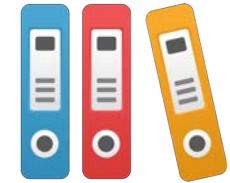


Dépôt Légal du Web de l'Ina

novembre 2025



Institut National de l'Audiovisuel



Conservation des archives
audiovisuelles, Formation, et
Recherche

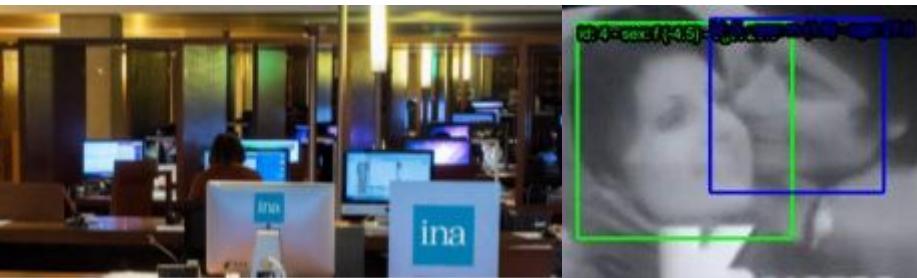


1974

Recherche sur
les médias

1992

Dépôt légal



1995

INA thèque

2009

Archivage du web

2022

le lab

1999

Numérisation

2004

INA mediapro

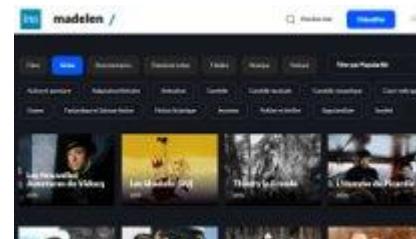
2020

mediaclip

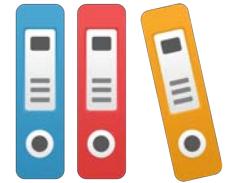
2006

Ina.fr Youtube | Facebook | Twitter | Instagram madelen data.ina.fr

L'INA, producteur et
éditeur de contenus



Dépôt Légal de l'Internet : Loi DADVSI, 2006

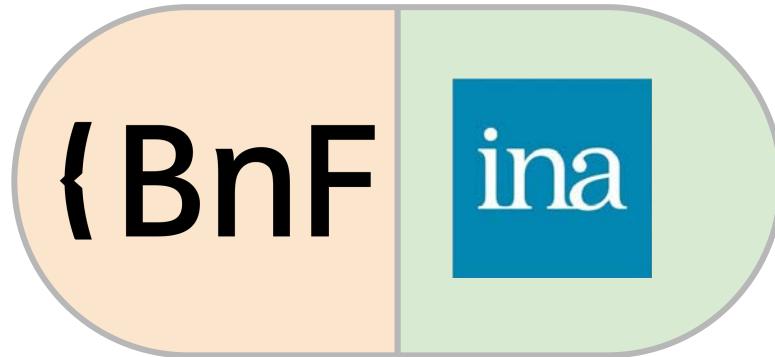
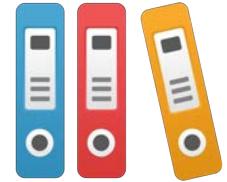


Extension du dépôt légal à l'Internet

Transposition de la Directive Européenne 2001/29/EC sur l'harmonisation des droits d'auteurs et droits voisins

Extension du dépôt légal aux «signes, signaux, écrits, images, sons ou messages de toute nature faisant l'objet d'une communication au public par voie électronique»

Dépôt Légal de l'Internet : Décret, 2011



- Le périmètre de l'Ina, parmi les sites français, correspond aux :
- Sites émanant des services des médias audiovisuels ;
- Web TV et Web radios ;
- Sites principalement consacrés aux programmes radio et télé ;
- Sites des organismes de l'environnement professionnel et institutionnel du secteur de la communication audiovisuelle.

La BnF archive les autres sites français.



