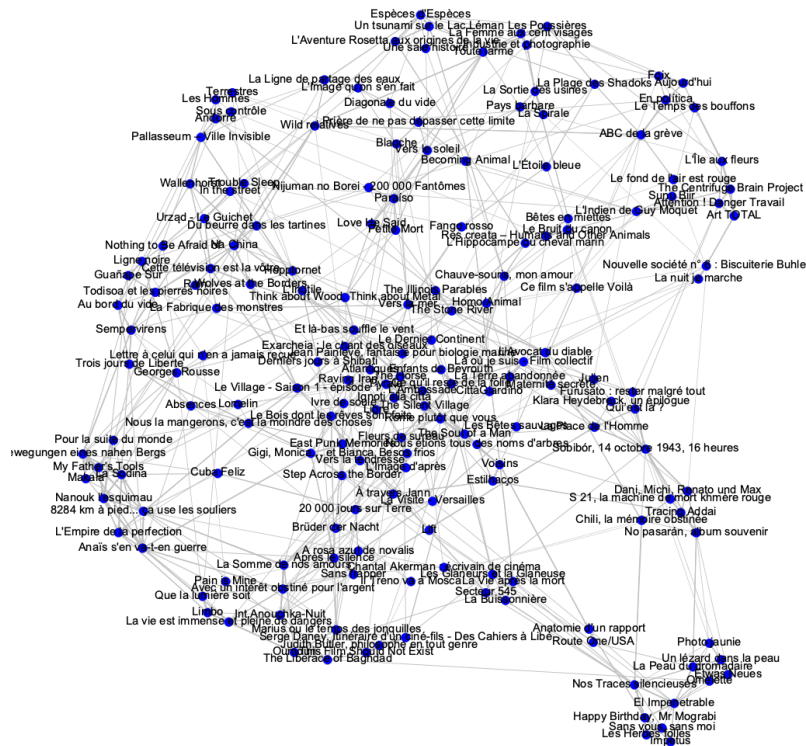
Au regard de ces deux matrices, la pondération par paliers semble la plus en adéquation avec les recommandations humaines : nous y trouvons moins de liens hors des groupes préétablis, et les liens entre films du même groupe sont plus forts et plus nombreux.

Outre les matrices d'adjacence, une autre piste explorée a été de filtrer l'affichage des liens sur un graphique nœud-lien en fonction d'un seuil de score ou d'ordre de la recommandation proposée dans Rumo. Cette solution permet d'améliorer la lisibilité des résultats de la recommandation et d'alimenter notre réflexion sur les modalités de tri des liens selon leur pertinence.

Grâce à ce tri des liens, les groupes de films composés par la recommandation ainsi que leurs relations internes et externes ressortent plus clairement. La mise en œuvre d'un calcul de modularité à des fins de classification des films en fonction de leurs relations permet encore d'améliorer l'appréhension de ces groupes.

On retrouve sur ce graphique la même disposition des films dans l'espace, mais le tri des liens et la classification permettent de mieux saisir le positionnement de chaque film. Les cinq films que l'on avait retenus pour leur tonalité comique se trouvent par exemple placés dans le groupe couleur cyan, en périphérie de la figure 9. Au contraire des films des groupes orange, rouge et violet, ils ne semblent pas faire de lien transclasses. On peut y voir la confirmation de la pertinence de la recommandation, conforme à la recommandation éditoriale. Mais cela peut être aussi interprété comme une limite de la démarche : si les films sont intégrés à des groupes de recommandation trop autonomes, l'outil ne décourage-t-il pas la découverte de nouveaux univers ?

Figure 8 – Représentation en réseau des liens (ordre inférieur à 5 ou score inférieur à 0,5) proposés par Rumo entre films du corpus. Représentations s'appuyant sur une indexation des dispositifs cinématographiques pondérés selon une échelle par paliers (1, 3, 5, 8, 10)



4 Discussion : limites de la démarche

4.1 Une expérimentation algorithmique

Pour évaluer notre expérimentation de la recommandation automatisée de films basée sur leurs métadonnées et appréhender les résultats du calcul algorithmique, nous nous sommes trouvés face à deux difficultés.

