



MÉMOIRE  
AUGMENTÉE

# L'archive "augmentée" : évolution des pratiques médiatiques et nouvelles approches patrimoniales

# Sommaire

- **01.** L'INA : missions et collections
- **02.** Au-delà du « linéaire »
- **03.** L'archive web de l'INA
- **04.** Le cas de l'archive Twitter

ina

MÉMOIRE  
AUGMENTÉE

01.

→ L'INA : missions  
et collections

# Les missions de l'INA

## Conserver



**100 km linéaires** d'archives conservées sur supports originaux  
**30 Po** d'archives numérisées ou nativement numériques

## Valoriser/Transmettre



**20 millions d'heures** numérisées

## Former



**491 offres de formation continue** pour **3500 professionnels/an**  
**14 diplômes universitaires** du BTS au Master

# Les collections de l'INA

22 millions d'heures de programmes  
télé et radio



179 chaînes de télé/radio captées 24h/24 7j/7  
60 millions de documents

2 millions de photos



10 Po d'archives du Web



# Extension de la mission de dépôt légal

- *Loi du 20 juin 1992* : Dépôt Légal de la Radio et de la télévision confié à l'INA
- *Loi DADVSI du 1<sup>er</sup> août 2006* : Dépôt légal des « signes, signaux, écrits, images, sons ou message de toute nature faisant l'objet d'une communication au public par voie électronique »
- *Décret du 19 décembre 2011*



- les « services de communication au public par voie électronique »,
  - édités par les chaînes de radio et télévision
  - Portant essentiellement sur les contenus et médias « broadcast »
  - les SMADs offrant des contenus produits par les médias broadcast



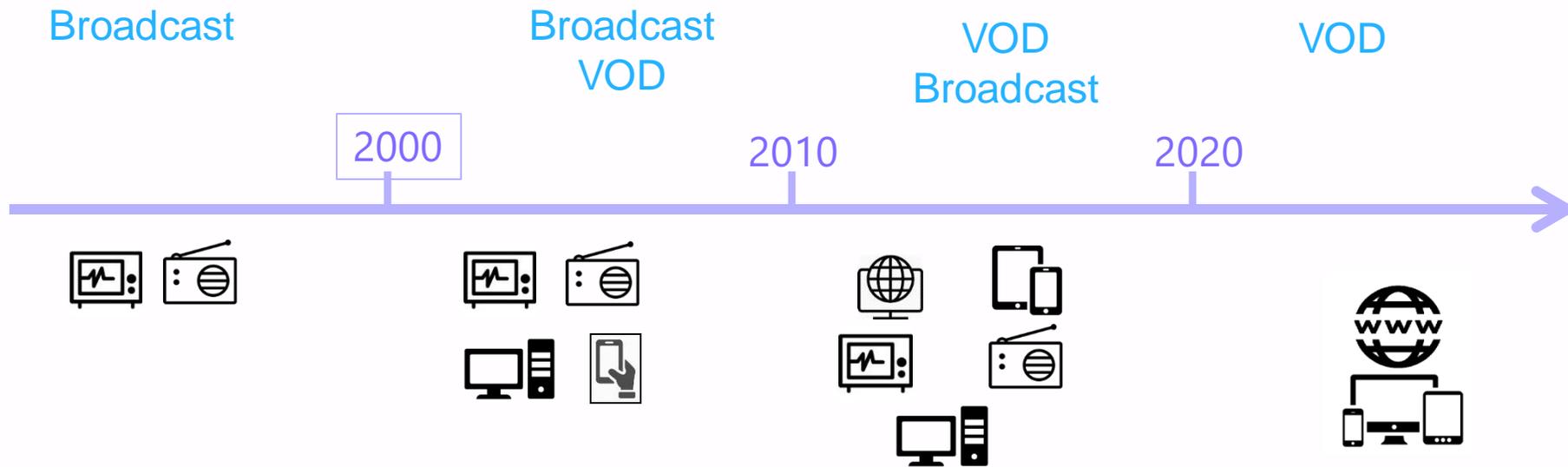
- Le domaine français à l'exception de ce qui est collecté par l'INA

**Cohérence et continuité des collections patrimoniales**

02.

→ Au-delà du « linéaire »

# Évolution des technologies de diffusion et pratiques de consommation



# Nouvelles pratiques éditoriales des médias *broadcast*

## Quels supports ?

- Présence démultipliée sur une variété de plateformes
- Création de comptes sur les réseaux sociaux (Youtube, Twitter, Facebook)
- Sites dédiés

## Quels contenus ?

- Programmes courts
- Vidéos verticales (9/16)
- Contenus web exclusifs
- Programmes VR/AR

## Quels formats ?

- Adaptation de l'éditorialisation à chaque modalité de publication
- Création de communautés
- Participation des usagers

# Nouvelles pratiques éditoriales des médias *broadcast*

## L'exemple d'Arte

**Blow Up, l'actualité du cinéma (ou presque) - ARTE**  
86 674 abonnés

ACCUEIL VIDEOS PLAYLISTS COMMUNAUTÉ CHÂÎNES À PROPOS

**Le Mariage au cinéma - Blow up - ARTE**  
56 380 vues · il y a 1 an  
C'est fou ce qu'on peut se marier au cinéma. De quoi bien évidemment faire un petit point sur la question.  
Abonnez-vous à la chaîne Youtube d'ARTE Blow Up : <https://www.youtube.com/channel/UCIE1...>  
Bonne nuit sur les réseaux sociaux - LIRE LA SUITE

LES CHÂÎNES ARTE

- BITS, magazine presqu...  
S'ABONNER
- Tu mourras moins bête...  
S'ABONNER
- Tout est vrai (ou presq...  
S'ABONNER

