

Informatique et MOOC | 24 juin 2014 | Paris

Enrichissement de vidéos pédagogiques et apprentissage actif : le projet **COCO**

*Colin de la Higuera, Yannick Prié, Camila Morais Canellas, Matthieu Riou
Université de Nantes*

Le projet

1

L'enrichissement des vidéos

2

L'alignement multimodal

3

Les traces

4

La plateforme

5



L'ÉQUIPE



Yannick
Prié



Olivier
Aubert



Colin de la
Higuera



Camila
Canellas



LE PROJET

coco

comin-ocw.org/

Open Education
Enrichissement/annotation vidéo
Recherche



LES OBJECTIFS



Creating open
multimodal
contents for
knowledge
diffusion

Exploring new
techniques for e-
learning

Leveraging
annotations in
video-centered
pedagogical
resources

Using machine
learning to
support OER
creation



POURQUOI LA VIDÉO ?



Souvent plus riche que d'autres médias : audio, image, mouvement, langage non verbal ...



Mais exploité principalement comme un flux d'information à sens unique. L'interaction est limitée.



VIDÉOS ENRICHIES ?

Des vidéos complétées par des contenus supplémentaires, comme des légendes, des images, des hyperliens, etc.



ANNOTATION VIDÉO ?

Une annotation vidéo est composée de données explicitement associés à un fragment spatio-temporel d'une vidéo.

