



HAL
open science

Les Échos visuels

Shiming Shen, Matteo Treleani, Aline Menin, Marco Winckler

► **To cite this version:**

Shiming Shen, Matteo Treleani, Aline Menin, Marco Winckler. Les Échos visuels. Humanistica 2024, Association francophone des humanités numériques, May 2024, Meknès, Maroc. hal-04563421v2

HAL Id: hal-04563421

<https://hal.science/hal-04563421v2>

Submitted on 23 May 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.



Distributed under a Creative Commons Attribution 4.0 International License

Les Échos visuels

Exploration des Phénomènes de Répétition dans les Archives audiovisuelles

Shiming Shen, Matteo Treleani, Aline Menin, Marco Winckler

Université Côte d'Azur

{shiming.shen, matteo.treleani}@univ-cotedazur.fr

{aline.menin, marco.winckler}@inria.fr

Résumé

Cette communication présente le projet *Crossing Borders Archives*, qui se consacre à l'étude de la diffusion des archives audiovisuelles relatives à la construction de l'Union européenne au sein des médias français et italiens. Le projet examine comment ces images façonnent l'imaginaire visuel collectif, en utilisant des données franco-italiennes datant de 2001 à 2021, totalisant 37 006 images d'archives dans les reportages. Le corpus comprend des vidéos, des images clés, et des métadonnées enrichies. Des outils computationnels ont été utilisés pour l'analyse d'image et la recherche de contenu. L'étude met en évidence l'importance des métadonnées et présente une plateforme interactive pour l'analyse multimodale, permettant aux utilisateurs de naviguer et explorer les archives par mots-clés et outils dédiés. Elle révèle également les défis liés à l'analyse de données culturelles, mettant en lumière le phénomène de répétition et d'homogénéisation dans les productions audiovisuelles liées à l'actualité.

1 Introduction

Le domaine de l'actualité fonctionne souvent comme un espace où les événements sont reconstitués ou simulés, et qui est imprégné de stéréotypes. Comprendre le langage visuel de ce domaine peut grandement améliorer notre compréhension de la manière dont les représentations sont façonnées. Nous supposons que le paysage visuel des actualités à travers divers médias est nettement homogénéisé, en grande partie à cause de l'utilisation d'images d'archives (Treleani et al., 2023; Shen et al., 2023).

Le projet ANR CROBORA (*Crossing Borders Archives*¹) vise à étudier les dynamiques entourant la construction de l'Union européenne dans les médias télévisuels et en ligne. CROBORA postule que la diffusion d'images d'archives favorise

un imaginaire européen commun, potentiellement sélectif, reposant sur un nombre limité d'éléments visuels. Le projet explore l'impact de diverses médiations – techniques, politiques, culturelles – sur le recyclage de ces matériaux, perçus comme des flux migratoires franchissant des frontières nationales, médiatiques et institutionnelles. Une étude empirique a été réalisée pour déterminer les images d'archives les plus réutilisées et analyser l'évolution de leur signification au fil du temps, de l'espace et des contextes. L'objectif est de décrypter les mécanismes de circulation de ces images et d'atteindre des objectifs tels que l'identification des contenus visuels fréquemment utilisés, la cartographie de leur utilisation et l'examen de leur transformation sémantique.

2 Etat de l'art

Dans le sillage du « tournant computationnel » (Berry, 2011), les chercheurs en sciences sociales et humanités numériques sont constamment poussés à renouveler, diversifier et réinventer leurs méthodologies. Cette nécessité émerge de l'évolution rapide des pratiques sociales, culturelles et scientifiques dans un monde de plus en plus dominé par le numérique. Parmi les innovations, la méthodologie des « données denses » (*thick data*) (Latzko-Toth et al., 2017), s'avère particulièrement appropriée pour l'analyse des données culturelles. Cette approche se distingue par l'utilisation de techniques qui dépassent la simple analyse computationnelle ou la collecte massive de données, en valorisant aussi la qualité et la profondeur des informations recueillies.