Dépôt Légal INA : Consultation

Consultation
experte

7 lieux
Inathèque
Délégations
régionales

Consultation
autonome

50 PCMs
Bibliothèques
partenaires

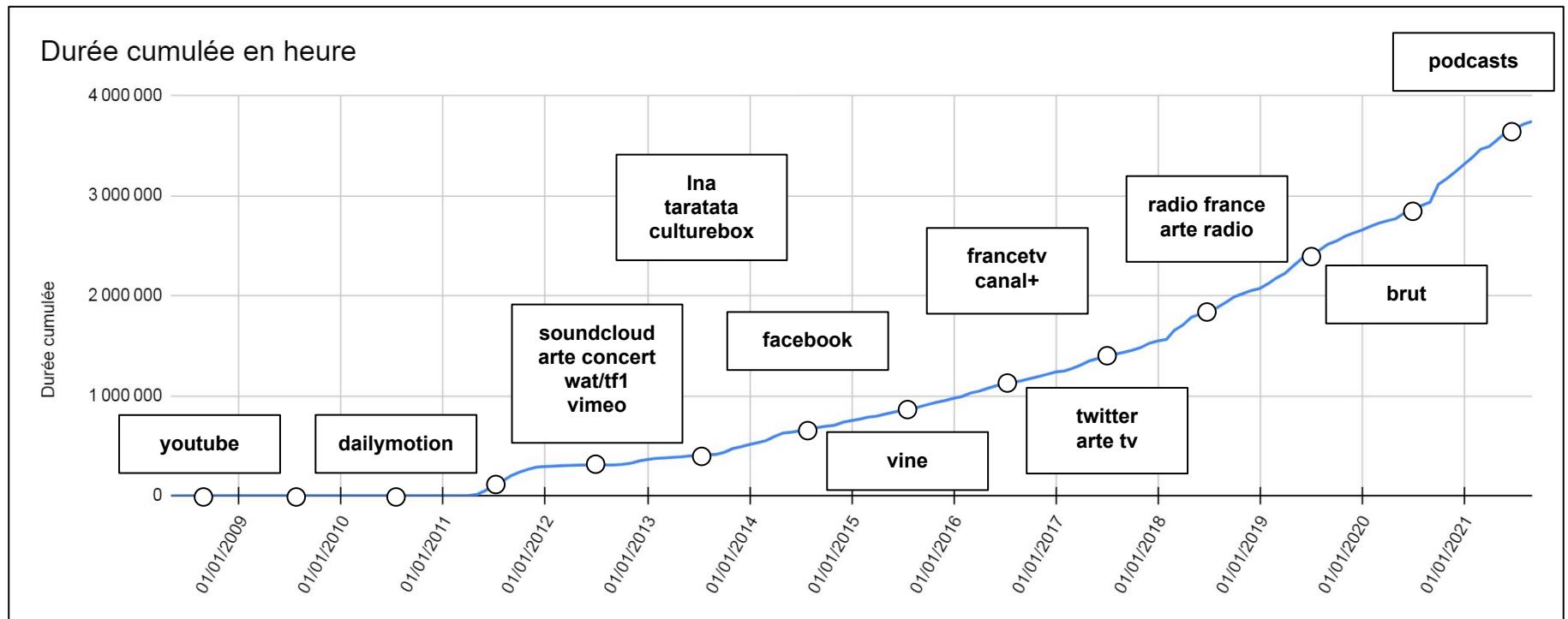


Périmètre : nomination et documentation

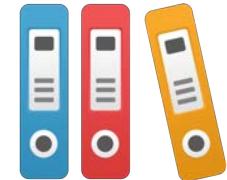


- Périmètre :
 - En lien avec la Télévision et la Radio en France
 - Médias audiovisuels natifs Français
 - Publications liées à des évènements à forte résonance en France
- Type de sources :
 - Sites web
 - Chaînes vidéo
 - Comptes et hashtags de réseaux sociaux
 - Programmes de Replay TV et vidéos natives
 - Emissions de Replay Radio et podcasts natifs

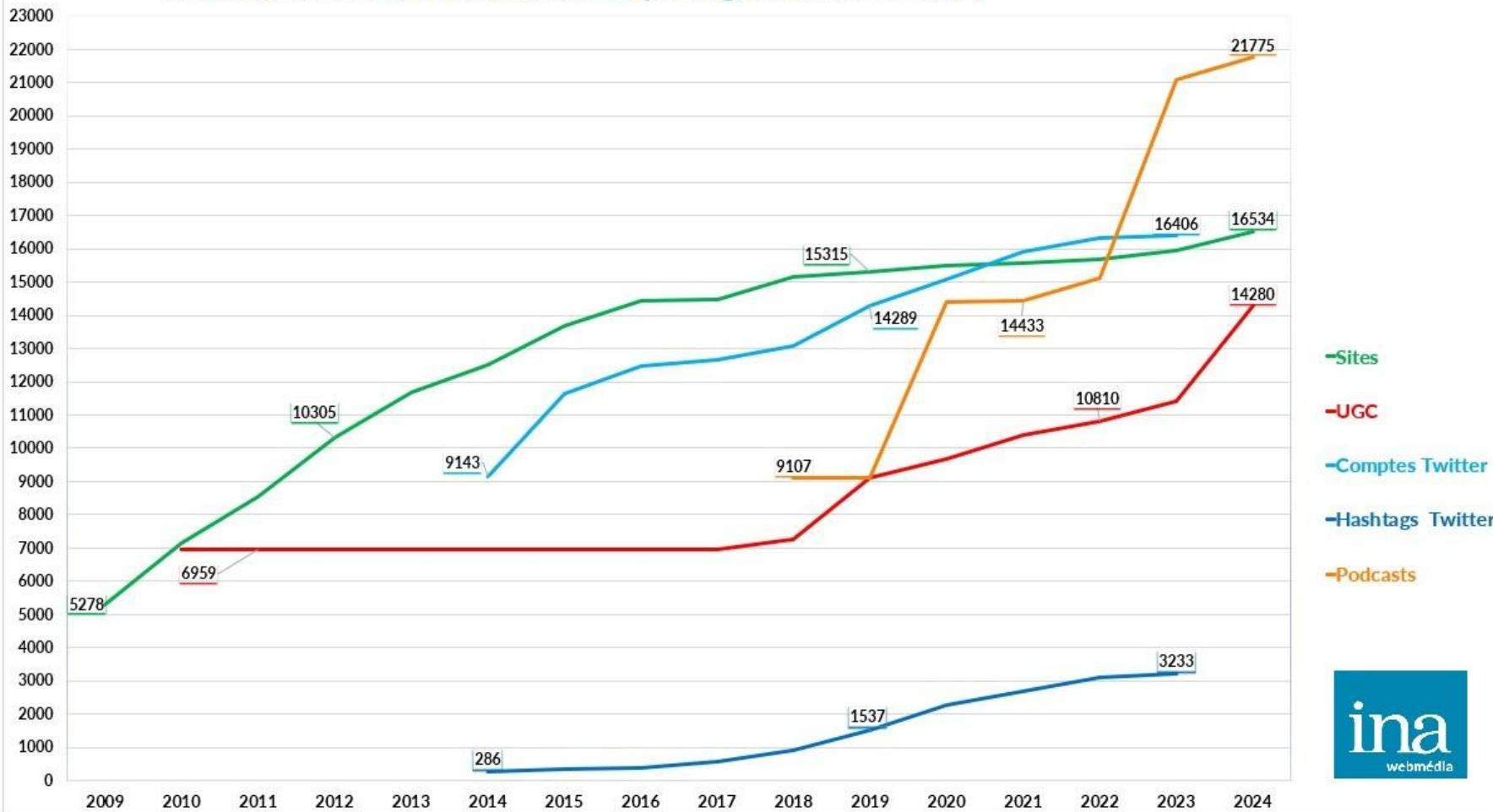
Du web des sites au web des plateformes



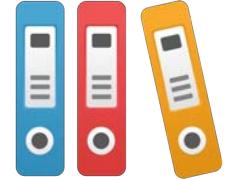
Evolution des sources de la collecte



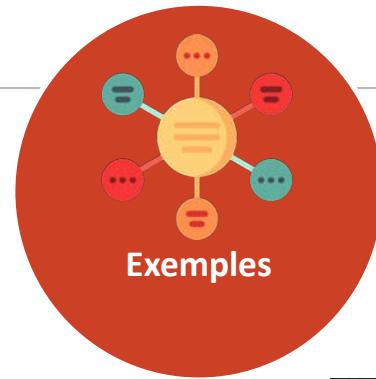
Évolution annuelle du nombre d'objets web inscrits à la collecte et consultables dans l'archive du Dépôt légal du web de l'INA



Archive des sites



170 milliards d'urls, dont 20 milliards de pages à partir de 29 000 sites web :
- éditeurs TV/Radio
- sites professionnels
- sites & blogs de fans



france•tv

6play

Brut.