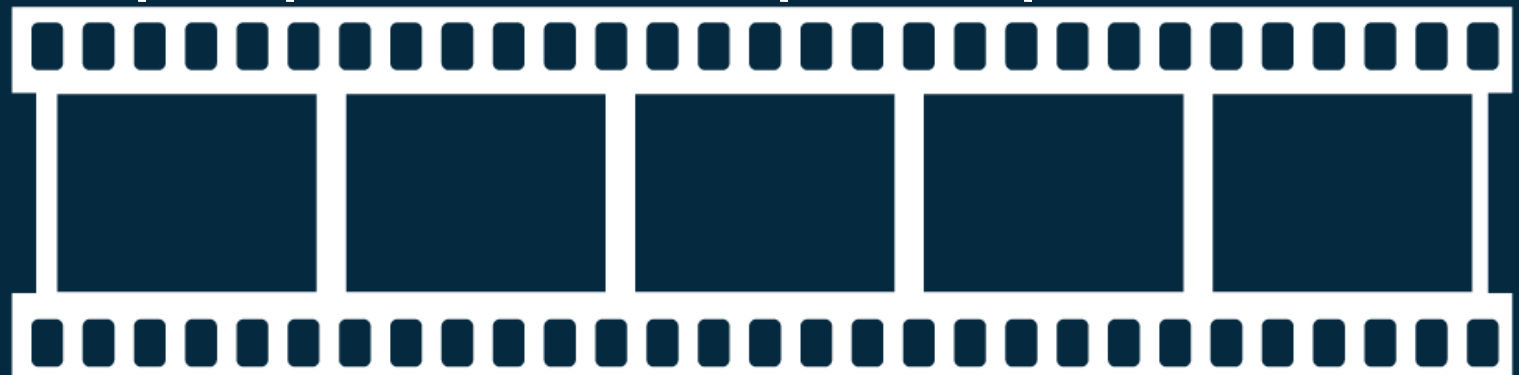


0'11

0'32

2'11

2'57



1'31

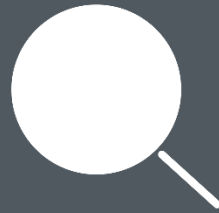
1'43

3'55





**DANS QUELS
DOMAINES ?**



Archives vidéo



Musique



Recherche



Sport



Spain vs Holland World Cup - LongoMatch

File Tools View Help

- ▶ Goal (3)
- ▼ Penalty Corner (4)
 - Name: Penalty Corner 001
Start: 0:10,000
Stop: 0:26,000
 - Name: Penalty Corner 002
Start: 0:24,513
Stop: 0:40,525
 - Name: Penalty Corner 003
Start: 0:25,740
Stop: 0:55,740
 - Name: Penalty Corner 006
Start: 8:42,086
Stop: 9:02,086
- Stroke (0)
- ▼ Green Card (4)
 - Name: Green Card 016
Start: 1:18,080
Stop: 1:38,080
 - Name: Green Card 009
Start: 2:20,549
Stop: 2:40,549
 - Name: Green Card 011
Start: 5:22,015
Stop: 5:42,015
 - Name: Green Card 018

Plays Spain Holland Tags

0:06/0:20 1X

Event	Start Time	End Time
Goal	0:10	0:26
Penalty Cor	0:24	0:40
Stroke	-	-
Green Card	1:18	1:38
Yellow Card	2:20	2:40
Green Card	5:22	5:42
Green Card	8:42	9:02
Yellow Card	-	-



**EN
PÉDAGOGIE ?**



Pédagogie

- Usages possibles
- Usage actuel



LES POSSIBILITÉS

1. Lecture active ➔
2. Annotation live
3. Annotation de performance
4. Devoir

VideoNot.es We're open-source, support us by [contributing on Github](#) or [buying us coffee](#)

Video (press Ctrl and space to play/pause) [Change video](#)

Change video

0:00 Beginning
0:04 Logos
0:10 Introduction
0:21 Type
5:07 This role was essential for the changes in legislation. This scene shows why.
5:12 The absence of the political context tells more about the director than we could imagine.
37:23 This journal was actually true!|
37:34 Should we reproduce this?
51:51 Note: do an analysis of this text.
51:42 The relationship between the two has a gap here.

Share

Feedback and Support

videonot.es/



LES POSSIBILITÉS

1. Lecture active ➤

2. Annotation live

3. Annotation de performance

4. Devoir

- ✓ reconnaître les **idées principales** d'un discours
- ✓ **identifier** les segments les plus importants
- ✓ **réduire** l'information en concepts clés
- ✓ **se souvenir** à plus long terme
- ✓ reproduire des idées en utilisant ses **propres mots**
- ✓ collaborer à la **création** et à l'**évaluation**



LES POSSIBILITÉS

1. Lecture active

2. Annotation live 

3. Annotation de
performance

4. Devoir

POLEMIC TWEET

Home Program About

日本語 Français English

Intégrer

Ecologie de l'attention #7

Le marché de l'attention et l'avènement de la publicité numérique. (09 mars 2014)

- **Philippe Tassi** : Docteur en statistique économique diplômé de l'ENSAE, Philippe Tassi fut d'abord chef des statistiques structurelles des entreprises de l'INSEE, avant d'être directeur des études de l'ENSAE de 1981 à 1989. Directeur scientifique et technique de MEDIAMETRIE depuis 1990, il est aujourd'hui directeur général adjoint de cet institut de sondages de radio et de télévision.
- **Philippe Lescaut** : Directeur délégué de l'Institut de Recherches et d'Études Publicitaires.

Nous assistons aujourd'hui à une migration de la croissance des marchés publicitaires des médias traditionnels vers les médias numériques. Les deux principales poches de croissance du marché de la publicité résident actuellement sur deux supports : la vidéo sur internet et la publicité sur mobile. La consommation de médias sur les terminaux mobiles (smartphones et tablettes) est déjà bien installée mais encore mal monétisée. A titre d'illustration, une comparaison simple entre le temps passé par les utilisateurs sur mobile et le marché publicitaire correspondant laisse espérer uniquement aux USA un potentiel de plus 20 milliards de dollars par an. Le marché de la publicité mobile devrait, dans tous les cas, à horizon 2016 dépasser les annonces sur Internet. Il reste encore et surtout à évaluer quel sera l'impact de l'entrée imminente de nouveaux acteurs de taille comme Facebook ou Twitter sur ces marchés, et sans oublier l'attention des industries publicitaires au futur coup porté sur