Les Pires looks de John Travolta - Blow Up - ARTE  
8,1 k vues · il y a 2 mois

Vous vous souvenez du festival de Cannes 1962 ? - ...  
3,7 k vues · il y a 2 mois

Les Génériques de Roger Coman - Blow Up - ARTE  
7,4 k vues · il y a 2 mois

Accueil Notifications Messages

Blow Up  
@blowup\_arte  
Le web-magazine cinéma d'ARTE.TV  
Paris  
[arte.tv/blowup](https://arte.tv/blowup)

Tweets 888 Abonnements 60 Abonnés 5 798 J'aime 15

Tweets Tweets & réponses Médias

Blow Up @blowup\_arte · 13 jul.  
#RogerComan, le Pape de la série B est dans Blow up :  
 Les Génériques de Roger Coman - Blow Up - ARTE  
Blow Up, l'actualité du cinéma (ou presque) par ARTE.

facebook

Blow Up  
@blowuparte

Accueil À propos Photos Événements Vidéos Arte Help Netquette Publications Communauté Infos et publicités

Créer une Page

Blow up #339

Photos

Rechercher publications sur la Page

Site web culture et société

Communauté

- 7 141 personnes aiment ça
- 87 041 personnes suivent ce lieu

À Propos

- Contact Blow Up on Messenger
- cinema.arte.fr/magazine/blow-up
- Site web culture et société

# Nouvelles pratiques éditoriales des médias *broadcast*

## L'exemple de France.tv

### Quels supports, Quels contenus ?

Sites dédiés, publics ciblés :

- Slash.tv : 15-25 (millennials)
- Studio 4 : 25-45 (familles)

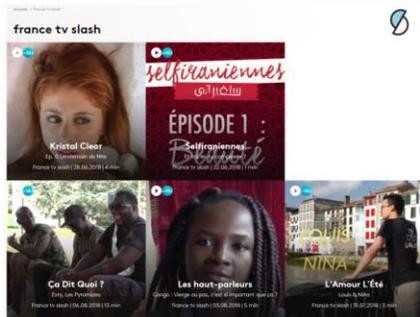
### Quels formats ?

Inspirés des usages initiés par les pure players (Netflix, Youtube)

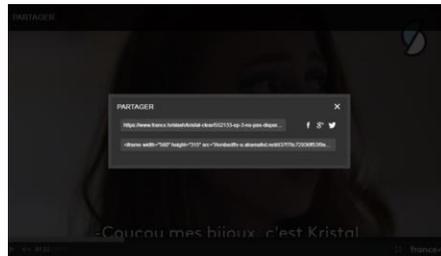
- Interfaces épurées
- Suggestions
- Favoris
- Playlists
- Evaluations, partages et commentaires

# Nouvelles pratiques éditoriales des médias *broadcast*

## L'exemple de France.tv



Interfaces épurées



Partages



Playlists



Suggestions

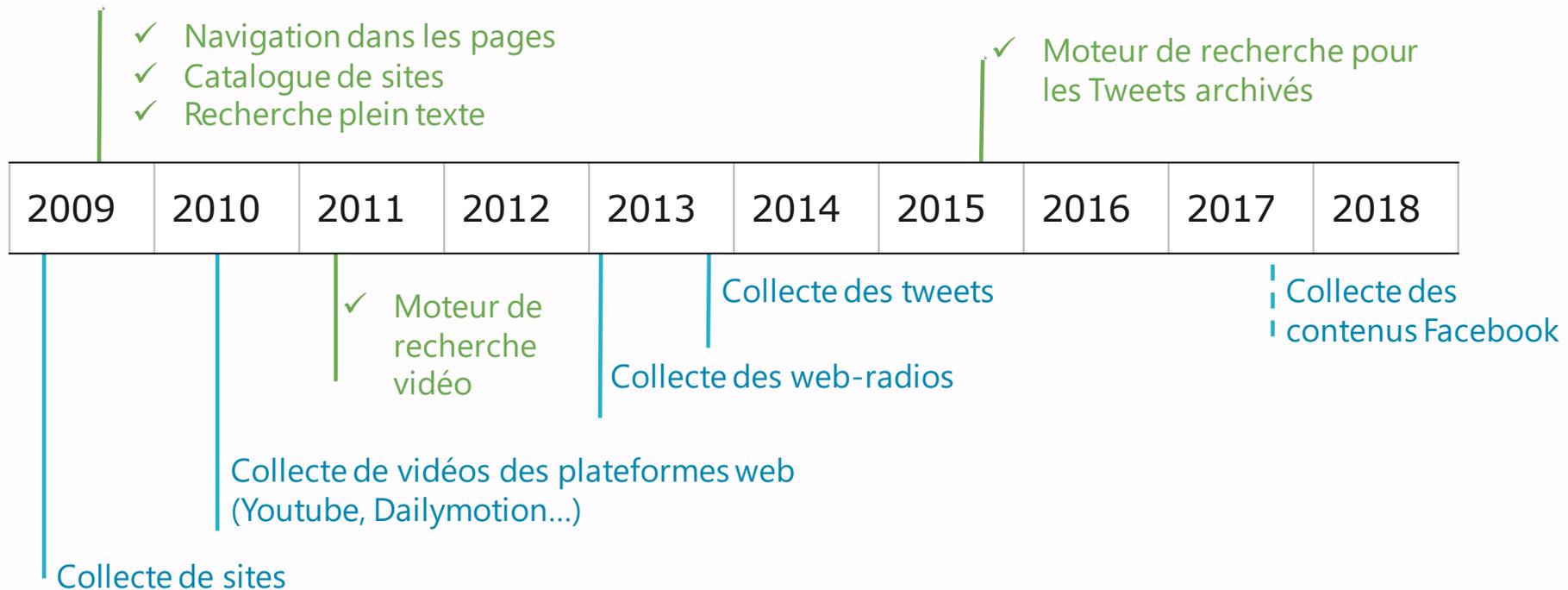
ina

MÉMOIRE  
AUGMENTÉE

03.