La première semble inhérente à notre terrain : constituer un panel de prescripteurs ayant vu la totalité des films que nous avons indexés était impossible, il nous a fallu concentrer notre regard sur un petit nombre de films dont la recommandation faisait consensus auprès des spécialistes consultés.

générer des métadonnées pertinentes, et du fait de la faiblesse des métadonnées existantes (Foulonneau, 2014). Ces deux aspects peuvent être compensés par la prise en compte des traces d'usage et par la délégation aux utilisateurs de la fabrication de ces métadonnées par la folksonomie, dispositif proposant aux utilisateurs d'indexer eux-mêmes les contenus, librement ou au moyen d'une liste de termes préétablie. Toutefois, ces méthodes trouvent souvent leurs limites dans le manque de cohérence des métadonnées qu'elles produisent.

4.3 La recommandation de contenu à contenu, une technologie de l'intelligence artificielle ?

Une autre question soulevée par cette contribution concerne la nature du dispositif de recommandation considéré : relève-t-il vraiment d'une intelligence artificielle ? Oui, dans le sens où il s'agit d'un système déductif (Farchy & Denis, 2020, p. 21-22), pensé pour reproduire par l'algorithmie un processus intelligent et remplacer un travail intellectuel humain par une action automatisée. Mais, pour mener à bien cette automatisation, un travail de documentation reste indispensable, travail dont dépend la pertinence de la recommandation algorithmique. En rejetant l'utilisation de traces d'usages, en n'automatisant pas l'indexation, en ne mobilisant pas d'algorithme prédictif ou de technologie d'apprentissage par la machine, la recommandation algorithmique de contenu à contenu dont il est question dans ce chapitre paraît finalement moins artificielle qu'humaine.

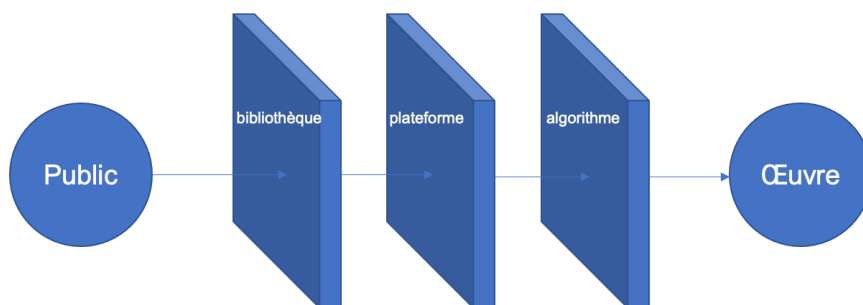
Cependant, l'automatisation de la recommandation de biens culturels au moyen d'algorithmes ne peut pas non plus être considérée comme anodine. Si les effets sociaux de tels dispositifs sont moins évidents que ceux des algorithmes utilisés pour automatiser les actions de police ou de justice, ils n'en sont pas moins réels. En donnant de la visibilité ou en n'affichant pas certains films, ils contribuent notamment à forger ou entretenir les goûts des utilisateurs des plateformes (Lawrence, 2015). Certes, ce phénomène est particulièrement sensible dans le cas de recommandations basées sur les traces d'usage (Beuscart et al., 2019 ; Micheau, Després-Lonnet, & Cotte, 2017 ; Roberge, Jamet, & Rousseau, 2019). Chercher à recommander des contenus à partir de leur similarité avec un autre tend toutefois également à renforcer l'exposition des usagers à un ensemble homogène de documents, et ainsi « enfermer les individus dans l'espace des possibles » (Coulangeon, 2020) dessiné par leur choix de départ.

4.4 Une multiplicité d'acteurs qui participent à la valorisation d'œuvres patrimoniales

Si l'utilisation d'algorithmes de recommandation par des plateformes VAD/VADA peut servir à améliorer la découvrabilité du patrimoine cinématographique, elle peut aussi s'avérer encombrante pour les acteurs de la médiation de ces œuvres qui ne l'ont pas entamée. Dès lors que la médiation algorithmique devient incontournable pour avoir accès à l'œuvre, elle se superpose avec les autres formes de médiation, notamment humaine.

De fait, le recours à un dispositif artificiel de recommandation ne vient pas régler la question de la multiplicité des intérêts des acteurs qui participent à la valorisation d'œuvres audiovisuelles patrimoniales, aussi variés que ceux de leur potentiel public. Parmi ces intermédiaires du patrimoine cinématographique, les médiathécaires et les bibliothécaires notamment. Ils interviennent théoriquement au terme de la circulation commerciale des films, en lien direct avec leurs usagers. En sélectionnant un film sur support physique pour l'intégrer à leurs collections, ils héritent bien sûr des actions de médiation antérieures mises en œuvre par leurs producteurs, distributeurs et éditeurs. Mais, depuis le développement de plateformes VAD/VADA, ces actions continuent de prendre place au sein même des médiathèques. Conservateurs et bibliothécaires n'y sont plus les seuls à assurer le lien entre des œuvres et leur public. Ils sont concurrencés sur ce terrain par les plateformes, qui font varier leur catalogue suivant une logique qui leur est propre, et par les dispositifs de recommandation algorithmique implémentés en leur sein, qui orientent à leur tour la prescription.