Dans les études visuelles, l'analyse en profondeur et à grande échelle des images devient réalisable avec l'avancement de la vision par ordinateur, qui optimise la rapidité et l'étendue de l'interprétation en agissant comme intermédiaire entre l'œil humain et l'image numérique (Arnold et Tilton, 2023). Récemment, la théorie de la multimodalité

1. Projet financé par l'Agence nationale de la recherche (2020-2024), <https://croborahuma-num.fr>.

a apporté un éclairage nouveau sur l'interaction entre textes et images à travers divers modes sémiotiques au sein d'un même artefact culturel (Hiipala, 2021; Smits et Wevers, 2023). Le principe de la multiplication des sens illustre que la signification d'un ensemble formé par une image et un texte – comme une photo et sa légende – change si l'on considère les deux éléments séparément (Bateman, 2014). Par exemple, Wevers et Smits (2019) ont démontré, en entraînant des réseaux de neurones convolutionnels pour distinguer les illustrations, les photographies et d'autres modes sémiotiques dans les archives de journaux, que les analyses informatiques des médias commencent à transcender les dichotomies simplistes entre textes et images.

Pour réaliser efficacement une analyse à grande échelle des artefacts multimodaux, l'utilisation d'approches numériques s'avère cruciale. Dans ce contexte, la méthode du *distant viewing* (Arnold et Tilton, 2019) peut être judicieusement combinée avec l'approche des données denses pour enrichir l'analyse. Cette méthode a été inspirée du concept de *distant reading* (Moretti, 2013a) appliqué à l'histoire littéraire, qui se concentre sur l'analyse extensive de textes. Généralement, lorsqu'on applique la vision par ordinateur à une image numérique, nous soutenons que cela constitue le *distant viewing* (Arnold et Tilton, 2023). Cette méthode se spécialise dans l'observation approfondie de matériaux visuels, qui comprend l'extraction de métadonnées informatives et représentatives de milliers d'images, telles que des bandes dessinées (Arnold et al., 2023), des photographies (Arnold et al., 2022), des séquences de télévision ou des films (Arnold et al., 2019), et l'usage d'algorithmes pour interagir avec ces métadonnées en vue de leur analyse.

En adoptant le cadre du *distant viewing* dans notre projet, nous nous proposons d'explorer les conventions visuelles spécifiques à un corpus centré sur l'Union européenne. L'objectif est d'identifier et d'examiner des éléments formels subtils au travers d'une série d'images, suivant l'approche qui met en avant l'importance de saisir les transformations de ces éléments sémantiques. L'usage de méthodes quantitatives offre une voie pour mettre au jour ces éléments souvent trop banals pour être immédiatement remarqués, tout en soulignant leur contribution essentielle à la construction de sens (Moretti, 2013a).

Auparavant, plusieurs projets se sont également

intéressés à la représentation et à la circulation des matériaux visuels. Ces projets impliquent généralement la sélection de cas d'étude spécifiques suivis de recherches ciblées dans des bases de données pour retrouver ces images (Shen et al., 2023). Cependant, notre projet diffère car nous ne connaissons pas *a priori* les images spécifiques à suivre; nous devons d'abord détecter des éléments récurrents dans notre corpus initial pour ensuite analyser leur circulation. Cette démarche ressemble à celle des études de suivi de textes, où des développements algorithmiques importants ont été réalisés pour identifier des textes récurrents. Cela inclut des initiatives comme le projet *Viral Texts*, qui vise à détecter des passages de texte fréquemment répétés à travers diverses éditions de journaux (Smith et al., 2015). Des efforts semblables sont consacrés à la détection de phrases ou de « mèmes » brefs répétés (Suen et al., 2013). Dans la section suivante, nous prévoyons de détailler notre méthode pour analyser le phénomène de répétition associé au corpus audiovisuel que nous avons collecté.

3 Collecte de données

Les données collectées couvrent la télévision en France et en Italie, incluant le dépôt légal de la télévision française (Institut national de l'audiovisuel/INA) et les archives du service public de télévision italien (Radiotelevisione italiana/Rai), ainsi que MFE - MediaForEurope (à l'origine Mediaset), un groupe de média italien présent dans le secteur télévisuel, la production audiovisuelle, la presse et l'Internet. L'accent est mis sur les journaux télévisés de 2001 à 2021 des principales chaînes nationales, y compris TF1, France 2 et ARTE pour la France, et Rai 1, Rai 2 et Canale 5 pour l'Italie. Pour la collecte de données sur le Web, deux approches ont été utilisées : le dépôt légal de l'INA de l'archive du Web et la collecte directe du Web vivant, en utilisant une recherche textuelle dans la base de données de l'INA, les moteurs de recherche du Web vivant, ainsi qu'un logiciel d'intelligence artificielle Snoop spécialisé dans la détection visuelle automatisée. Jusqu'à ce jour, le nombre total d'occurrences d'images d'archives dans notre corpus est de 37 006, réparties entre différents sujets de JT et vidéos en ligne.