→ L'archive web INA

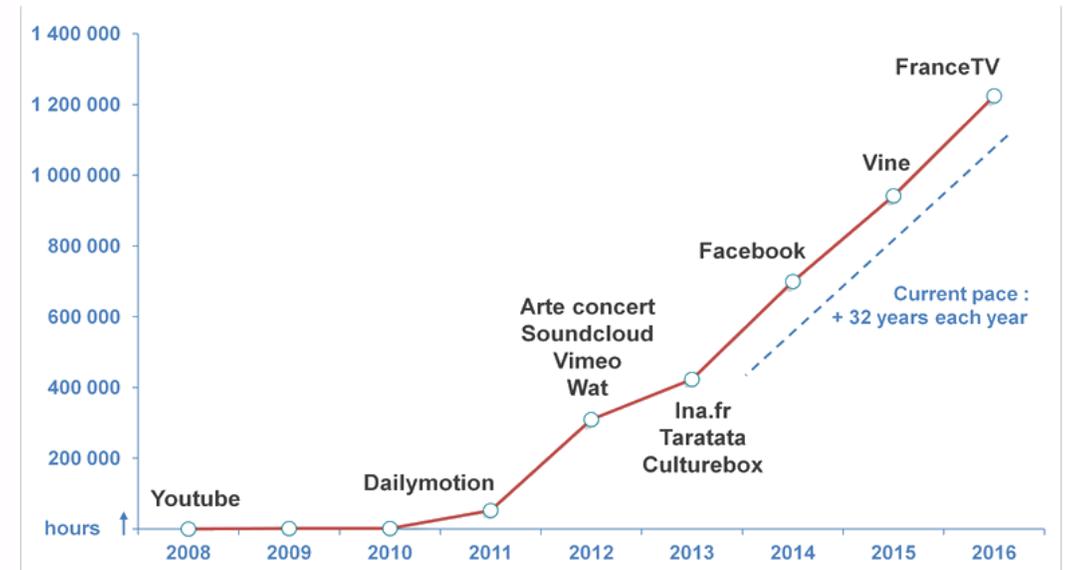
# Chronologie de l'archivage des contenus nés-numériques



# L'archive web en genres et en nombres

## Une déclinaison d'objets :

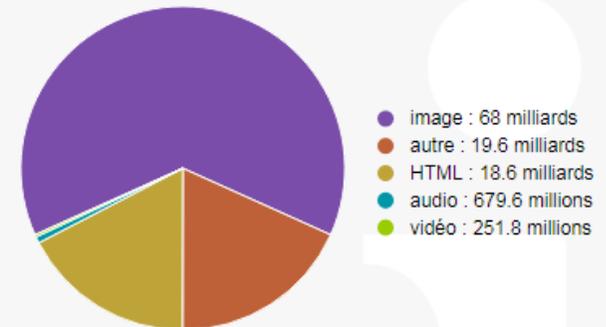
- Sites ( 15 450 )
- Réseaux sociaux :
  - comptes Twitter ( 15 750 )
  - hashtags ( 1500 )
- Plateformes vidéos : Youtube, Dailymotion, Vimeo... ( 9 400 chaînes )
- Plateformes de podcasts : ( 5 300 émissions )



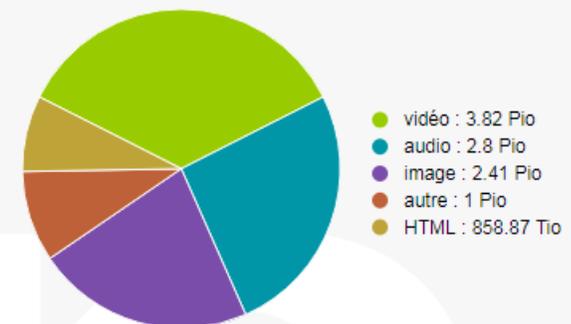
# L'archive web en genres et en nombres

- Archive: 11 Petabytes
- 30 millions de vidéos = 3 millions d'heures
- 2 milliards de tweets

Nombre de ressources archivées par type



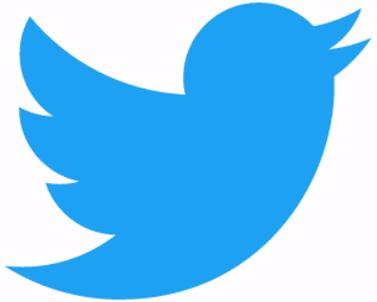
Taille totale des ressources archivées par type



04.

→ Le cas de l'archive  
Twitter

# Twitter : extension du périmètre de l'archive INA

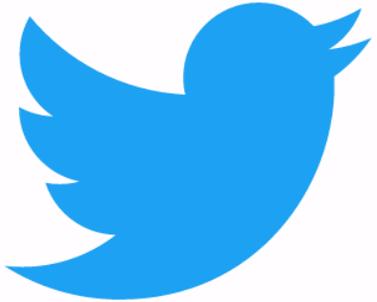


- 326 millions d'utilisateurs actifs mensuels
- 12,8 millions d'utilisateurs actifs mensuels en France
- Twitter est la plateforme des conversations autour de la télévision

## Sources :

- <https://www2.harris-interactive.com/social-life-2020>
- [https://marketing.twitter.com/content/dam/marketing-twitter/insights/tvconversation/InsightTVConversation\\_Infographic.png](https://marketing.twitter.com/content/dam/marketing-twitter/insights/tvconversation/InsightTVConversation_Infographic.png)