| MYTF1 |

Chérie
FM

TLM

S*

grenouille
88.8fm
grenouille888.org

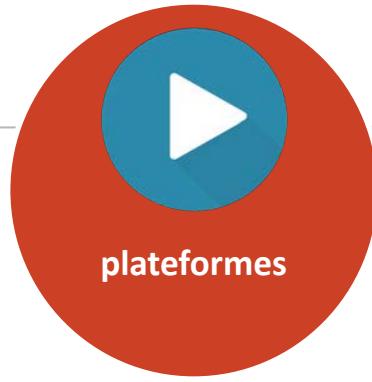
ina

Archive des vidéos



42 millions de vidéos sur 15 000 comptes ou chaînes sur 15 plateformes
- TV/Radio replay/exclus
- Média web natifs
- Youtubers populaires

5 millions d'heures



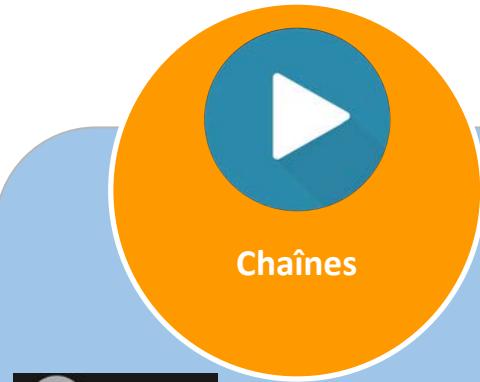
france.tv

YouTube

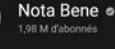
arte
dailymotion

| MYTF1 |
wat

Brut.



Nota Bene



1,98 M d'abonnés

AFPO

LEMPDA

SQUEEZIE

Mcfly et Carlito

LE HUFFPOST



564 k abonnés

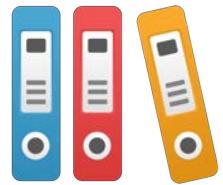


ina

Collecte YouTube limitée depuis septembre 2024

2 milliards d'heures

Archive des audios & podcasts

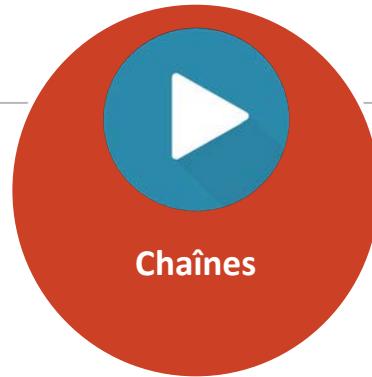


Audio

5 millions de sons
sur 30 000 chaînes ou
émissions
- Replay Radio
- podcasts indépendants et
natifs

2 millions d'heures

5 millions d'heures



Chaînes



arte

nova

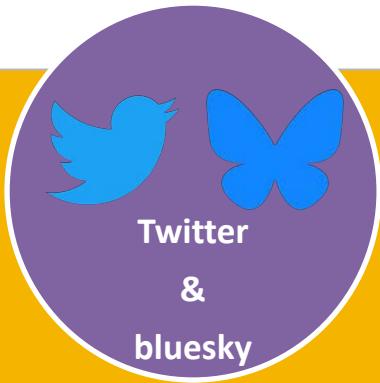
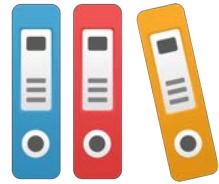
:le billet sciences



radiofrance



Archive des réseaux sociaux



Twitter :

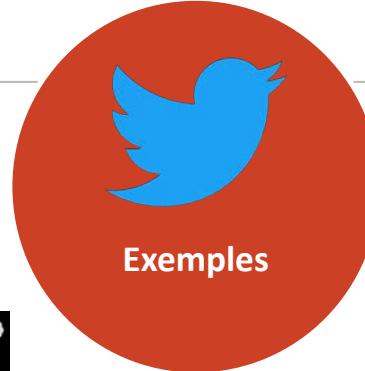
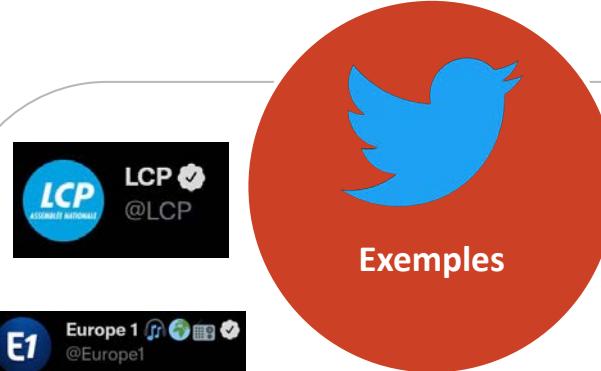
3.1 milliards de tweets à partir de 13 500 comptes & 2 700 hashtags :

- chaînes, émissions et personnalités TV/Radio
- actualités des médias
- grands évènements
- comptes @elysee, PR

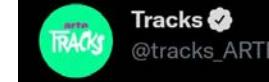
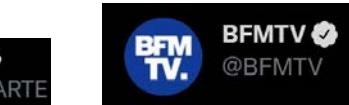
Bluesky :

Comptes des TVs et Radios

Comptes des TVs et Radios
Bluesky :



Exemples



#jesuischarlie

#13hfoot

#GiletsJaunes

#NuitDebout

#KohLanta

#tpmp

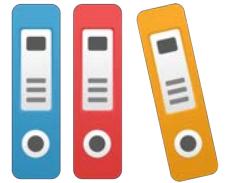
#Covid19

Collecte Twitter stoppée en juillet 2023

Collecte Bluesky démarrée en juin 2025



Dépôt Légal et grandes plateformes

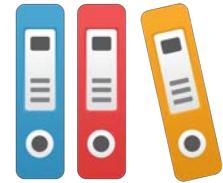


Aujourd'hui, la loi de dépôt légal ne s'applique pas aux acteurs bénéficiant du statut d'hébergeur.

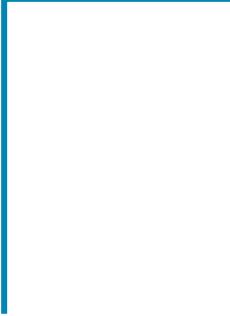
Ce qui exclut pour l'heure les grandes plateformes numériques de devoirs envers le dépôt légal en France.

L'INA n'a pas de facilité particulière d'accès aux données de ces plateformes, et peut-être empêché de collecter.

