

Décrypter le présent au prisme de la mémoire audiovisuelle et numérique : **les collections et les données de l'INA** au service de la Recherche

Média patrimonial de service public, l'Institut national de l'audiovisuel a pour singularité de conjuguer ses missions de collecte, conservation, traitement, valorisation et exploitation des archives dans une démarche de décryptage du présent et de création audiovisuelle, de transmission des savoirs et de recherche technologique.

Fortes de près de 28 millions d'heures de contenus TV et radio, de 2 millions de photographies, de plus de 3 km linéaires d'archives écrites, rassemblant plus de 17 000 sites web pour 140 milliards de versions d'URL depuis 1996, 16 500 comptes Twitter pour 3 milliards de tweets collectés et 11 600 comptes de plateformes pour 39 millions de vidéos web archivées, les collections de l'Institut national de l'audiovisuel (INA), et la gestion de masse de la data qu'elles entraînent, représentent aujourd'hui un outil précieux pour l'analyse des écosystèmes médiatiques.

L'historique de leur constitution est marqué par la fin de l'ORTF (1975), suivie des lois successives instaurant le dépôt légal de la radiotélévision (1992), puis du Web (2006). À l'éclatement de l'ORTF, l'INA hérite de l'ensemble du fonds constitué depuis les débuts effectifs de la radio et de la télévision en France. Ceux-ci remontent pour la radio

La même logique de réemploi des images se traduit également dans les statuts de l'INA, depuis toujours, par une mission de distribution et commercialisation adossée à la conservation des archives, et étendue aux fonds tiers déjà acquis par l'ORTF, comme celui des Actualités françaises (1940-1969), fonds de presse filmée à destination des salles de cinéma. L'articulation de ces missions, patrimoniale et commerciale, a conduit de nombreux producteurs de contenus audiovisuels à confier leur catalogue à l'Institut, sous forme de don, dépôt, ou mandat.

LES ARCHIVES AUDIOVISUELLES, SOURCES À PART ENTIÈRE POUR LA RECHERCHE

L'expertise développée par l'Institut en matière de sauvegarde et de valorisation des archives audiovisuelles lui vaut très tôt d'être

tion d'un savoir critique, la loi sur le dépôt légal confère à l'archivage de la radiotélévision une nouvelle dimension, et crée les conditions d'une possible objectivation des médias audiovisuels. Cette nouvelle dimension se traduit également par l'accroissement de son périmètre, à mesure que le paysage audiovisuel français se décline en régions métropolitaines et ultramarines, sur le câble, le satellite et la TNT : aujourd'hui plus de 180 chaînes TV et radio sont captées 24 h/24h. La loi DADVSI, promulguée le 1^{er} août 2006, permet à l'INA d'accompagner le déploiement des médias audiovisuels sur le Web, pensé dès les années 2000 comme un prolongement en ligne de la radio et de la télévision, comme un antidote à la possible érosion des médias traditionnels. Grâce à un travail assidu de veille prospective, indispensable compte tenu de la nature réactive et volatile des modes de publication sur le Web, l'INA collecte en continu les publications textuelles et vidéo liées aux médias audiovisuels, quel que soit leur format de diffusion : sites Web, médias *pure players*, podcasts et flux de Web radio, plateformes et réseaux sociaux, intégrant la collecte des mots-dièse les plus représentatifs des phénomènes médiatiques.

DES DÉFIS SPÉCIFIQUES À LA CONSERVATION ET LA MÉDIATION DES DOCUMENTS AUDIOVISUELS

La construction de ce périmètre, unique par la cohérence, la continuité et la profondeur historique qu'il offre — des médias audiovisuels traditionnels aux plus innovants — ne va pas sans un certain nombre de défis. Défis technologiques, auxquels l'INA a répondu en recourant à des outils très finement adaptés à la diversité des médias audiovisuels et capables de les lire dans le temps indépendamment de leur technologie native. Le choix d'une interface d'accès qui ne soit pas figée mais capable d'évo-

L'INA a créé des outils documentaires capables de restituer les logiques de diffusion propres aux objets audiovisuels.

à la fin des années 1930, et au lendemain de la Seconde Guerre mondiale pour la télévision. Leur archivage, lui, naît relativement tôt du besoin des chaînes de réutiliser certains documents pour illustrer ou nourrir en rediffusion les grilles de programmes. Dès 1952, la collecte des premiers « bobineaux films » permet ainsi de former, au sein de la RTF d'alors, un embryon de cinémathèque. Il ne s'étoffera véritablement qu'en 1970 avec la création du Service central de la conservation des archives audiovisuelles, et sa transposition, inscrite dans les statuts fondateurs de l'Institut. La mission d'archivage que met en œuvre l'INA à compter de 1975 se structure autour de la collecte continue des programmes diffusés chaque jour par les sociétés nationales de production.