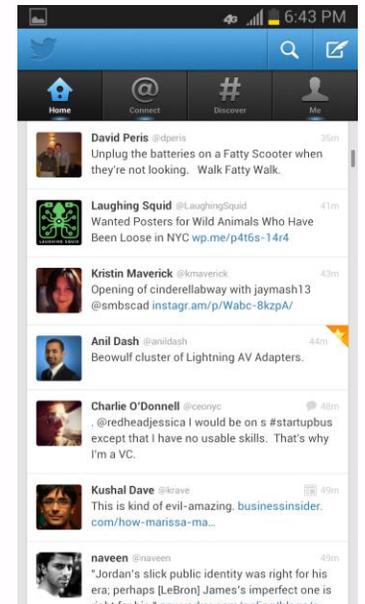
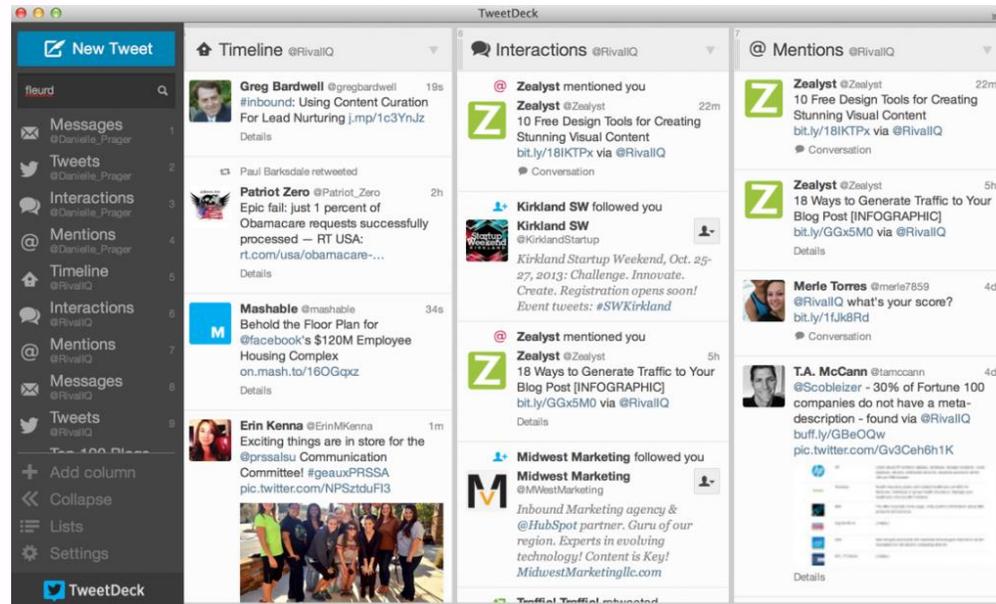
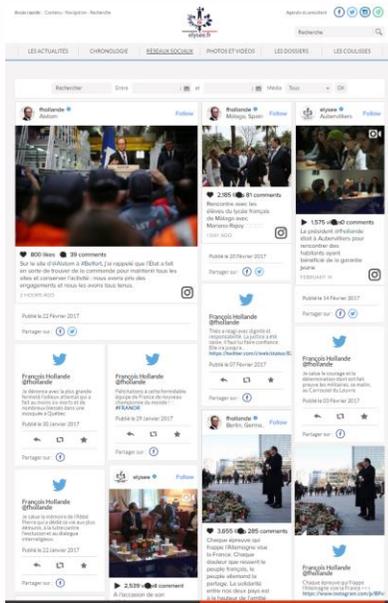
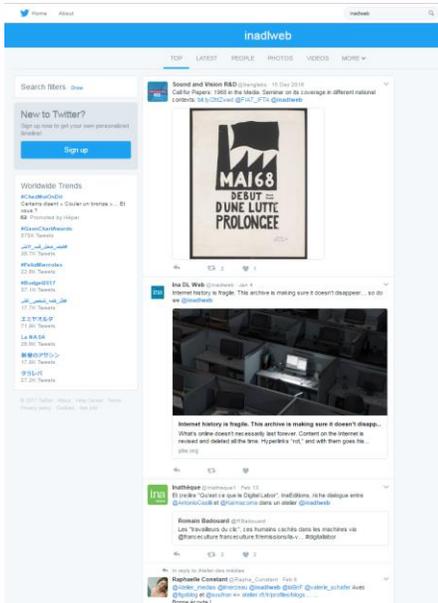
# Twitter : extension du périmètre de l'archive INA



- Renouvellement des pratiques de production et de consommation de l'actualité
- Réactions aux évènements “médiatiques”
  - évènements sportifs
  - catastrophes, attentats
  - élections présidentielles
- Live tweet événementiel
  - conférence scientifique
  - keynote technologique
- Social TV / Second Screen
- Données ‘facilement analysables’
  - intelligence économique, marketing
  - sciences sociales

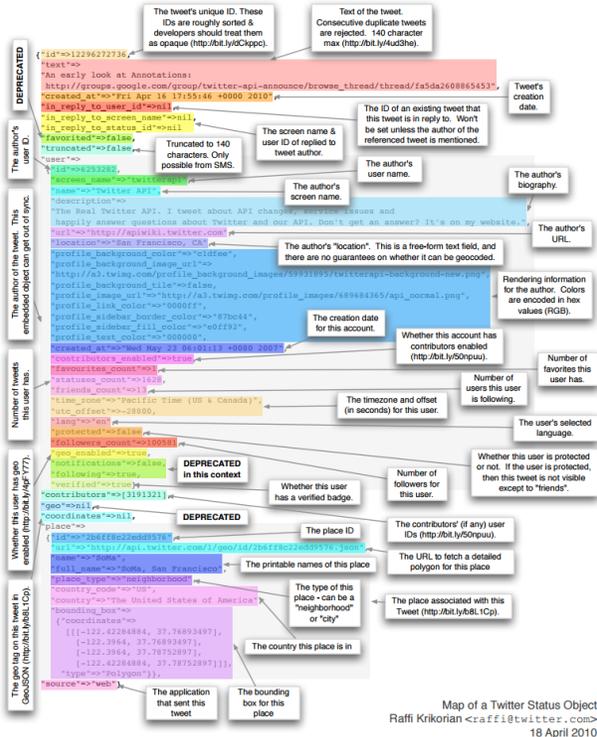
# Twitter : extension du périmètre de l'archive INA

- Depuis février 2014
- 2 milliards de tweets
- Captation par les APIs Twitter (donnée)



# Twitter : extension du périmètre de l'archive INA

- 140/280 caractères et +30 métadonnées
- Entre 3000 et 4000 caractères en tout
- Le texte d'un tweet représente environ 5% de l'information disponible.