Projets en cours



- Collecte de Contenus (S)VOD :
 - Automatisation de screen recording
 - problématiques de fiabilité, fidélité et qualité à estimer et parer si possible
 - Test de contournement des DRMs ?
- Collecte de Vidéos Live :
 - Live Twitch, Youtube
 - Fast TV, ...
- Mise à jour des applications et moteurs de recherche
 - Consultation de l'archive des médias du web
 - Consultation de l'archive des réseaux sociaux
 - Intégration des données Bluesky à la consultation

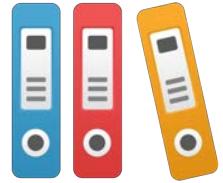


Démo

Webmédia



Consultation des sites & pages web



Navigation : URL https://versions.dlweb/search?q=http://www.france2.fr&version=1420110684

ina webmédia/ http://www.france2.fr/ Explorer

Navigation URL

Sites Chaînes Comptes Hashtags Doc diffuseurs Corpus

Recherche Lab

Besoin d'aide ? Aide contextuelle Dernière mise à jour 24/02/2025

229 868 versions archivées pour cette URL france2.fr

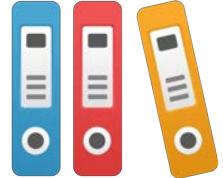
Année	Nombre de versions
2025	1 219
2024	12 056
2023	15 167
2022	14 885
2021	14 248
2020	13 883
2019	12 644
2018	14 112
2017	17 501
2016	24 750
2015	25 515
janvier	1 251
février	1 181
mars	1 972

0h 1h 2h 3h 4h 5h 6h 7h 8h 9h 10h 11h 12h 13h 14h 15h 16h 17h 18h 19h 20h 21h 22h 23h

01	0:04	2:01	3:01	4:19	5:28		7:11	8:03	-	10:00	12:22	13:21	14:13	15:06	16:08	17:12	18:32	19:14	20:12	21:22	22:04			
02	0:05	-	2:05	4:05	-		7:21	8:03	-	10:07	-	12:12	13:13	14:42	15:01	-	17:34	18:27	19:23	20:11	21:04	22:19	-	
03	0:12	-	2:05	3:13	4:24	-	7:03	8:02	-	10:22	-	12:25	13:08	14:13	15:09	-	17:15	18:02	19:17	20:18	21:04	22:19	23:04	
04	0:00	-	2:01	3:34	4:20	5:06	6:42	7:48	8:47	9:18	-	11:05	12:29	13:09	14:21	15:08	16:22	17:40	18:10	19:06	20:08	21:34	22:30	
05	0:13	-	2:07	3:03	4:26	5:26	-	7:04	8:10	-	10:01	-	12:24	13:01	14:24	15:04	16:09	17:16	18:18	19:11	20:15	21:03	22:17	
06	0:13	1:13	2:49	3:03	4:00	5:03	6:34	7:25	-	9:01	10:41	11:05	12:37	13:02	14:06	15:19	16:26	17:29	18:21	19:19	20:04	21:16	22:00	
07	0:32	-	2:31	3:09	4:07	5:18	6:36	7:31	8:18	9:44	10:33	11:14	12:03	13:26	14:08	15:02	16:40	17:34	18:06	19:10	20:20	21:13	22:06	23:03
08	0:02	-	2:01	4:15	5:03	6:43	7:09	8:17	-	10:21	-	12:02	13:18	14:13	15:55	16:09	17:06	18:15	19:15	20:56	21:01	22:01	23:20	
09	0:25	-	2:04	3:24	4:19	5:00	-	7:09	8:01	-	10:00	-	12:01	13:03	14:06	15:04	-	17:21	18:06	19:08	20:08	21:07	22:05	
10	0:04	-	2:01	3:22	4:26	5:17	-	7:25	8:03	-	10:00	-	12:10	13:03	14:00	15:01	16:01	17:06	18:00	19:02	20:01	21:03	22:01	
11	0:06	-	2:02	3:01	4:10	-	-	7:07	8:07	-	10:04	-	12:01	13:02	14:03	15:03	16:51	17:06	18:01	19:15	20:02	21:02	22:01	
12	0:01	-	2:05	3:03	4:46	5:31	-	7:16	8:10	9:14	10:09	-	12:03	13:05	14:00	15:07	16:30	17:04	18:03	19:02	20:02	21:01	22:01	
13	0:00	-	2:06	3:01	4:27	-	-	7:13	8:21	-	10:12	-	12:02	13:02	14:08	15:02	16:08	17:22	18:06	19:00	20:09	21:08	22:13	
14	0:02	-	2:01	3:12	4:16	-	6:51	7:17	8:39	-	10:18	-	12:09	13:05	14:02	15:07	-	17:08	18:00	19:06	20:01	21:03	22:12	

ina

Consultation des sites & pages web



ukraine guerre

Entre le dd-mm-yyyy et le dd-mm-yyyy sur ex: tf1.fr trié par les plus pertinentes

valeurs relatives

1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025

4 220 296 résultats

[guerre Ukraine -](https://vl-media.fr/tag/guerre-ukraine/)

vl-media.fr

26-03-2022 13:32:25 – totalement dégénéré ces dernières heures entre la Russie et l'Ukraine, focus sur les cinq questions à se poser sur cette guerre "surprise". Que s'est-il passé ces derniers jours ? L'hypothèse d'une... Soyez informés Categories (...) 9 mars 2022 En plein contexte de guerre, Alexis Corbière s'est retrouvé dans une situation délicate après avoir évoqué partiellement le contenu des (...) livraisons d'armes françaises à l'Ukraine en direct. Alexis Corbière était l'invité de Marc Fauve... 24 février 2022 Alors que la situation a

voir tous les résultats pour ce site (1 513 ou plus)

<http://www.france-ukraine.com/La-guerre-n-a-pas-un-visage-de.html>

'La guerre n'a pas un visage de femme', de Svetlana Alexievitch - France-Ukraine.com

28-09-2010 16:57:58 – 'La guerre n'a pas un visage de femme', de Svetlana Alexievitch Le 20 novembre 2004 Svetlana Alexievitch est née en Ukraine, en 1948. Journaliste

voir tous les résultats pour ce site (47 970 ou plus)

<https://www.cnews.fr/guerre-en-ukraine>

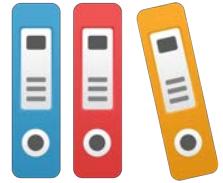
Guerre en Ukraine | CNEWS

cnews.fr

02-12-2022 02:19:40 – Guerre en Ukraine Retrouvez toute l'actualité autour de la guerre en Ukraine sur CNEWS

voir tous les résultats pour ce site (22 214 ou plus)

Consultation des sites & pages web



The screenshot shows the France 2 website (france2.fr) with a blue sidebar on the left containing navigation links like 'Navigation', 'Recherche', 'Lab', and 'Information'. The main content area features a large video player for 'Télématin' news. Below it, there's a section for 'POINT ROUTE' and other news highlights. The right side has a sidebar for 'INFO FRANCE 2' and a 'VIDÉO' section with links to 'Vidéo Service', 'Vidéo Globe Sud-Ouest', 'Vidéo JT', 'Vidéo Ski Challenge', and 'Vidéo Europe Foot'.