Liste des Annotations

Mots-clés

<http://polemictweet.com/>



LES POSSIBILITÉS

1. Lecture active

2. Annotation live ➤


3. Annotation de performance

4. Devoir

- ✓ poser une **question** au moment où elle vient à l'esprit
- ✓ ajouter des **informations pertinentes** à une partie d'un discours dès que vous les reliez
- ✓ marquer des points à **réviser**
- ✓ ajouter des **commentaires précis** du public



LES POSSIBILITÉS

1. Lecture active
2. Annotation live
3. Annotation de performance 
4. Devoir

The screenshot displays the VISU software interface. At the top, there are navigation tabs: "VISU", "Accueil", "Utilisateurs", "Séances", "Salon synchrone", and "Bilans". Below these, a title bar reads "Les loisirs", enregistrée le 20-10-2011 à 15:14. The main area features a video player showing a woman wearing a headset, with the name "V. Caroline" in the top right corner. A play button and a timestamp "00:05:17" are visible below the video. To the right, a detailed timeline shows the video's duration from 00:00:00 to 00:08:29, with a red circle highlighting the current time "00:05:17". Below the timeline, a pink bar contains the text "vous voyer / voier" with a speech bubble icon. The bottom section is a yellow workspace with various icons for annotations, including a key, a speech bubble, and a play button. On the left side, a sidebar contains a "Comments" section with a green plus icon, and a "Markers" section for "S. Serguei" with a dropdown arrow. Below this, there are sections for "Messages", "Documents", "Images", and "Videos".

<http://visu.advene.org/>



LES POSSIBILITÉS

1. Lecture active

2. Annotation live

3. Annotation de
performance



4. Devoir

✓ annoter une performance pour

l'auto-réflexion

OU

partager une analyse



LES POSSIBILITÉS

1. Lecture active
2. Annotation live
3. Annotation de performance
4. Devoir



The screenshot displays the Mediathread interface. At the top, the 'mediathread' logo is on the left, and user information 'User Name_3' and navigation links 'Tools', 'Reports', 'Help', 'Contact Us' are on the right. The main content area is divided into two panels. The left panel, titled 'Assignment', shows an assignment titled '01 Bullet Memo-Birth of a Nation' by 'User Name_392, User Name_619'. It contains instructions for a focused paragraph and a list of steps to fulfill the assignment. The right panel, titled 'Discussion', shows a post by 'User Name_2972' titled 'Silent Stereotypes and Barthesian Myths'. It includes a video player with a play button and a video titled 'cameron meeting slaves' from 'Birth of a Nation'.

Mediathread (Columbia University)



LES POSSIBILITÉS

1. Lecture active
2. Annotation live
3. Annotation de performance
4. Devoir



- ✓ analyser certains aspects de la vidéo
- ✓ les annotations peuvent servir de base pour produire un document

résumé
OU
collage vidéo



ÉTUDE

Mais comment est-ce utilisé dans l'apprentissage en ligne ?

Plateformes MOOC

- EdX
- Coursera
- Canvas Network
- Iversity
- Open2Study

Outils dédiés

- VideoANT
- VideoNot.es
- Annotated HTML
- Mediathread
- YouTube
- Matterhorn



ÉTUDE

Les fonctionnalités ?