# Twitter : Collecte par les APIs

## API : Application Programming Interface

solution informatique fournie par un site tiers pour permettre l'accès à des services ou des données

### ➤ REST API

- Les tweets historiques, max 3200
- 450 demandes toutes les 15 min

### ➤ Streaming API

- Pas des données historiques
- Le flux temps réel des messages répondant à une requête
- Seuil à 50 tweets par seconde max

# Sélection des sources

- Curation manuelle par les documentalistes de l'INA en lien avec :
  - les collections audiovisuelles
  - Les événements à forte portée médiatique
- Identification de la présence en ligne des acteurs des médias *broadcast* :
  - Chaînes de TV/Radio
  - Émissions
  - Personnalités

# Les collections de Tweets archivés

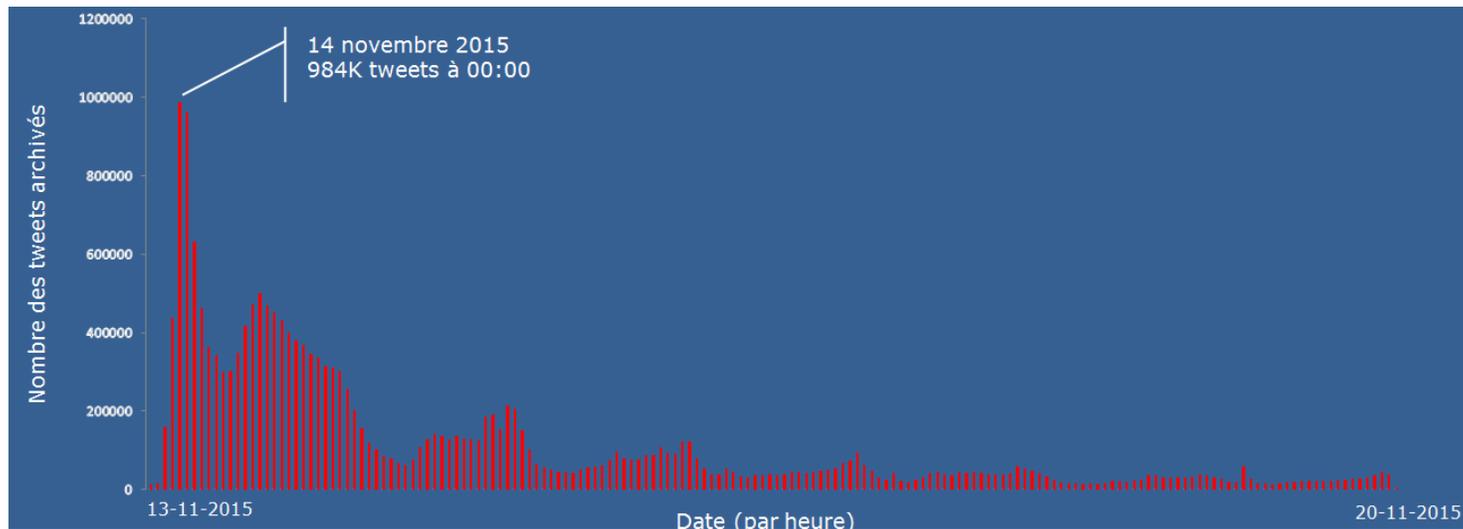
➤ **Collection large** en lien avec l'audiovisuel français

- 15 750 utilisateurs ( tf1, france2, rfi, journaliste, animateur, ...)
- 1 500 hashtags en lien avec l'audiovisuel et les médias (#tpmp, #quotidien, #Le20h, ...)
- 1 000 hashtags en lien avec l'actualité (#jesuischarlie, #metoo, #covid-19, ...)

# Les collections de Tweets archivés

## ➤ Collections événements

- Attentats
- Elections2017, Elysee
- COVID19
- ...



- ## ➤ Attentats du 13 novembre
- + 20 millions de tweets
  - À partir de 23:00
- #Bataclan  
#JeSuisParis  
#PrayForParis  
#ParisAttacks

...