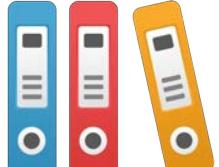
The screenshot shows the France 2 television website. The top navigation bar includes links for 'Accès direct à la chaîne', 'Programme TV', 'Films', 'Vidéos', 'Amp', 'Jeux', and 'Partenaires'. A search bar is present, along with a link to 'Répondez à la crise' and a 'Suivez-nous' section. On the left, there's a sidebar for 'Navigation', 'Recherche', 'Lab', and 'Date de diffusion' (with a link to '04-06-2013 11:06:43'). Below that is a 'Date de diffusion' section for '05-06-2013 11:06:43'. At the bottom left is an 'Information' section with links to 'Accès direct à la chaîne', 'S'abonner', 'Paramètres', and 'Aide'. The main content area features a large video player showing a tennis player in action, with a red overlay for 'ROLAND GARROS'. Below it, a section titled 'VOS SOIREEES' lists two programs: '20h55' and '23h05'. Further down, sections for 'LES JT sur franceinfo' and 'VOIR ET REVOIR' are shown, each with thumbnail images of news anchors.

A screenshot of the France 2 website. The main header says "France 2 Replay" and "france-tv". Below it is a large image for the TV show "LA SAINTE FAMILLE" featuring two men and a woman. A play button icon is overlaid on the image. To the right of the image, the show's title is displayed in large letters. Below the title, a subtitle reads "Un homme coincé entre le D de désir et le D de Dieu". At the bottom of the main image, there are five smaller thumbnail images for other shows: "Généalogie", "Le Journal de l'info", "L'Émission politique", "Télé-Docs", and "Secret Story". On the left side of the page, there is a sidebar with navigation links like "Navigation", "Recherche", and "Lab", as well as a "Date de diffusion" section showing "01-02-2008 00:11:00" and "Date de la reprise: 01-03-2008 00:00:00".

A screenshot of the France 2 website's homepage. The top navigation bar includes links for 'France 2 - émissions', 'Chaine', 'Séries & Séances', 'Documentaires', 'Cinéma', and 'Plus'. A search bar and a 'mon agenda' button are also present. The main content area features several program cards: 'Mme la FRANCE', 'Les BRONOS D'OR', 'CONCERT à PARIS', 'LA GRANDE SOIREE DU 31 DÉCEMBRE', 'ZOKRO', and 'VU DE L'ANNÉE'. Below this, a section titled 'En direct sur France 2' shows a thumbnail for 'ZOKRO'. At the bottom, there is a row of images from the 'ZOKRO' series.

ina

Consultation des vidéos/podcasts



ina webmédia /

Navigation Recherche Pages Vidéos Twitter Social beta Lab Besoin d'aide ? Aide contextuelle Dernière mise à jour 24/02/2025

Recherche

Plateforme(s) Auteur(s)

Entrez votre requête... Chercher

Entre le dd-mm-yyyy et le dd-mm-yyyy trié par les plus pertinentes Options avancées

Nombre de vidéos publiées par année (2000-2025)

27 887 résultats

Information

Anonyme Europe/Paris Français

Elise Lucet fait passer des tests anti...
Touche pas à mon ... 12-02-2025 22:42

La dépêche d'Andrea du 12 février
Touche pas à mon ... 12-02-2025 22:41

C8 : Le conseil d'Etat va-t-il revenir ...
Touche pas à mon ... 12-02-2025 10:10

Conseil d'Etat : Une tribune contre ...
Touche pas à mon ... 12-02-2025 09:53

Le concert de Baba et Kelly
Touche pas à mon ... 12-02-2025 09:51

Les chroniqueurs dévoilent leur sal...
Touche pas à mon ... 11-02-2025 09:45

Danielle Moreau fond en larmes à l...
Touche pas à mon ... 11-02-2025 09:42

Quel avenir pour TPMP ? Les chron...
Touche pas à mon ... 11-02-2025 09:42

Cyril Hanouna 2027 ? Tout le mond...
Touche pas à mon ... 10-02-2025 09:36

Mathias Barthès : le nouveau roi de...
Touche pas à mon ... 10-02-2025 09:34

L'indic de la rédac' : Sasha Elbaz dé...
Touche pas à mon ... 10-02-2025 09:33

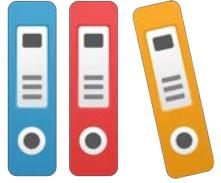
Fêtes, alcool, drogue : les folles rév...
Touche pas à mon ... 07-02-2025 11:05

Alain Delon : nouvelles révélations ...
Touche pas à mon ... 07-02-2025 11:03

Cyril Hanouna, candidat à l'électio...
Touche pas à mon ... 07-02-2025 10:54

ina

Consultation des vidéos/podcasts



ina webmédia /

Navigation Recherche Pages Vidéos Twitter Social beta Lab Besoin d'aide ? Aide contextuelle Dernière mise à jour 24/02/2025

Information Anonyme Europe/Paris Français

TikTok "M. Birch ne travaille pas pour le ch... ôter 24-02-2025 09:03

TikTok #SixNationsRugby | Élu meilleur... France tv sport 24-02-2025 09:01

TikTok Procès de Mohamed Amra : "J'esp... RTL 24-02-2025 08:48

TikTok FIGHT WEEK POUR GOMIS À l'... RMC Sport Combat 24-02-2025 08:35

france•tv Les dossiers d'ICI : La police municipale recrute France 3 24-02-2025 08:26

france•tv ICI 19/20 édition toutes régions : É... France 3 24-02-2025 08:22

Tropical Fever Radio Show & Party SOUNDCLLOUD "Tropical Fever" vol.133 "Tropical ... radiomega 24-02-2025 08:21

Dossier Edition Toutes Régions 19/20

Immersion, coulisses, décryptage : la rédaction vous emmène chaque jour au cœur de nos territoires. En semaine, un dossier de fond pour mieux comprendre les enjeux de l'actualité. Le week-end, un grand reportage pour suivre au plus près les acteurs de la société française et de ses changements.