Figure 10 – Schématisation de l'enchâssement des acteurs et outils de la médiation du cinéma documentaire. Outils diffusés par les plateformes de VAD mis à disposition de leurs usagers par les bibliothèques



Or la principale mission des bibliothécaires et des médiathécaires reste d'établir des liens entre usagers et œuvres, voire entre usagers :

On ne peut pas faire l'économie de ce lien avec l'utilisateur. Lorsqu'on pense un accès à distance comme la VOD, il faut réfléchir à quel lien on maintient avec les usagers. On ne peut pas simplement rester à un lien qui serait celui d'un algorithme¹².

Dans cette perspective, le recours à des algorithmes de recommandation par les plateformes semble un frein à l'activité de médiation des professionnels de bibliothèques. Pour autant, les dispositifs de recommandation algorithmiques pourraient également être mobilisés directement par ces professionnels pour les aider dans leurs missions, en particulier pour suggérer de nouveaux liens entre films, au même titre que les outils documentaires qu'ils utilisent au quotidien.

5 Conclusion

L'efficacité des algorithmes de contenu à contenu pour faciliter la médiation du patrimoine cinématographique est dépendante des données auxquelles il est appliqué. Les liens entre films qu'ils mettent en avant sont avant tout tissés par les stratégies documentaires établies en amont par les structures qui les implémentent. Ils permettent toutefois, combinés à des outils de visualisation de la recommandation, d'adapter la documentation sciemment. Ils constituent donc une solution de rechange solide, plus éthique et plus humaine, aux algorithmes de recommandation classiquement basés sur les traces d'usage d'utilisateurs. En perpétuant des usages sociaux lui préexistant, le recours à un algorithme de recommandation de contenu à contenu ne résout pas la difficulté à concilier les logiques concurrentes des acteurs de la médiation du cinéma patrimonial. Mais il constitue un outil comme un autre dont les acteurs de la médiation traditionnelle peuvent aussi s'emparer pour en faire une technologie collective conforme à leurs valeurs.

12. Jean-Yves de Lepinay (président d'Images en bibliothèque), « Quelle place et quelles missions pour les bibliothèques ? Appréhender les différentes offres et l'évolution des pratiques en médiathèque », *VàD et ressources numériques*, formation professionnelle organisée en ligne par l'association Images en bibliothèque à destination de médiathécaires, 29 avril 2021.

Références

- Adary, A., & Domenget, J.-C. (2017). Entretien : pour une intégration des data-déontologues dans les équipes communication. *Communication & professionnalisation*, 6, 152-160. doi: 10.14428/rcompro.v0i6.3483
- Bachimont, B. (2007). Nouvelles tendances applicatives : de l'indexation à l'éditionnalisation. In P. Gros (Ed.), *L'Indexation multimédia* (pp. 313-326). Paris : Lavoisier Hermes science publications.
- Benghozi, P.-J., & Paris, T. (2014). L'économie culturelle à l'heure du numérique : une révolution de l'intermédiation. In L. Jeanpierre & O. Roueff (Eds.), *La culture et ses intermédiaires* (pp. 175-188). Paris : Éditions des archives contemporaines.
- Benghozi, P.-J., & Paris, T. (2016). The cultural economy in the digital age : A revolution in intermediation? *City, Culture and Society*, 7(2), 75-80. doi: 10.1016/j.ccs.2015.12.005
- Beuscart, J.-S., Coavoux, S., & Maillard, S. (2019). Les algorithmes de recommandation musicale et l'autonomie de l'auditeur. *Réseaux*, 213(1), 17-47. doi: 10.3917/res.213.0017
- Bonenfant, M., Crémier, L., & Lafrance St-Martin, L. I. (2018). Quelques réflexions sémiotiques sur le circuit des données massives. In A. Mondoux & M. Ménard (Eds.), *Big Data et société. Industrialisation des médiations symboliques* (pp. 131-164). Québec : Presses de l'Université du Québec.
- Bullich, V. (2012). L'intégration de la production cinématographique aux technologies numériques de l'information et de la communication. In M. Bourgatte & V. Thabourey (Eds.), *Le cinéma à l'heure du numérique. Pratiques et publics* (pp. 22-43). Paris : MkF éditions.
- Bullich, V., & Clavier, V. (2018). Dossier 2018 : Production des données, « Production de la société ». Les Big Data et algorithmes au regard des Sciences de l'information et de la communication. *Les Enjeux de l'information et de la communication*, 2018(2), 5-177.
- Cardon, D. (2015). *À quoi rêvent les algorithmes. Nos vies à l'heure des big data*. Paris : Seuil.
- Casilli, A. A. (2019). *En attendant les robots. Enquête sur le travail du clic*. Paris : Seuil.
- Coulangéon, P. (2020). Comment se forment les goûts culturels aujourd'hui ? In S. Paugam (Ed.), *50 questions de sociologie* (pp. 169-176). Paris : Presses universitaires de France.
- Danard, B. (2020). *L'économie des films de patrimoine | CNC* (Étude). Paris : CNC.