4 Différents types de données

Notre corpus de recherche est constitué des données et des métadonnées suivantes :

1. **Vidéos** : Les artefacts audiovisuels sont le pilier central de notre projet. Ces vidéos, souvent non-segmentées, intègrent des images d'archives dans leur contexte original. Le défi majeur réside dans l'extraction et l'analyse de ces images d'archives spécifiques, noyées dans un flot audiovisuel continu.
2. **Image clé de chaque image d'archives** : La principale solution que nous avons adoptée pour extraire les images d'archives a consisté à identifier manuellement les séquences dans les vidéos et à capturer une image clé pour chacune. Les captures d'écran sont un substitut manipulable par lequel nous opérationnalisons (Moretti, 2013b) la tâche de travailler avec des séquences vidéo. Comme l'écrit également Lev Manovich (2020), pour visualiser une collection de vidéos, il est généralement plus pratique de sélectionner quelques images qui capturent les propriétés et les motifs de chaque vidéo et de travailler avec elles.
3. **Métadonnées originales des vidéos** : Ces métadonnées, accompagnant nos collections vidéo, sont issues de sources diverses, telles que les documentations de l'INA ou les descriptions fournies par des comptes officiels sur YouTube. Ces textes ne se contentent pas de décrire les vidéos ; ils imposent également des catégories culturelles aux données. En tant qu'éléments clés de la documentation des artefacts numériques, ces métadonnées offrent une fenêtre sur le contexte et la signification des objets audiovisuels collectés.
4. **Métadonnées CROBORA des images d'archives** : L'annotation détaillée des images d'archives constitue un complément précieux aux métadonnées institutionnelles. Les métadonnées CROBORA sont des descripteurs nouvellement ajoutés à chaque image d'archive par les participants du projet. Contrairement aux métadonnées institutionnelles qui se concentrent généralement sur un résumé du contenu vidéo dans son ensemble, notre objectif est d'établir un recodage focalisé exclusivement sur les images d'archives. Nous avons opté pour une description de chaque séquence selon quatre critères : personnalité, événement, lieu et illustration. L'anglais est utilisé pour les reportages en français et en italien. De plus, chaque séquence se voit at-

tribuer un identifiant unique (UID).

Concernant le dernier point, l'étiquetage automatique est un domaine bien développé dans l'analyse visuelle, avec de nombreux outils comme Distant Viewing Toolkit (DVT)² capables de produire automatiquement des métadonnées résumant le contenu (personnes/acteurs, dialogues, scènes, objets) et le style (angle de la caméra, éclairage, cadrage, son) des images. Par exemple, dans une étude réalisée sur de grandes collections historiques des Archives de la radio allemande, une approche de recherche de vidéos basée sur le contenu a été adoptée (Mühling et al., 2018).

Le but est d'attribuer automatiquement des tags sémantiques aux séquences vidéo. Cette méthode pourrait être très intéressante pour notre processus d'annotation. Néanmoins, nous n'avons pas trouvé de logiciel qui pourrait satisfaire nos besoins spécifiques. En raison de l'absence d'un thésaurus/lexique préparé et d'algorithmes formés dédiés aux représentations visuelles dans un contexte européen, la plupart des outils d'annotation automatique se concentrent sur une annotation de bas niveau axée sur des caractéristiques générales et plastiques (personnes, voiture, arbre, lumineux, portrait...), alors que nous avons besoin d'une forme d'annotation bien plus précise et figurative (Angela Merkel, le traité de Rome, Capitole, le contrôle des frontières, etc.). De plus, des millions d'images d'entraînement sont nécessaires pour construire un réseau neuronal convolutif (CNN) à partir de zéro, ce qui dépasse largement les limites de notre projet en termes de coût et de temps.

De plus, nous sommes confrontés à un autre défi majeur : notre intérêt se porte non sur la totalité des contenus vidéo, mais plutôt sur les courtes images d'archives, souvent dissimulées au sein de chaque vidéo. Cette particularité exige une approche manuelle et qualitative, nécessitant que les chercheurs examinent chaque reportage avec soin pour identifier et recueillir exclusivement ces images cachées. Ainsi, nous avons préféré opter pour la collecte et l'annotation manuelles des données audiovisuelles.