Tropical Fever Radio Show & Party SOUNDCLLOUD "Tropical Fever" vol.132 "caribes ..." radiomega 24-02-2025 08:16

TikTok Quel rôle jouent les réseaux sociaux... france3hautsdefra... 24-02-2025 08:15

TikTok Le dernier établissement de la myt... BFMTV 24-02-2025 08:13

Europe 1 LE JOURNAL DE LA RÉDACTION Podcast Europe 1-Le Journ... 24-02-2025 08:13

Europe 1 LE JOURNAL DE LA RÉDACTION Podcast Europe 1-L'Info en... 24-02-2025 08:13

france•tv Journal 08h00 : Édition du lundi 24... France 2 24-02-2025 08:10

TikTok Louis Boyard contre Cyril Hanouna... TPMP Officiel 24-02-2025 08:09

RTL 0:55 Les infos 3:01 Les infos 3:01 Les infos 3:00 3:31 Les infos 3:00 2:06

Consultation des publications RS



ina webmédia /

Navigation

- Pages
- Vidéos
- Twitter
- Social beta

Recherche

805 832 publications (44ms)

Filtres **Térme : elonmusk**

1 / 67 153 ▶ Les plus anciennes

Texte 661 635
Photo 108 214
Vidéo 20 658
Gif 15 325

Langue

Hashtags

Rechercher par Hashtags

Elonmusk 119 558
Ukraine 63 824
Covid19 52 403
Mahsaamini 50 978
Crypto 31 409
Nft 26 134
Vr 24 000
Metaverse 20 977
Bitcoin 19 654
Cryptocurrency 18 581

Voir plus ↗

Emojis

Emoji	Nombre
🌙	17 962
😂	11 306
🚀	10 685
🔥	9 566
🔥	7 573
🙏	6 520
❤️	5 587
👤	4 751
🤔	4 003
💥	3 682
SOS	3 536
👉	3 533
👍	3 531
❗	3 517
😉	3 014
💲	2 939
💯	2 751
💰	2 666
😍	2 546
👤	2 494
✓	2 464

SpaceX @SpaceX 30/04/2009 17:02:35 (Europe/Paris)
@ElonMusk Elon Musk does not have a Twitter account. Please discontinue this account immediately. Your fraud has been reported.

Ted Patrick @_ted 14/02/2013 18:32:55 (Europe/Paris)
Dammit, now I want a car with an API, thanks @elonmusk

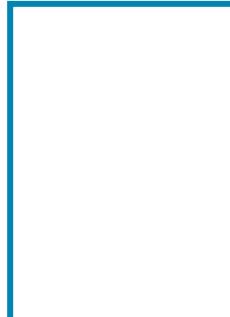
mathieu Kassovitz @kassovitz1 19/03/2013 13:47:20 (Europe/Paris)
@elonmusk Sir I have a S Signature it's time to make it fun let us tweak with the cpu driving modes should be incorporated in the system.

mathieu Kassovitz @kassovitz1 a répondu 23/03/2013 03:03:11 (Europe/Paris)
@elonmusk like... Super super dead ?

Mark Rechtin @MarkRechtin 09/04/2013 02:37:41 (Europe/Paris)
You say you want a revolution? Well, you know, Elon wants to change the world.
<http://t.co/7OjKKRaUqq> @TeslaMotors @ElonMusk

Dominique Delport @domdelport 30/05/2013 06:03:41 (Europe/Paris)
ELECTRIC : #ElonMusk: #Tesla Supercharger network to triple by the end of next month
<http://t.co/2yPYygD365>

ina

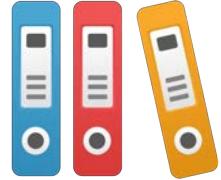


Ina / le lab

accompagnement des chercheurs



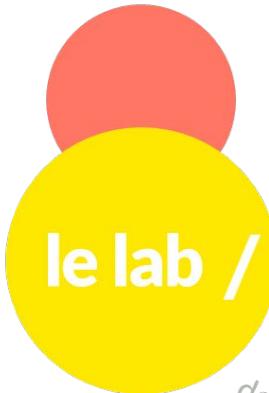
INA / Le lab



Universitaires
Chercheurs
Étudiants



Comité scientifique



Outils
& données



Direction
Data & Technologies

Expertise
documentaire



Direction
Patrimoines

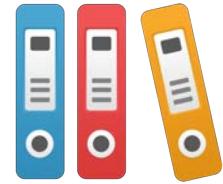




Une offre à facettes

- **Accompagnement méthodologique**
 - Permanence sur RDV
 - Sprints (1-2 mois)
 - Projets en résidence (AAP annuel)
- **Mise à disposition de données et d'outils**
 - Animation scientifique
 - Séminaire
 - Ateliers

INA / le lab : missions

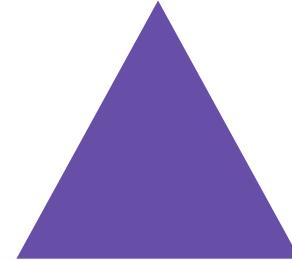


Participer à l'Analyse
des données

Déployer, former
à des Outils

Réaliser des
Traitements IA, ML

Fournir des corpus
d'Archives



OUTILS

Vision par ordinateur
Détection du genre (Inaspeech)
Transcription (Whisper)
OCR, NER...