Visualisation

- Sous-titres
- Transcription
- Enrichissements synchronisés
- Liste des annotations
- *Timeline*
- Enrichissements interactifs

Éditer/partager

- Commentaire
- Marqueur
- Outil d'annotation interne/externe
- Exportation
- Partage d'annotation interne/externe

ÉTUDE



		EdX	Coursera	Canvas LMS
Visualization	Subtitles		X ^{1,2}	X
	Transcription	X	X	X
	Other synchronized adjacent enrichments			
	List of annotations			
	Timeline with annotations			
	Interactive enrichments		X ⁴	
	Comment	X		
	Video markers	X ⁶		
External annotation				
Internal annotation	X			



ÉTUDE

Dans les MOOCs

Fonctionnalités facilitant la compréhension du discours

*Largement
présentes*



- possibilité d'ajuster la vitesse de la vidéo
- sous-titres
- transcriptions



ÉTUDE

Dans les MOOCs

Autres formes d'interaction

- quizz 
- commentaires/forum 



ÉTUDE

Outils dédiés

Plus de fonctionnalités pour le processus d'annotation



- *timelines* interactives
- exportation des données des annotations
- partage des annotations



ÉTUDE

Les outils sont disponibles.

L'usage dans l'apprentissage en ligne est encore basique.

Plusieurs scénarios pédagogiques intéressants se basent sur les annotations vidéo.



L'ALIGNEMENT MULTIMODAL

Stage de Matthieu Riou



Faire correspondre des segments de deux documents



L'ALIGNEMENT MULTIMODAL

Trois étapes

1) Découper la vidéo et le texte en segments cohérents

- Vidéo : temps des slides
- Texte : paragraphes



L'ALIGNEMENT MULTIMODAL

Trois étapes

2) Comparer les segments de la vidéo et du texte

- Tokenisation : découpage en mots
- Suppression des mots trop courants, les *stop words*
- Lemmatisation : chaque mot est remplacé par sa racine



L'ALIGNEMENT MULTIMODAL

Trois étapes

3) Trouver une correspondance globale entre la vidéo et le texte

a) TF-IDF (*term frequency–inverse document frequency*)
- Donner une valeur au mot, selon sa fréquence d'apparition dans le segment étudié ainsi que selon sa fréquence d'apparition dans l'ensemble des segments.

b) Mesure Cosinus

- Vecteurs de mots à n dimensions (n mots)
- Regarder l'angle entre deux vecteurs (plus l'angle est petit, plus les vecteurs sont proches)
- Utilisation du cosinus de l'angle (valeur croissante ; entre 0 et 1)



TEST DE SYNCHRONISATION

Accueil

Editer



standable, (2) semantically relevant, (3) covering the whole topic well, and (4) discriminative across topics. Without prior domain knowledge, extracting understandable candidate labels is non-trivial. Since a topic model and a label have different representations, it is also difficult to compare their semantics. As a result, there is no existing method to measure the semantic relevance between a topic model and a label. Even with a good measure for semantic relevance, it is still unclear how we can ensure that the label would fully cover the meaning of a topic and also capture the difference between different topic models.

In the next section, we propose a probabilistic approach to generating labels for topic models automatically.

3. PROBABILISTIC TOPIC LABELING

To generate labels that are *understandable*, semantically *relevant*, *discriminative* across topics, and of *high coverage* of each topic, we first extract a set of understandable candidate labels in a preprocessing step, then design a relevance scoring function to measure the semantic similarity between a label and a topic, and finally propose label selection methods to address the inter-topic discrimination and intra-topic coverage problems.

3.1 Candidate Label Generation

As discussed in Section 4, compared with single terms and sentences, phrases appear to be more appropriate for labeling a topic. Therefore, given a reference collection \mathcal{C} , the first task is to generate meaningful phrases as candidate labels. Phrase generation has been addressed in existing work [7, 26, 15, 6]. In general, there are two basic approaches:

Chunking/Shadow Parsing: Chunking (Shadow Parsing) is a common technique in natural language processing, which aims at identifying short phrases, or “chunks” in text. A chunker often operates on text with part of speech tags, and uses the tags to make decisions of chunking according to some grammar, or through learning from labeled training sets. In our work, we extract the chunks/phrases frequently appearing in the collection as candidate labels.