# Les collections de Tweets archivés : ce qui manque

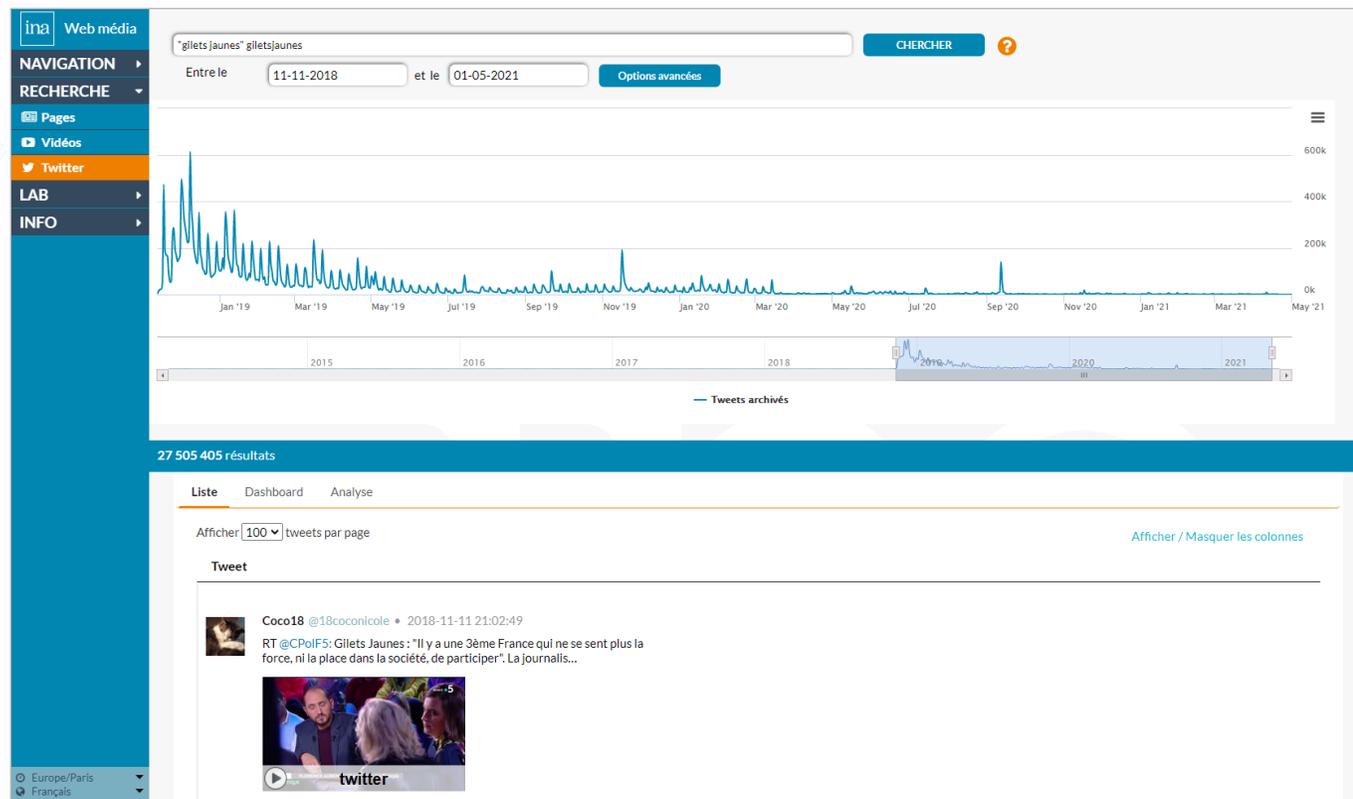
## ➤ Interactions et *followers*

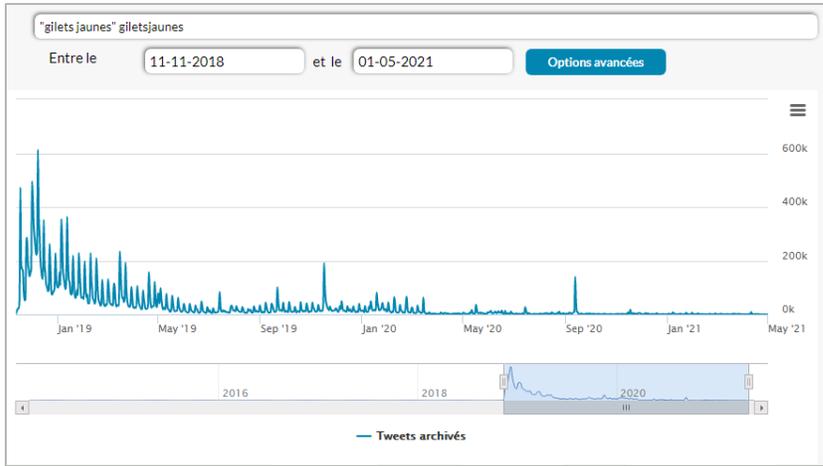
- liste des abonnés d'un compte
- L'évolution dans le temps du nombre de retweets :  1
- Les réponses à un tweet déjà archivé : En réponse à @
- Les mentions : @

# Contextualiser pour des usages scientifiques

- Représentation canonique ?
  - forme
  - données
  - Authenticité
  
- Analyse / fouille
  - Représentativité
  - Outils spécialisé ou générique
  - Ouverture des données ?
  
- Social TV / Second Screen
  - Synchro TV
  - Indexation

# Contextualiser pour des usages scientifiques : les interfaces





quelques Français vont finir par sortir en masse» https://cc0.fr/8m/ku...



Oeildelynx @Oeilde146148608 • 2019-11-11 09:01:19  
 RT @MediasUpr: @UPR\_Asselineau sur #radiocourtoisie : « M.#Macron n'est pas le Général de Gaulle, il n'a pas les mêmes fonctions. Il n'...