5 Analyse des données

Pour analyser les artefacts multimodaux collectés, il est crucial pour nous de recourir à divers outils numériques afin d'explorer notre corpus, qui se compose principalement d'images et de textes.

2. <https://github.com/distant-viewing/dvt>.

5.1 Analyse d'image

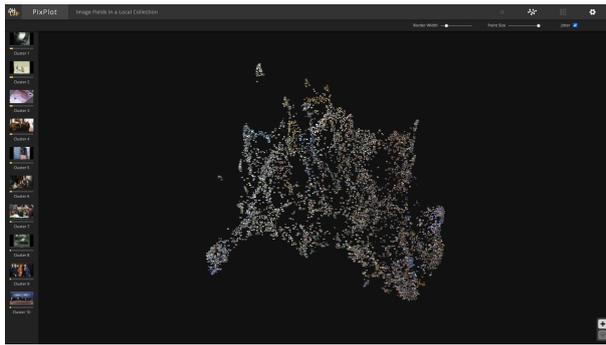


FIGURE 1 – La cartographie formée par les images d’archives collectées sur PixPlot.

Premièrement, pour obtenir une vision globale du corpus, nous avons utilisé plusieurs outils visuels. Par exemple, PixPlot³, un logiciel *open source* développé par le Yale Digital Humanities Lab (DHLab), fonctionne en effectuant un large ensemble de correspondances. Il utilise un réseau neuronal convolutif et la technique d’*Uniform Manifold Approximation and Projection* (UMAP) pour regrouper dans un navigateur web des images visuellement similaires (cf. fig. 1).

Les images, lorsqu’elles sont traitées par ces outils, se regroupent en clusters spatiaux basés sur des similitudes de couleurs, de formes, ou une combinaison des deux. Nous observons dans notre corpus une homogénéité marquée parmi les éléments visuels : des motifs visuels récurrents tels que le drapeau européen, les foules, les réunions et la population apparaissent de manière prédominante à travers l’ensemble du corpus. Cette récurrence révèle un intérêt éditorial commun pour des représentations centrées sur les assemblées publiques, les discussions politiques et les thèmes relatifs à la population dans un contexte européen. Par ailleurs, les images représentant des paysages urbains, des bâtiments, des arrivées, des usines et des migrants illustrent une couverture qui se concentre sur le développement urbain, les activités économiques, la mobilité, les secteurs industriels et les problématiques migratoires.

Bien que des outils comme PixPlot soient efficaces pour identifier des motifs visuels récurrents, ils présentent une contrainte notable : leur difficulté à capter les phénomènes uniques masqués par ces grandes tendances. Par exemple, notre analyse préliminaire a révélé que certaines séquences

3. <https://github.com/YaleDHlab/pix-plot>.

d’archives spécifiques, telles que la signature du traité de Rome, bien que peu nombreuses (69 occurrences) dans l’ensemble du corpus, ont marqué les deux décennies couvertes par notre étude. Lors des reportages sur l’histoire de la construction de l’UE, cette séquence a été fréquemment utilisée comme preuve de la construction historique. Cette observation met en évidence l’importance de ne pas se limiter aux tendances visuelles globales et de porter une attention particulière aux événements distinctifs qui peuvent influencer de manière significative l’imaginaire collectif.

5.2 Analyse de contenu

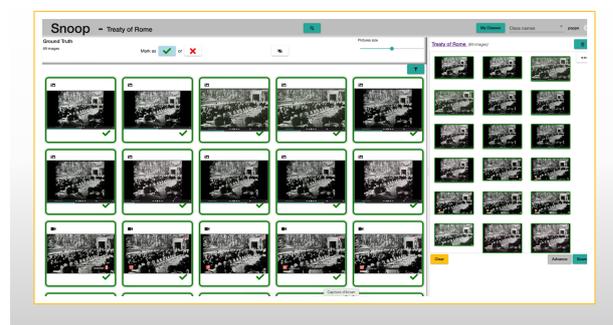


FIGURE 2 – La généalogie de la séquence d’archives sur la signature du traité de Rome construite sur Snoop.