- Delcroix, G. (2019). Les données et les algorithmes, nouvelles muses de la prescription culturelle? In B. Chapelain & S. Ducas (Eds.), *Prescription culturelle : avatars et médiamorphoses* (pp. 347–360). Villeurbanne : Presses de l'enssib.
- Dessinges, C., & Perticoz, L. (2019). Les consommations de séries télévisées des publics étudiants face à Netflix : une autonomie en question. *Les Enjeux de l'information et de la communication*, 20(1), 5–23. doi: 10.3917/enic.026.0005
- Direction des études, des statistiques et de la prospective. (2021). *Le Bilan 2020 du CNC* (Bilan). Paris : CNC.
- Doueihy, M. (2013). *Qu'est ce que le numérique ?* Paris : Presses universitaires de France.
- Du Val, J.-L. (2017). "Toutes les images du monde" au plus près de chez soi : donner accès aux films dans les bibliothèques. In D. Rousselet, J. Guillaumot, & M. Palesse (Eds.), *Du cinéma en bibliothèque* (pp. 69–76). Paris : ABF; Images en bibliothèques.
- Epron, B., & Vitali-Rosati, M. (2018). I. Pour une définition de l'édition à l'ère numérique. In B. Epron (Ed.), *L'édition à l'ère numérique* (pp. 5–33). Paris : La Découverte.
- Farchy, J., & Denis, J. (2020). *La culture des données. Intelligence artificielle et algorithmes dans les industries culturelles*. Paris : Mines ParisTech.
- Flichy, P. (2021). Produire les données, à la frontière de plusieurs mondes sociaux. *Réseaux*, 228(4), 9–19. doi: 10.3917/res.228.0009
- Foulonneau, M. (2014). Recommandeurs et diversité. Exploitation de la longue traîne et diversité des listes de recommandations. In G. Kembellec, G. Chartron, & I. Saleh (Eds.), *Les Moteurs et systèmes de recommandation* (pp. 85–104). London : ISTE.
- Gantier, S. (2020). Construction de pratiques cinéphiles sur une plateforme de vidéo à la demande : Enjeu du design des personae. *Les Enjeux de l'information et de la communication*, 21(1), 53–73.
- Gantier, S., Givois, È., & Kergosien, É. (2020). Recommander autrement les films documentaires ? Design d'une preuve de concept basée sur l'indexation des dispositifs de réalisation. *RIHM - Revue des Interactions Humaines Médiatisées*, 21(2), 1–26.
- Givois, È., Bolka-Tabary, L., Kergosien, É., Gantier, S., & El Haoudi, B. (2021). Élaboration itérative d'un thésaurus pour indexer le cinéma documentaire à partir des procédés filmiques. *I2D - Information, données documents*, 2(2), 122–150. doi: 10.3917/i2d.212.0122
- Henry, N., & Fekete, J.-D. (2008). Représentations visuelles alternatives pour les réseaux sociaux. *Réseaux*, 152(6), 59–92.