The advantage of using an NLP chunker is that the phrases generated are grammatical and meaningful. However, the accuracy of chunking usually depends heavily on the domain of the text collection. For example, if the model is trained with news articles, it may not be effective on sci-

The advantage of such an ngram testing approach is that it does not require training data, and is applicable to text collection of any ad hoc domains/topics. The disadvantage is that the top ranked ngrams sometimes are not linguistically meaningful, and it usually only works well for bigrams.

In our experiments, we compare both approaches to extract the set of candidate labels.

3.2 Semantic Relevance Scoring

We propose two relevance scoring functions to rank labels by their semantic similarity to a topic model.

3.2.1 The Zero-Order Relevance

The semantics of a latent topic θ is fully captured by the corresponding multinomial distribution. Intuitively any reasonable measure of the semantic relevance of a label to a topic should compare the label with this distribution in some way.

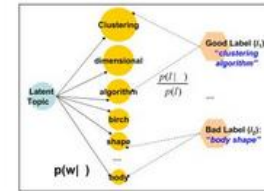


Figure 1: Illustration of zero-order relevance

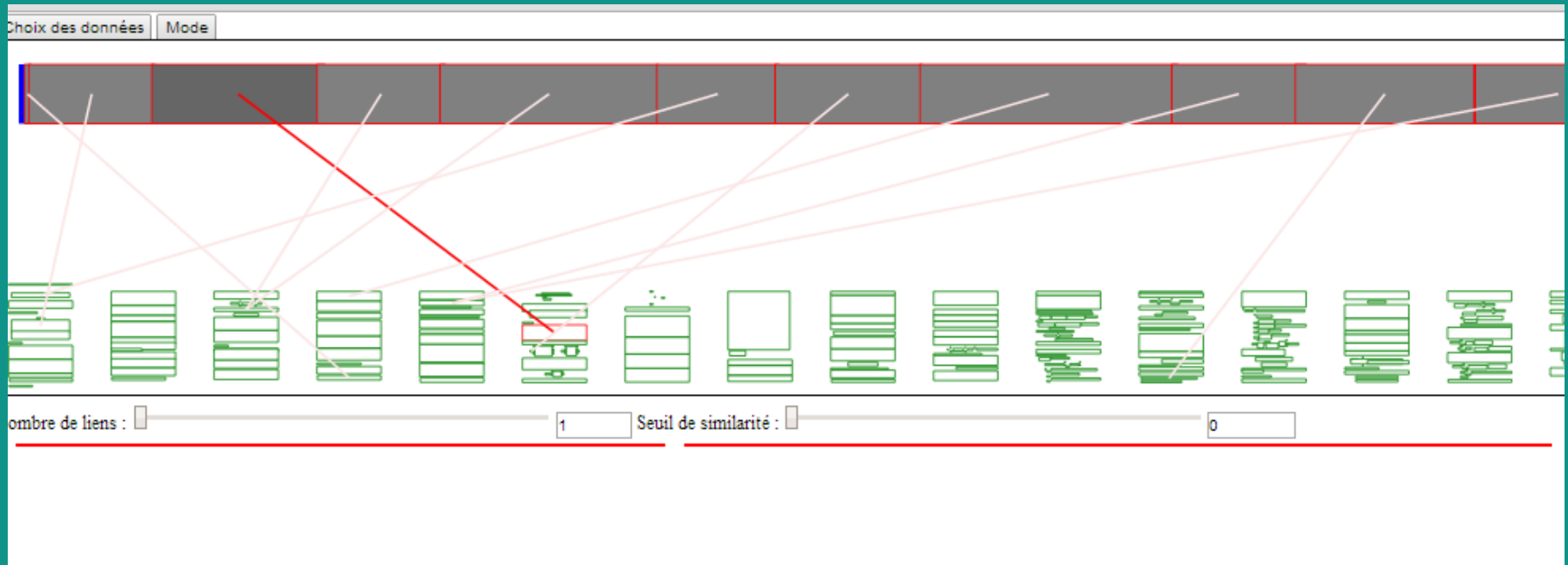
A larger circle means a higher probability.