Cette approche nous conduit à une autre méthode de détection basée sur le contenu, exécutée avec l’outil Snoop⁴ développé par l’INA. Ce moteur de recherche visuel interactif fonctionne avec une entrée initiale d’images sélectionnées par l’utilisateur, qui reflètent son intention de recherche. Snoop analyse ces images et permet à l’utilisateur de marquer les résultats visuels similaires fournis par le logiciel, soit comme positifs (correspondant à la recherche), soit comme négatifs (faux positifs). Cette méthode de recherche de similarité est intrinsèquement subjective, variant selon l’utilisateur et la situation spécifique. Snoop utilise ces retours d’informations pour affiner le modèle de l’intention de l’utilisateur. En intégrant des images d’archives que nous avons sélectionnées, nous avons utilisé Snoop pour analyser un corpus de 6 000 programmes d’actualités provenant du Web, totalisant environ 56 000 nouvelles vidéos, afin de retrouver des séquences d’archives spécifiques que nous avons identifiées *a priori* à partir du premier corpus,

4. <https://www.ina.fr/institut-national-audiovisuel/equipe-recherche/projet-snoop>.

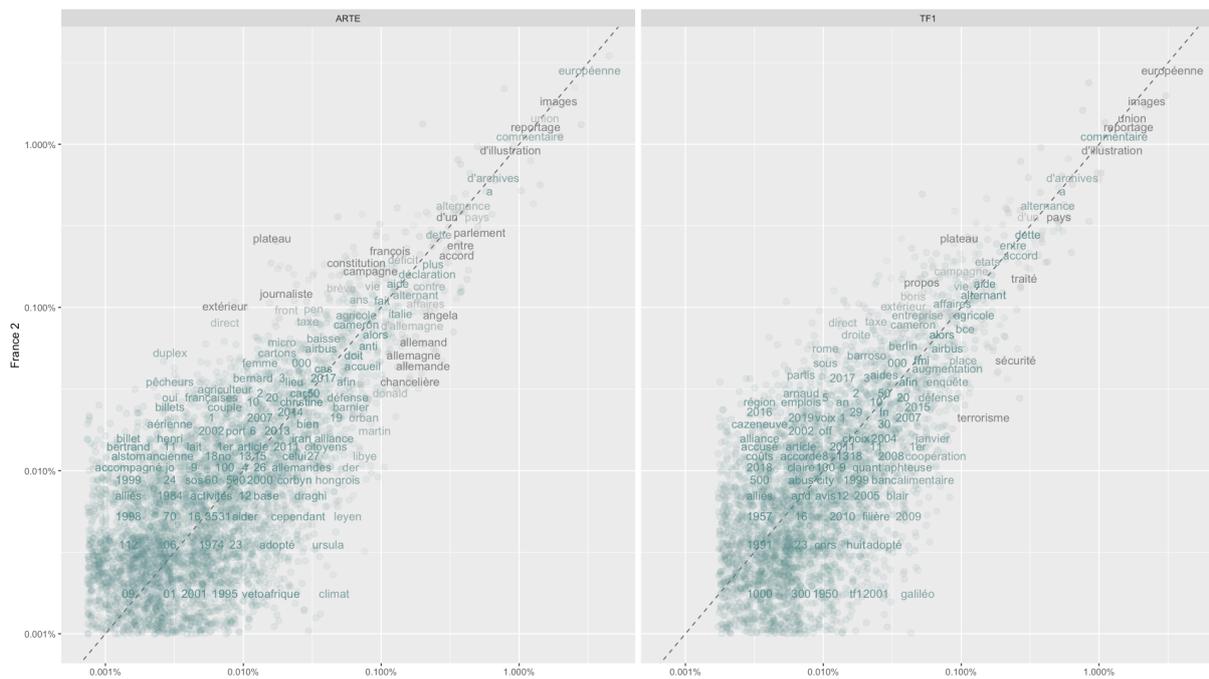


FIGURE 3 – La comparaison des fréquences de mots dans les métadonnées de TF1, France 2 et ARTE.

telles que celles de la signature du traité de Rome (cf. fig 2).

En appliquant systématiquement la même méthode à d'autres images d'archives sélectionnées, nous visons à tracer toutes les occurrences de ces séquences vidéo, permettant ainsi une étape supplémentaire de « lecture attentive » : l'examen du processus décontextualisation-recontextualisation pendant chaque réutilisation, tout en recherchant des tendances et des motifs.