- Jeanpierre, L., & Roueff, O. (2014). *La culture et ses intermédiaires. Dans les arts, le numérique et les industries créatives*. Paris : Éditions des Archives contemporaines.
- Kilborne, Y. (2008). *L'expérience documentaire. Approche communicationnelle du cinéma de réalité* (Thèse de doctorat). Paris 8, Paris.
- Laurent, N. (2020a). Le patrimoine est-il soluble dans le marché? In C. Gauthier (Ed.), *Patrimoine et patrimonialisation du cinéma* (p. XX-YY). Paris : École nationale des chartes.
- Laurent, N. (2020b). « L'ivresse patrimoniale », et après? *1895*, 91(2), 8–30. doi: 10.4000/1895.7994
- Lawrence, E. (2015). Everything is a recommendation netflix, altgenres and the construction of taste. *Knowledge Organization*, 42(5), 358–364.
- Le Deuff, O. (2006). Folksonomies. Les usagers indexent le web. *Bulletin des bibliothèques de France (BBF)*, 51(4), 66–70.
- Le Ludec, C., Wahal, E., Casilli, A. A., & Tubaro, P. (2020). Quel statut pour les petits doigts de l'intelligence artificielle? Présent et perspectives du micro-travail en France. *Les Mondes du travail*, 24–25, 99–112.
- Lesauvier, M.-E. (2019). Une plateforme au service d'un monde professionnel mobilisé. Enjeux symboliques et économiques. *tic&société*, 13(Vol. 13, N° 1-2), 225–246. doi: 10.4000/ticetsociete.3299
- Lindeperg, S., & Szczepaska, A. (Eds.). (2017). *À qui appartiennent les images? : le paradoxe des archives, entre marchandisation, libre circulation et respect des œuvres*. Paris : Éditions Maison des sciences de l'homme.
- Menger, P.-M. (2017). Introduction. In P.-M. Menger & S. Paye (Eds.), *Big data et traçabilité numérique : les sciences sociales face à la quantification massive des individus* (pp. 7–23). Paris : Collège de France.
- Micheau, B., Després-Lonnet, M., & Cotte, D. (2017). La recommandation musicale entre inscriptions documentaires, pratiques sociales, et dispositifs d'écoute. *Études de communication*, 49(49), 33–56. doi: 10.4000/edc.7014
- Pariser, E. (2011). *The Filter Bubble. What The Internet Is Hiding From You*. London : Viking.
- Roberge, J., Jamet, R., & Rousseau, A. (2019). *L'impact social des algorithmes de recommandation sur la curation des contenus musicaux francophones au Québec. Enquête qualitative*. (Étude exploratoire). Québec : Institut national de la recherche scientifique.
- Simonnot, B. (2018). Conduire des recherches en régime numérique : vers un cadre conceptuel de réflexion éthique. In L. Balicco, E. Broudoux, G. Chartron, V. Clavier, & I. Pailliant (Eds.), *L'éthique en contexte info-communicationnel numérique. Déontologie, régulation, algorithme, espace public. Actes du*

- colloque « Document numérique et société » (pp. 11–21). Louvain-la-Neuve : De Boeck Supérieur. doi: 10.3917/dbu.balic.2018.01.0071
- Taillibert, C. (2015). Festivals et patrimoine cinématographique à l'heure du numérique. La situation française. *Secuencias. Revista de historia del cine*, 39, 83–99.
- Taillibert, C. (2020). *Vidéo à la demande : une nouvelle médiation ? Réflexions autour des plateformes cinéphiles françaises*. Paris : L'Harmattan.
- Taillibert, C. (2021). Penser l'accompagnement éducatif cinéphile dans un environnement numérique. *Communiquer. Revue de communication sociale et publique*, 31, 75–96. doi: 10.4000/communiquer.7786
- Thuillas, O., & Wiart, L. (2019a). Les plateformes de VOD cinéphiliques : des stratégies de niche en questions. *Les Enjeux de l'information et de la communication*, 20(1), 39–55. doi: 10.3917/enic.026.0039
- Thuillas, O., & Wiart, L. (2019b). Plateformes alternatives et coopération d'acteurs : quels modèles d'accès aux contenus culturels ? *tic&société*, 13(1-2), 13–41. doi: 10.4000/ticetsociete.3043
- Tiffon, G. (2017). La contribution des internautes aux big data : un travail ? In P.-M. Menger & S. Paye (Eds.), *Big data et traçabilité numérique : les sciences sociales face à la quantification massive des individus* (pp. 69–81). Paris : Collège de France.
- Vander Wal, T. (2004, août). *You Down with Folksonomy ?*
- Vayre, J.-S. (2017). Une histoire de machines à recommander des biens de consommation : de l'Internet documentaire à l'Internet des données. *Études de communication. langages, information, médiations*, 49, 89–106. doi: 10.4000/edc.7178
- Vayre, J.-S. (2018). Comment décrire les technologies d'apprentissage artificiel ? Le cas des machines à prédire. *Réseaux*, 211(5), 69–104. doi: 10.3917/res.211.0069