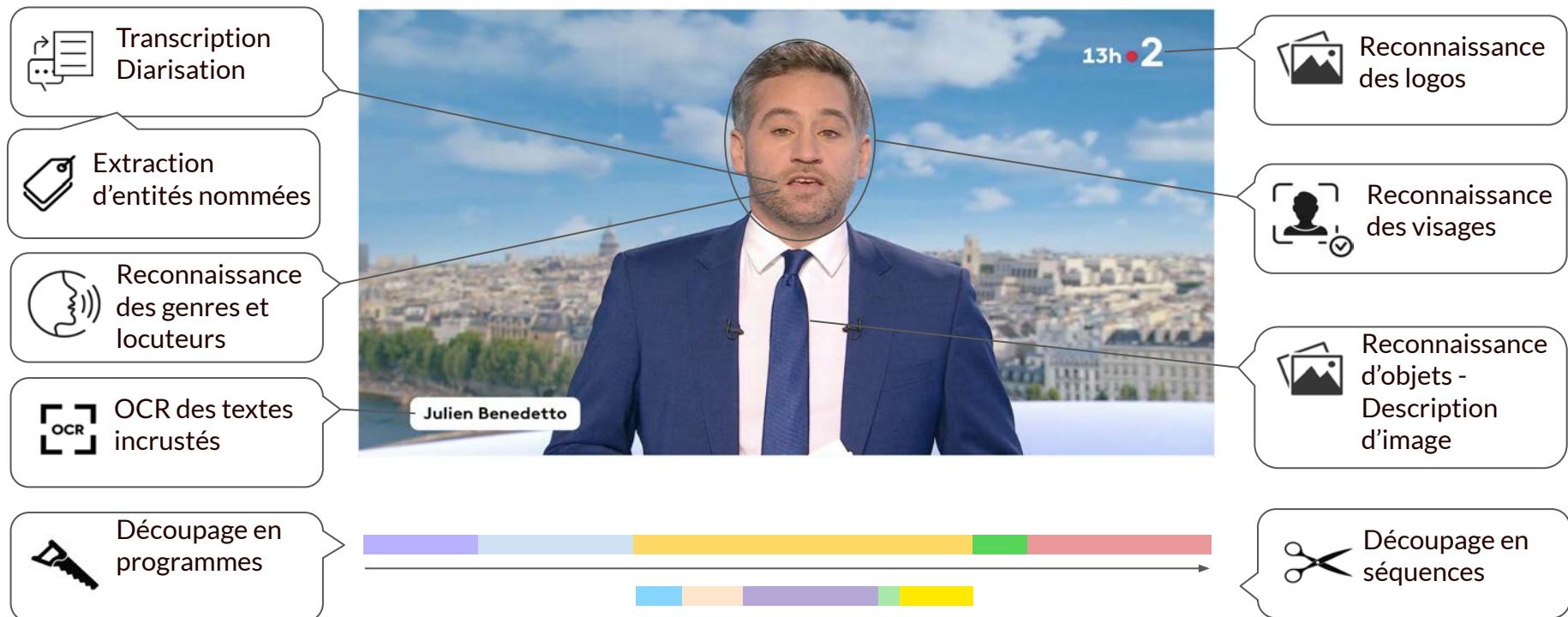
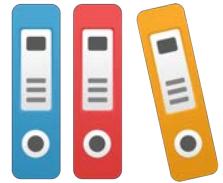
Données documentaires

Médias numérisés

DONNÉES



INA / Le lab : traitements & outils



INA / Le lab : accompagnement de projets



105 chercheurs rencontrés depuis 2022

84 projets soutenus

dont **54** l'année passée

Par type d'approche

44% de thèse

25% de recherche individuelle

10% de projet collectif

10% travaux de master

Par discipline

27% en information et communication

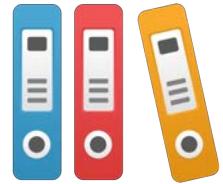
20% en sociologie

17% en économie

8% en informatique

3% en histoire

INA / Le lab : accompagnement de projets



Données utilisées (2024-2025)

Par type d'approche

85% télévision

10% web

5% radio

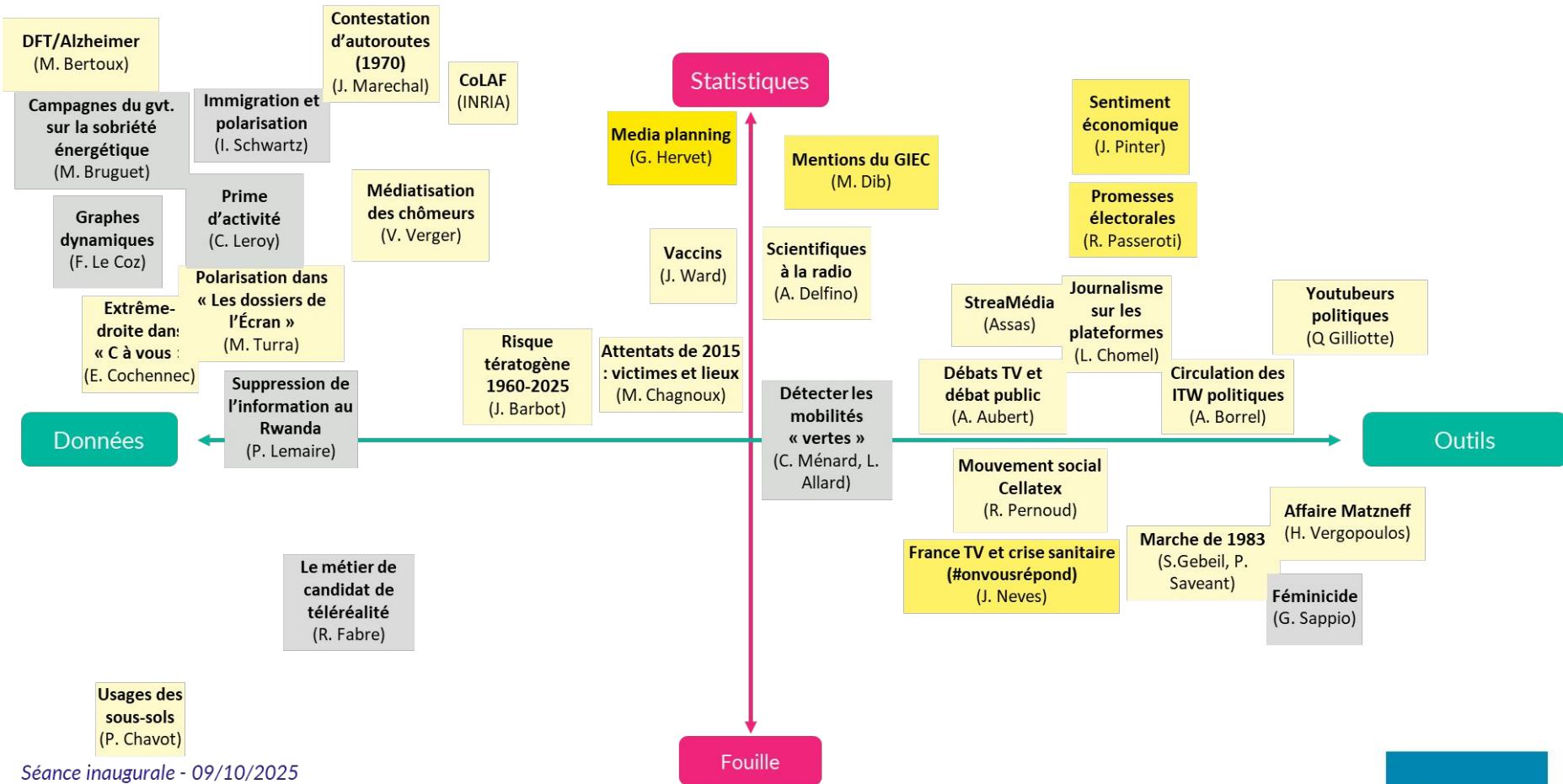
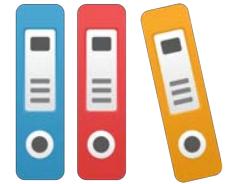
Par discipline