5.3 Analyse des métadonnées

Comme l'exprime Bruno Bachimont (1998), les métadonnées sont en elles-mêmes des données. La description de l'audiovisuel ne se cantonne plus à des documents purement descriptifs, mais ces derniers deviendront eux-mêmes des éléments documentaires exploitables en parallèle avec les contenus audiovisuels qu'ils détaillent. Par exemple, nous avons analysé les métadonnées produites par l'INA pour plusieurs chaînes de diffusion (cf. fig 3). Notre analyse comparative des fréquences de mots dans ces chaînes révèle une standardisation marquée dans le langage employé par les documentalistes institutionnels. Cette uniformité apparaît clairement dans la façon dont les sujets sont catégorisés, le genre des contenus (reportage, documentaire, etc.) est défini, et la nature des images est décrite dans les métadonnées. En plus, pour évaluer de manière quantitative la similarité et les

différences dans les fréquences de mots, nous avons employé un test de corrélation. Les résultats, corroborés par nos graphiques, indiquent une forte corrélation entre ces chaînes.

5.4 Analyse de métadonnées CROBORA

Les métadonnées CROBORA enrichissent l'analyse des archives audiovisuelles, offrant une couche d'interprétation supplémentaire. Des techniques comme l'analyse TF-IDF, la corrélation de mots-clés et les analyses de réseau via R, Iramuteq ou encore Gephi révèlent les relations et tendances au sein du corpus. Par exemple, l'analyse de réseau des personnalités mentionnées dans les reportages montre comment elles interagissent et se positionnent dans le paysage médiatique.

Le graphe (cf. fig 4) montre un réseau dense de connexions avec certains nœuds plus centralement situés et ayant plus de connexions, suggérant que ces individus peuvent être des figures clés au sein du réseau. La taille des étiquettes des nœuds tels que « Angela Merkel », « Viktor Orbán » et « Jean-Claude Juncker » indique leur importance ou un degré plus élevé de connectivité au sein du réseau, impliquant qu'ils peuvent être des figures centrales ou influentes dans le contexte analysé. Les différentes couleurs des lignes pourraient représenter divers types de relations ou d'interactions, telles que les affiliations politiques, le travail collaboratif ou d'autres formes de connexions.

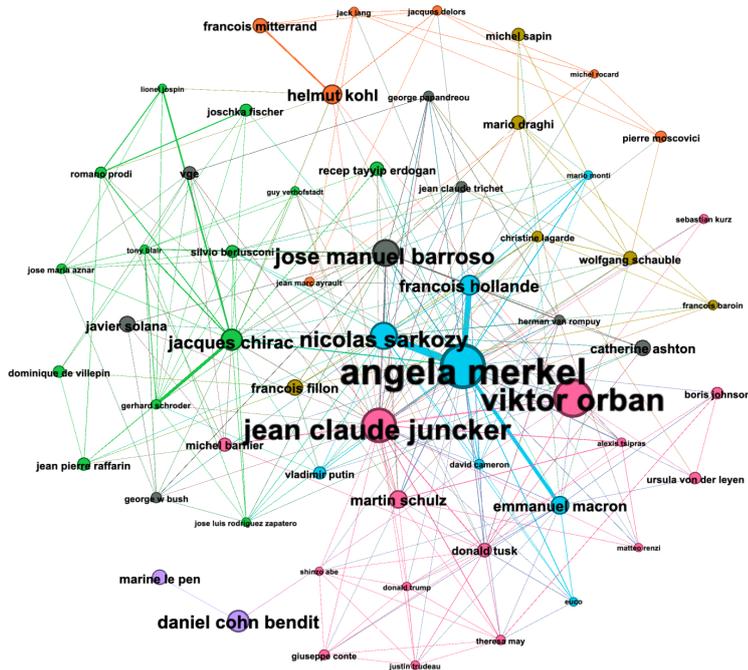


FIGURE 4 – La co-occurrence entre personnalités dans les métadonnées CROBORA.

6 Plateforme en construction

L'aboutissement du projet CROBORA sera son intégration dans une plateforme visuelle, actuellement en cours de développement par les membres de l'équipe du projet. Cette plateforme servira de ressource centralisée, offrant un accès fluide à une richesse de données, y compris des métadonnées et du contenu audiovisuel issus de la télévision française et italienne ainsi que du Web. Conçue en pensant à l'utilisateur final, la plateforme sera équipée de fonctionnalités avancées qui élèveront non seulement le processus analytique mais amélioreront également l'engagement global de l'utilisateur.

La plateforme permettra aux utilisateurs de naviguer dans le corpus en combinant différents mots-clés. Elle offrira également la possibilité de réaliser une analyse exploratoire des données grâce à deux fonctions distinctes : Distributions et ARViz. Nous intégrerons un logiciel nommé Muvin⁵ dans la fonction Distributions de notre plateforme. Muvin est une technique de visualisation innovante qui combine les graphiques de flux et les diagrammes de nœuds-liens pour faciliter la navigation dans

5. <http://dataviz.i3s.unice.fr/muvin/>.

