

Explorations numériques des archives de l'INA et de l'INHA

Séminaire des M2 CRDM et DEFI
Organisé en collaboration avec l'EUR ArTec



Institut
national
d'histoire
de l'art



Organisation du séminaire

- **Lundi 3 mai** 9h30-17h Présentation des projets, des corpus des données INA et INHA et des outils de recherche - choix des groupes.
- **Mardi 4, mercredi 5, jeudi 6 mai** 9h30-17h : activité de groupe
- **Vendredi 7 mai** matin 9h30-12h30 présentation des projets (même salle)

Aujourd'hui

9h30-10h15 Présentation des projets

10h-10h45 Présentation des corpus INA

10h45-11h Pause

11h – 12h30 Présentation des données et des outils de recherche INHA (DEFI en salle 1 et CRDM en salle 2)

12h30-14h Choix des groupes – échange avec les encadrants

14h – 16h Présentation des outils et projets de recherche INA (amphithéâtre)

Rendu final - évaluation

Chaque groupe devra produire :

- Un **présentation** pour vendredi matin
- un **article** de blog qui devra détailler :
 - Questions de recherche
 - Méthodologie
 - Principaux résultats

Délai: lundi 10

Méthodologie du séminaire

- Questions de recherche
- Corpus de données
- Méthodes qualitatives et quantitatives
- Expérimentation
- Travailler en groupe
- Travailler dans un court délai

1) Le corps

- Corpus : données INA
- Mot-clé : corps (toute la période)
- JT / documentaires / reportage (corpus très large)
- Extraction des notices
- Analyse quantitative
- Encadrante : Camille Claverie (Antonin Segault)
- Salle : BFC204
- Mercredi à l'INA

2) Les parfums

- Corpus : données INA
- Mot-clé : parfum (toute la période)
- Corpus des publicités
- Extraction notices
- Traitement : reconnaissance des voix (homme / femme)
- Retranscriptions
- Analyse quantitative et qualitative
- Encadrante : Antonin Segault (Virginie Sonet)
- Salle : BFC204

3) Les vaccins

- Corpus : données INA
- Mot-clé : vaccin*
- JT / documentaires / reportage
- Extraction des notices
- Analyse quantitative et qualitative
- Encadrant : Fabrizio Defilippi (Marta Severo)
- Salle : BFC205
- **jeudi à l'INA**

4) Campagnes de santé publique

- Corpus : données INA
- Mot-clé : campagne de santé (drogue, alcoolisme, etc.)
- JT / documentaires / reportage
- Extraction des notices
- Analyse quantitative et qualitative
- Encadrant : Virginie Sonet (Fabrizio Defilippi)
- Salle : BFC205
- **mardi à l'INA**



5) Protections périodiques

- Corpus : données INA
- Mot-clé : nom de marques
- Corpus de publicités + JT / documentaires / reportage
- Extraction des notices
- Analyse qualitative
- Encadrant : Marta Severo (Romain Bigay)
- Salle : BFC223



6) Les animaux dans le cirque

- Corpus : données INA
- Mot-clé : cirque
- JT / retransmission
- Extraction des notices
- Reconnaissance d'image (Snoop)
- Analyse qualitative et quantitative
- Encadrant : Romain Bigay (Marta Severo)
- Salle : BFC223
- **Mercredi à l'INA**



7) Provenance

- L'objectif de ce projet est de s'intéresser à la mobilité des oeuvres dans leurs parcours de vie et donc de s'intéresser à leurs provenances
- visualiser des transferts d'oeuvres (de leurs lieux de production à leurs lieux de conservation actuel, etc.)
- outils envisagés : excel, openrefine, flourish
- Salle BFC202
- https://antoinecourtin.github.io/seminaire_M2_InfoCom_ParisNanterre_2021/projets.html

8) Matérialité

- L'objectif de ce projet est de s'intéresser à la matérialité des objets beaux-arts et notamment la question des dimensions et de la représentation visuel de cette information dans les interfaces web.
- afficher et comparer la taille d'objets culturels
- outils envisagés : excel, openrefine, rawgraph, R
- Salle BFC203
- https://antoinecourtin.github.io/seminaire_M2_InfoCom_ParisNanterre_2021/projets.html

9) Iconographie et zones annotées

- L'objectif de ce projet est de s'intéresser à la possibilité d'utiliser le corpus visuel de wikimedia commons pour réaliser des annotations "zonées" via des données structurées de wikidata.
- A partir d'un sujet de votre choix, annoter un corpus visuel, savoir requêter et visualiser ses zones d'annotation outils envisagés : wikidata, wikimedia commons, cropper IIIF
- Salle BFC202
- https://antoinecourtin.github.io/seminaire_M2_InfoCom_ParisNanterre_2021/projets.html

Choix du groupe

- Chaque groupe entre 3-5 personnes
- <https://mensuel.framapad.org/p/mip2020-9n6m>