Gustave Eiffel en 1891, conservé dans le fonds des Archives de la Parole; par leur rareté, à l'instar des 265 heures d'archives du procès de Rivonia¹ enregistrées sur des dictabelts, supports cylindriques en vinyle souple, qui ont pu restituer la plaidoirie de Nelson Mandela et des membres du Congrès national africain en 1963; ou encore par leur singularité, à l'image du fonds dit des « Noticiéros »², films d'actualité tournés par l'ICAIC entre 1960 et 1990 et diffusés chaque semaine à Cuba.

Autre moment fondateur des collections de l'INA, l'instauration du dépôt légal de la radiotélévision par la loi du 20 juin 1992 consacre les archives audiovisuelles comme sources à part entière pour la recherche. En lui attribuant une finalité de constitu-

le dépositaire de documents remarquables par leur ancienneté, tel l'enregistrement de la voix de



luer au fil des mutations de l'écosystème audiovisuel et numérique, permet, depuis la mise en œuvre du dépôt légal, de déjouer l'obsolescence des formats et de garantir la « rejouabilité » des archives.

Défis méthodologiques également : avec la conviction que l'accès aux archives ne saurait suffire et qu'une contextualisation était indispensable, l'INA a créé des outils documentaires capables de restituer les logiques de diffusion propres aux objets audiovisuels, et a effectué une collecte systématique des métadonnées les accompagnant. Ainsi, outre la constitution d'une collection incontournable pour la mémoire des médias, l'archivage du dépôt légal a permis l'agrégation et la création d'une masse unique de données descriptives de référence, premières garantes de la découvrabilité des documents archivés. La réflexion sur la mise à disposition des fonds du DL a été menée en parallèle de la collecte, depuis l'ouverture de l'INAtèque en 1998, où sont mis à disposition les fonds et des outils d'interrogation « experte », conçus au plus près des besoins des usagers, dans le cadre d'ateliers méthodologiques. Dans les années 2010, pour couvrir le territoire, 50 postes de consultations décentralisés ont été installés en métropole et en Outre-mer, notamment dans des médiathèques régio-

nales. Enfin, en 2023, la création du « lab » est venue compléter l'offre de services à la recherche.

UN « LAB » DÉDIÉ À LA FOUILLE DE DONNÉES

Pour suivre des besoins en constante évolution, toujours en articulation étroite entre son service de la Recherche et ses communautés d'usagers, l'INA a successivement développé des moyens d'accès aux contenus (numérisés ou nativement numériques), aux jeux de données, mais également aux outils savants développés par l'INA, d'analyse, de segmentation et d'annotation de l'image et du son. L'essor des humanités numériques, le « tournant data » que connaît la recherche, ont conduit l'INA à structurer l'accompagnement de ces nouvelles modalités de fouilles de données dans un lab, pensé comme une passerelle entre les savoir-faire technologiques et documentaires de l'Institut. La richesse des métadonnées et les innovations de l'INA pour leur traitement automatique par des outils d'intelligence artificielle forment ainsi le socle des services proposés par le lab à la recherche universitaire.

Les collections, dont la profondeur historique offre la possibilité d'analyses longitudinales, de focales sur certaines périodes ou de comparaisons dans le temps long de

l'histoire, constituent une source singulière pour l'ensemble des sciences humaines et sociales, mais également pour les disciplines des sciences appliquées, qui trouvent dans les jeux de données disponibles un terrain fécond d'expérimentation.

La pluridisciplinarité du lab en fait un lieu de dialogue et de collaboration dans l'exploration des nouvelles frontières de la recherche. C'est le sens donné au dispositif d'appels à projets³ du lab. En 2024, quatre projets lauréats sont incubés et bénéficient d'un accompagnement technologique et méthodologique. Les questions qu'ils soulèvent touchent à des domaines aussi divers que l'évaluation du temps de parole des femmes dans les podcasts natifs français, la circulation et la reprise de la parole des candidats à la présidentielle de 2022, dans les médias audiovisuels et sur les réseaux sociaux, la catégorisation des inégalités dans les médias, ou encore l'impact des journaux télévisés sur les prévisions économiques des ménages. L'ensemble des projets accompagnés par le lab depuis sa création en 2023 (4 par an), présentés à rythme régulier dans son séminaire, constituent autant d'approches méthodologiques innovantes, qui viennent renouveler notre lecture du présent, en décryptant le débat public à travers le temps long des pratiques et des discours médiatiques.

MILEVA STUPAR

Directrice des patrimoines de l'INA
mstupar@ina.fr

[1] <https://www.ina.fr/actualites-ina/l-ina-artisan-de-la-memoire-de-nelson-mandela>

[2] <http://www.inatheeque.fr/publications-evenements/publications-2022/noticiero-icaic-30-ans-d-actualit-s-cin-matographiques-cuba.html>

[3] <https://inalelab.hypotheses.org/6808>