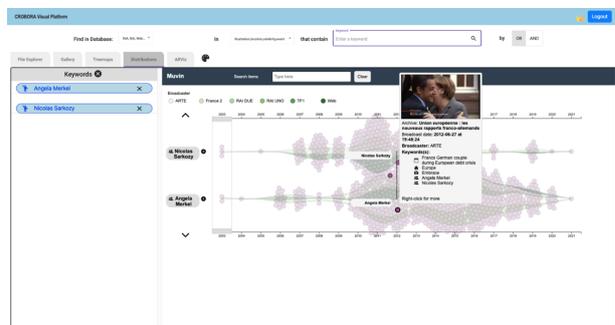


FIGURE 5 – Combinaison des mots-clés « Angela Merkel » et « Nicolas Sarkozy » à l'aide de la fonction Distributions.

les réseaux de collaboration multidimensionnels (Menin et al., 2022). Avec Muvin, les utilisateurs pourront explorer la co-occurrence de mots-clés dans les données CROBORA et suivre la progression chronologique de leur utilisation (cf. fig 5). En ce qui concerne la fonction ARViz⁶, elle sera conçue pour aborder le problème de l'exploration des règles d'association dans le corpus CROBORA. Notre interface utilisateur favorisera l'exploration

6. <http://covid19.i3s.unice.fr:8080/arviz>.

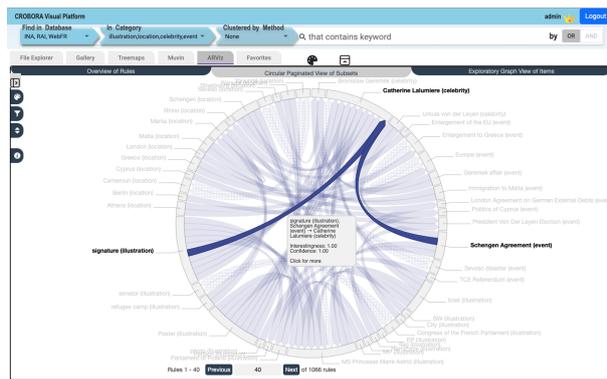


FIGURE 6 – Quelques règles d’association dans le corpus CROBORA à l’aide de la fonction ARViz.

des données, offrant une vue d’ensemble complète des règles via un graphique en nuage de points, une vue détaillée des sous-ensembles de règles à travers un diagramme en forme de corde, et des ensembles d’éléments via un graphique d’association (cf. fig 6) (Menin et al., 2021). Parallèlement, les utilisateurs peuvent importer et annoter de nouveaux contenus audiovisuels, tout en explorant leurs annotations grâce aux outils disponibles sur la plateforme, dans leur espace utilisateur personnalisé.

7 Conclusion

Cette communication détaille le projet CROBORA, qui explore la diffusion des archives audiovisuelles sur la construction de l’Union européenne dans les médias français et italiens. Le projet se penche sur la manière dont ces images façonnent l’imaginaire visuel collectif. Les données collectées de 2001 à 2021, tant de la télévision que du Web, comprennent 37 006 images d’archives. Le corpus inclut des vidéos, des images clés de chaque séquence d’archives, et des métadonnées enrichies. Ces métadonnées, à la fois originales et complétées par l’équipe CROBORA, offrent une description détaillée de chaque séquence. Pour l’analyse, des outils numériques tels que PixPlot, pour regrouper des images similaires, et Snoop, un moteur de recherche visuel, ont été employés. Parallèlement, l’étude souligne le rôle crucial des métadonnées comme ressources exploitables, et des analyses quantitatives révèlent des motifs récurrents ainsi que des événements distinctifs significatifs. Enfin, le projet intègre une plateforme utilisateur-centrée qui facilite une exploration approfondie des données à l’aide de fonctions telles que Muvin et ARViz, permettant aux utilisateurs de naviguer

dans le corpus via différents mots-clés.

Pour l’avenir, il serait enrichissant de comparer nos données à celles issues de médias alternatifs ou locaux, afin d’avoir une compréhension plus approfondie. Cette approche pourrait dévoiler des perspectives nouvelles sur la culture contemporaine à travers l’analyse des données audiovisuelles.