75% utilisent la transcription

22% utilisent uniquement les notices documentaires et métadonnées

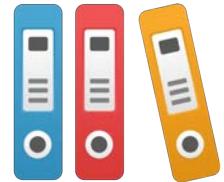
8% utilisent le média en lui même

INA / Le lab : accompagnement de projets



Séance inaugurale - 09/10/2025

INA / Le lab : accompagnement de projets



Janine BARBOT, Directrice de recherche en **sociologie** à l'INSERM

« *La trajectoire du risque tératogène dans les médias français depuis l'affaire de la thalidomide (1960-2025)* »

Marie CHAGNOUX, Maîtresse de conférences en **SIC** à l'Université Paris 8 Vincennes-Saint-Denis

« *Médiagénie des actants : appréhender en dynamique le rôle des médias dans la construction mémorielle des attentats de novembre 2015* »

Philippe CHAVOT, Maître de conférences en **SIC** à l'Université de Strasbourg

« *La construction des représentations publiques des usages des ressources du sous-sol* »

Laurent CHOMEL, Docteur et chercheur associé en **SIC** à l'Université Sorbonne-Paris-Nord

« *Reconfigurations du journalisme critique sur les plateformes : analyse comparée de 6 médias alternatifs vidéographiques français (2021-2025)* »

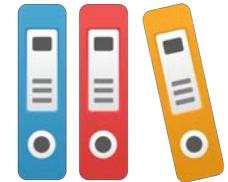
Anouck DELFINO, Doctorante en **SIC** à Sorbonne Université

« *Les scientifiques à l'antenne : étude de la visibilité et du traitement médiatique des chercheurs dans les médias sonores français (radio et podcasts natifs) entre 2018 et 2025* »

Rémi PASSEROTI, Doctorant en **Sciences politiques** à l'Université Grenoble Alpes

« *La trajectoire des promesses électorales : dynamiques médiatiques & stabilité programmatique lors de la présidentielle de 2022* »

INA / Le lab : accompagnement de projets



Séminaire 2025-2026

Replays : www.youtube.com/@inarecherche

Jeudi 9 octobre
2025

Séance inaugurale:
Critique de l'économie politique de l'espace public

Nikos Smyrnaios

14 novembre 2025 Santé

Body Capital (Solene Lellinger et Christian Bonah)
Jeremy Ward et Semiao Yang
Marius François et Mohamed Hassan

16 janvier 2026 Écologie et climat

Quota Climat (Jean Sauvignon)
Charles Menard et Laurence Allard
Samuel Gache

13 mars 2026 Emploi et précarité

Michael Zemmour
Vincent Verger
Claire Leroy

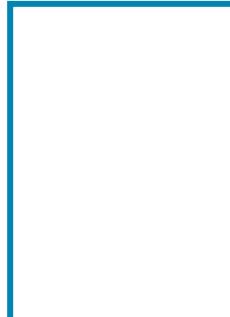
22 mai 2026 Mouvements sociaux

Paola Sedda
Raphaël Pernoud
Thibault Grison

19 juin 2026 Intervention politique et information

TBD
Julien Boyadjian
Pauline Lemaire

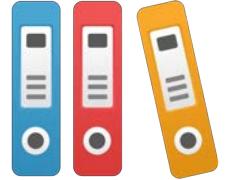




Impacts de l'IA



L'IA à l'INA



L'équipe 2IA : Intégration de l'IA

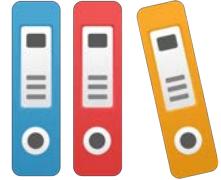
Industrialisation de traitements à des archives Radio, TV et web

Analyses image/vidéo, audio et textes

- transcription de la parole en texte
- segmentation paroles/musiques & hommes/femmes & locuteurs
- reconnaissance des locuteurs
- référentiel des visages des personnalités publiques
- analyse des génériques
- classification thématique
- recherche multi-modale : plein-texte / vectorielle



L'IA à l'INA



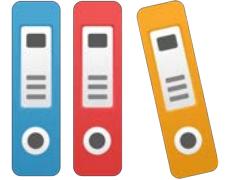
Le service de la recherche

La description des contenus audiovisuels est au cœur de la création des activités de recherche de l'INA

- Gender Equality Monitor : Rapport annuel ARCOM/INA
- Pantagruel (ANR) / ArGIMI (BPI) : création et évaluation de LLMs souverains
- STREAMédia (ANR) : reconnaissance et classification des genres éditoriaux des vidéos d'actualité



L'IA à l'INA



Applications au dépôt légal du web

Traitements par lots de médias du web en fonction :

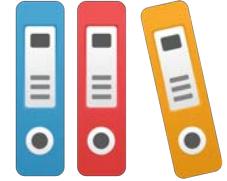
- besoins internes
- demandes de chercheurs

Principalement des traitements audio aujourd'hui

- transcription de la parole en texte
- segmentation paroles/musiques & hommes/femmes & locuteurs
- reconnaissance des locuteurs

360 000 heures en cours de traitements

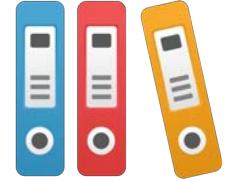
Usages des données dépôt légal & IA



Toutes demandes d'export de données pour des projets de recherche passent une analyse scientifique et juridique.

Les droits et devoirs des chercheurs d'un projet font l'objet d'une licence d'utilisation de données.

Impacts de l'IA sur la collecte du Web

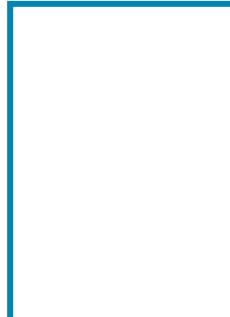


Les grands modèles de langages se sont nourris des données présentes sur le web.

Depuis quelques années nous observons une fermeture des APIs permettant l'accès aux données des plus grandes plateformes.

Ces mêmes plateformes se protègent plus efficacement du scraping, ce qui impacte aussi l'archivage du web.

C'est aussi le cas des sites produisant des contenus de valeurs comme la presse et les plateformes de TV et Radios.



Merci