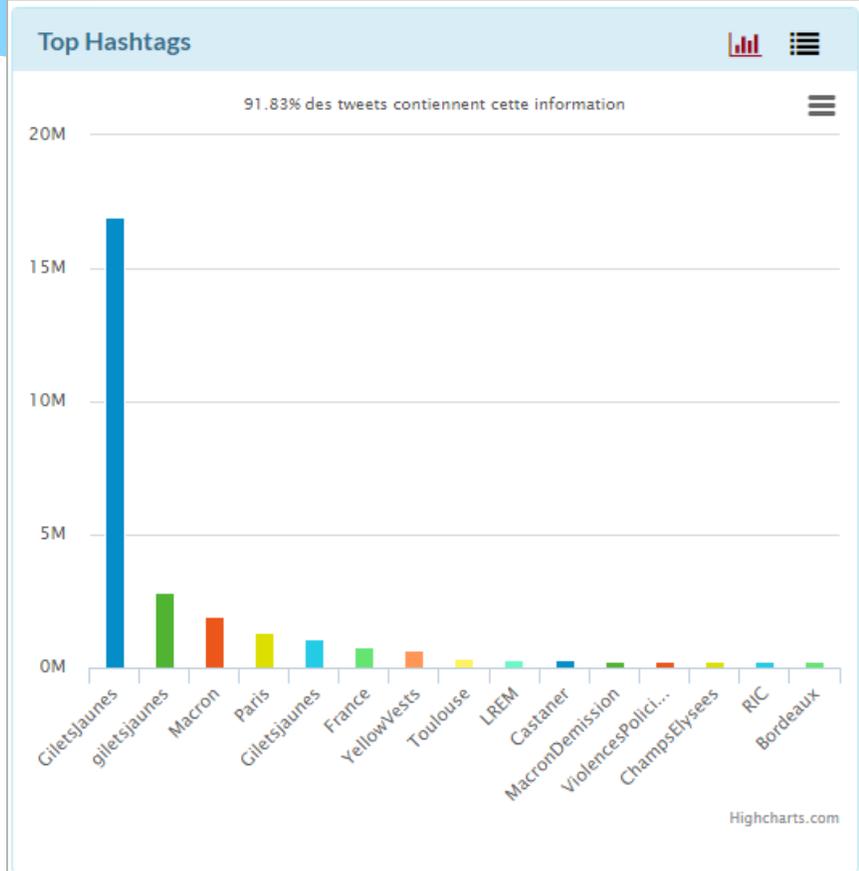
FRANÇOIS ASSELINEAU  
 EN DIRECT SUR

17 AVRIL 2019 twitter

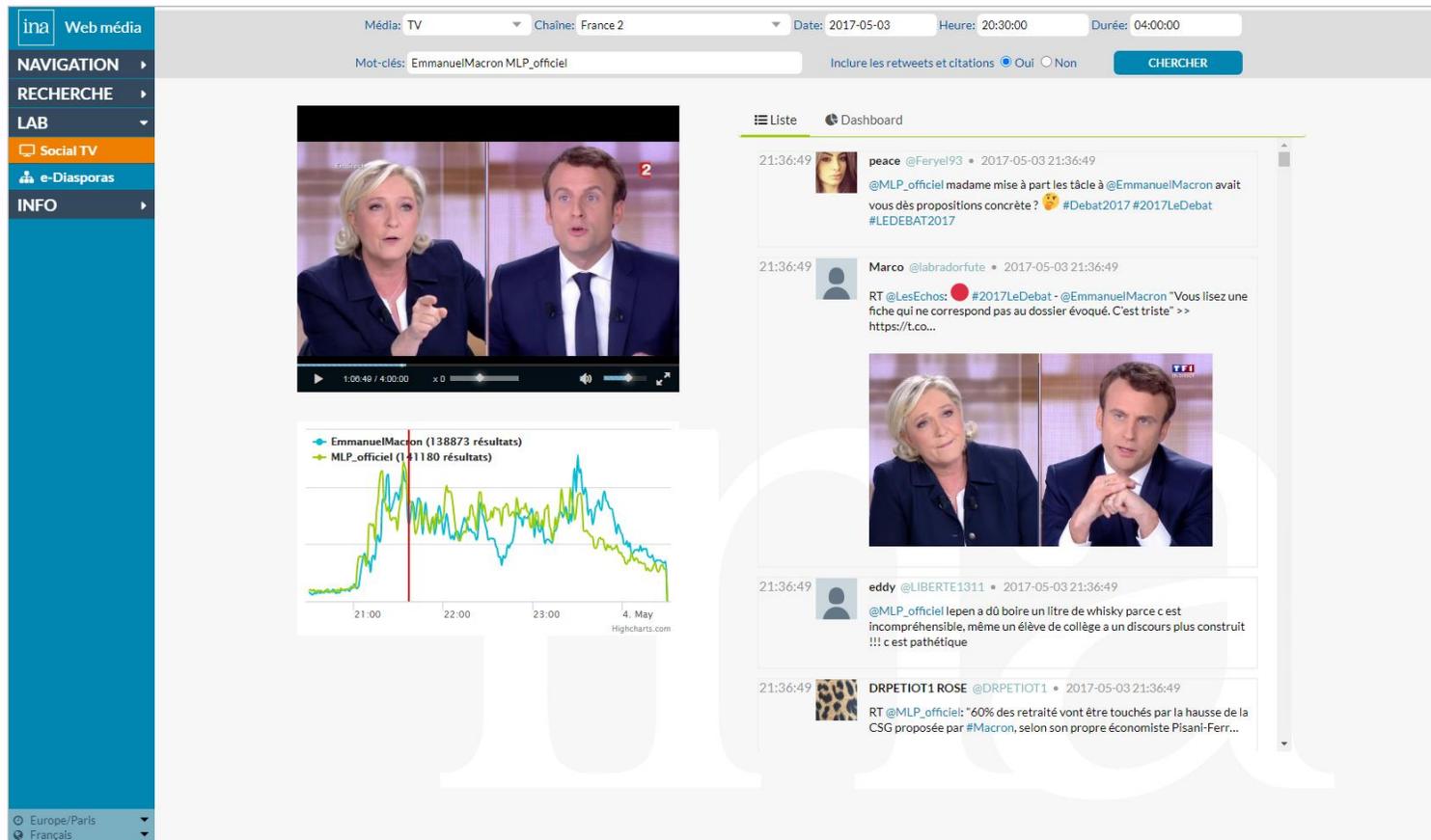
chaumulot @christianechau2 • 2019-11-11 09:12:42  
 RT @ScenicDoc: #LaPrecariteTue : Si certains avaient écouté Mélenchon en 2017,le cas d'#Etudiant du #CROUS de #Lyon,aurait pu être évité N...



edira @edira2 • 2019-11-11 09:16:49  
 RT @NPolony: Il y a un an, la colère des gilets jaunes éclatait. Mais la politique d'Emmanuel Macron, héritière de 40 ans d'échecs d'un mod...