Bibliographie

- Taylor Arnold et Lauren Tilton. 2019. *Distant viewing: analyzing large visual corpora*. *Digital Scholarship in the Humanities*, 34(Suppl. 1) :i3–i16.
- Taylor Arnold et Lauren Tilton. 2023. *Distant Viewing: Computational Exploration of Digital Images*. The MIT Press.
- Taylor Arnold, Lauren Tilton, et Annie Berke. 2019. *Visual style in two network era sitcoms*. *Cultural Analytics*, 4(2).
- Taylor Arnold, Lauren Tilton, et Justin Wigard. 2022. *Automatic identification and classification of portraits in a corpus of historical photographs*. In *Proceedings of the Computational Humanities Research Conference 2022*. CEUR Workshop proceedings.
- Taylor Arnold, Lauren Tilton, et Justin Wigard. 2023. *Understanding peanuts and schulzian symmetry: Panel detection, caption detection, and gag panels in 17,897 comic strips through distant viewing*. *Journal of Cultural Analytics*, 8(3).
- Bruno Bachimont. 1998. *Bibliothèques numériques audiovisuelles: des enjeux scientifiques et techniques*. *Document numérique*, 2(3-4).
- John A. Bateman. 2014. *Text and Image : A Critical Introduction to the Visual/Verbal Divide*. Routledge.
- David M. Berry. 2011. *The computational turn: Thinking about the digital humanities*. *Culture Machine*, 12.
- Tuomo Hiippala. 2021. *Distant viewing and multimodality theory: Prospects and challenges*. *Multimodality & Society*, 1(2).
- Guillaume Latzko-Toth, Claudine Bonneau, et Mélanie Millette. 2017. *Small data, thick data : Thickening strategies for trace-based social media research*. *SAGE Handbook of Social Media Research Methods*, pages 199–214.
- Lev Manovich. 2020. *Cultural Analytics*. The MIT Press.
- Aline Menin, Michel Buffa, Maroua Tikat, Benjamin Molinet, Guillaume Pelerin, Laurent Pottier, Franck Michel, et Marco Winckler. 2022. *Incremental and multimodal visualization of discographies: exploring the wasabi music knowledge base*. In *WAC 2022 - Web Audio Conference 2022*. WAC Proceedings 2022.
- Aline Menin, Lucie Cadorel, Andrea Tettamanzi, Alain Giboin, Fabien Gandon, et Marco Winckler. 2021. *Arviz: Interactive visualization of association rules*

- for rdf data exploration. In *2021 25th International Conference Information Visualisation (IV)*. IEEE Xplore.
- Franco Moretti. 2013a. *Distant Reading*. Verso.
- Franco Moretti. 2013b. “Operationalizing”: or, the function of measurement in modern literary theory. *New Left Review*.
- Markus Mühling, Manja Meister, Nikolaus Korfhage, Jörg Wehling, Angelika Hörth, Ralph Ewerth, et Bernd Freisleben. 2018. Content-based video retrieval in historical collections of the german broadcasting archive. *International Journal on Digital Libraries*, 20 :167–183.
- Shiming Shen, Matteo Treleani, Dario Compagno, et Marco Winckler. 2023. From stock shots to ghost data: Tracking audiovisual archives about the european union. *VIEW Journal of European Television History and Culture*.
- David A. Smith, Ryan Cordell, et Abby Mullen. 2015. Computational methods for uncovering reprinted texts in antebellum newspapers. *American Literary History*, 27(3) :E1–E15.
- Thomas Smits et Melvin Wevers. 2023. A multimodal turn in digital humanities. using contrastive machine learning models to explore, enrich, and analyze digital visual historical collections. *Digital Scholarship in the Humanities*, 38(3) :1267–1280.
- Caroline Suen, Sandy Huang, Chantat Eksombatchai, Rok Susic, et Jure Leskovec. 2013. Nifty: a system for large scale information flow tracking and clustering. In *WWW '13 : Proceedings of the 22nd international conference on World Wide Web*, page 1237–1248, Rio de Janeiro, Brésil.
- Matteo Treleani, Shiming Shen, Aline Menin, et Marco Winckler. 2023. Circulation et répétition des images de stock dans les médias audiovisuels
la représentation visuelle des migrants dans les jt européens. *Les Cahiers du Numérique*.
- Melvin Wevers et Thomas Smits. 2019. The visual digital turn: Using neural networks to study historical images. *Digital Scholarship in the Humanities*, 35(1) :194–207.