# Second screen / Social TV



The screenshot displays the 'ina Web média' interface. At the top, there are filters for 'Média: TV', 'Chaîne: France 2', 'Date: 2017-05-03', 'Heure: 20:30:00', and 'Durée: 04:00:00'. A search bar contains the text 'EmmanuelMacron MLP\_officiel'. Below the search bar, there is a video player showing a debate between Emmanuel Macron and Marine Le Pen. To the right of the video player is a list of tweets. Below the video player, there is a line chart comparing the number of results for 'EmmanuelMacron' (138873) and 'MLP\_officiel' (141180) over time. The chart shows a significant spike in results for both terms around 21:00 on May 4th. The tweets are as follows:

- 21:36:49 peace @Feryel93 • 2017-05-03 21:36:49 @MLP\_officiel madame mise à part les tacle à @EmmanuelMacron avait vous des propositions concrète? 🤔 #Debat2017 #2017LeDebat #LEDEBAT2017
- 21:36:49 Marco @labradorfute • 2017-05-03 21:36:49 RT @LesEchos: 📌 #2017LeDebat - @EmmanuelMacron "Vous lisez une fiche qui ne correspond pas au dossier évoqué. C'est triste" >> <https://t.co...>
- 21:36:49 eddy @LIBERTE1311 • 2017-05-03 21:36:49 @MLP\_officiel lepen a dû boire un litre de whisky parce c est incompréhensible, même un élève de collège a un discours plus construit !!! c est pathétique
- 21:36:49 DRPETIOT1 ROSE @DRPETIOT1 • 2017-05-03 21:36:49 RT @MLP\_officiel: "60% des retraité vont être touchés par la hausse de la CSG proposée par #Macron, selon son propre économiste Pisaní-Ferr...

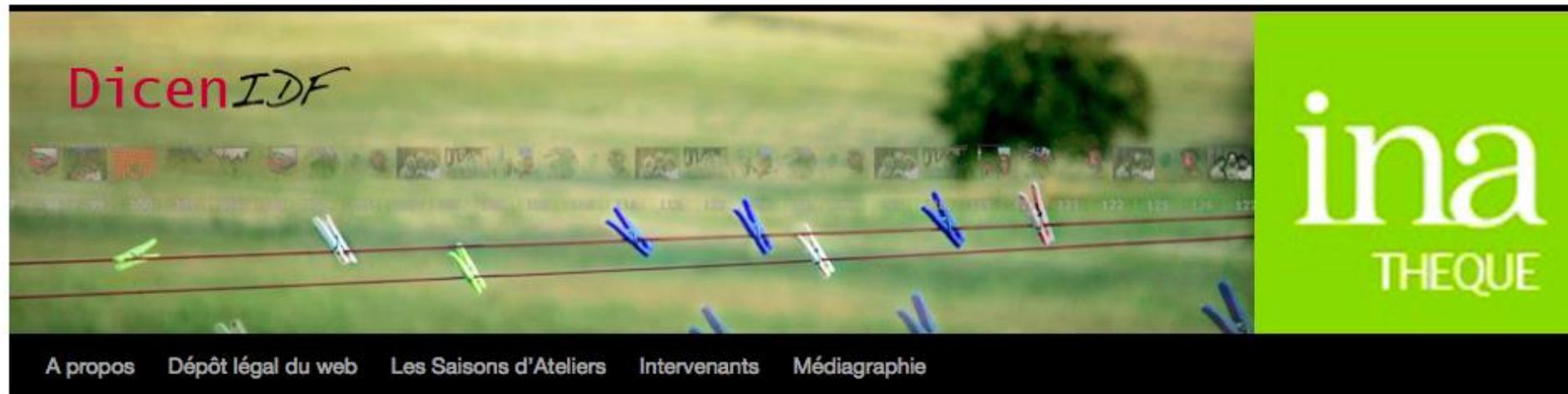


# Accompagner les usages scientifiques

2009-2014 : 6 années d'atelier de recherche co-pilotés par Louise Merzeau, acculturation et découverte

## Ateliers – Dépôt Légal du Web

*Les ateliers de recherche  
méthodologique du dépôt légal du  
web à l'Ina*



# Accompagner les usages scientifiques

2016-2017 : 2 séances de Labs, extractions, manipulations de données pour analyse



# Accompagner les usages scientifiques

Les projets :

- BodyCapital : Individus, sociétés de marchés et politiques corporelles dans une Europe audiovisuelle du 20e siècle ; Christian Bonah (Université de Strasbourg)
  - Collecte d'un corpus de tweets relatifs au procès du Mediator
  
- RwandaMAP2020 : International research network on the (archival, heritage and memorial) traces of the Genocide against the Tutsi in Rwanda ; François Robinet (Université Paris-Saclay) & Rémi Korman (EHESS)
  - Collecte d'un corpus de tweets autour des commémorations des 25 ans du génocide au Rwanda

# Accompagner les usages scientifiques

Les projets :

- La communication des chaînes de télévision TF1, France 2 et M6 sur Twitter, Facebook et Instagram ; Laurence Leveneur (Université Toulouse 1 Capitole/IUT de Rodez)
  - Mise à disposition des tweets publiés par TF1, M6 et FranceTV
  
- Ecrans en lutte. Mémoires des mouvements sociaux sur les WebTV françaises depuis la fin des années 1990 ; Sophie Gebeil (Université Aix-Marseille)
  - Mise à disposition d'un corpus de sites de WebTV et vidéos archivées à partir des plateformes de partage de vidéos et de replay

→ [jthievre@ina.fr](mailto:jthievre@ina.fr)

Merci !