One possibility is to define the semantic relevance score of a candidate phrase $l = u_0 u_1 \dots u_m$ (u_i is a word) as

$$Score = \log \frac{p(l|\theta)}{p(l)} = \sum_{0 \leq i \leq m} \log \frac{p(u_i|\theta)}{p(u_i)}$$

where the independence of u_i 's is assumed. The basic idea of this approach is illustrated in Figure 1. Basically, a phrase containing more “important” (high $p(u_i|\theta)$) words in the topic distribution is assumed to be a good label. $p(u_i)$ is

Video player controls: play button, progress bar, volume icon, full screen icon, time 0:06:21/0:17:54, and a checked box for 'activer synchro'.



<http://alignement.comin-ocw.org/demo/visualisation/visualisation.php?document=Regrets>

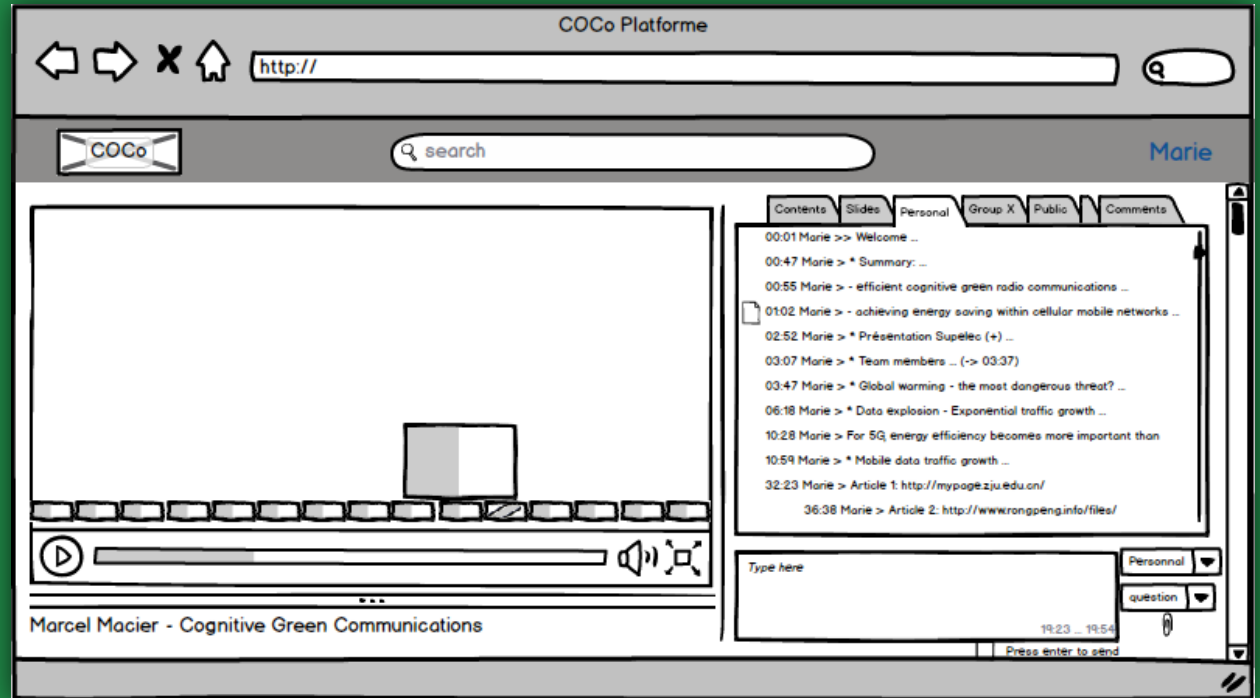


TRACES

- ✓ haut niveau et bas niveau
- ✓ utilité pour l'utilisateur
- ✓ navigation dans l'interface



PLATEFORME



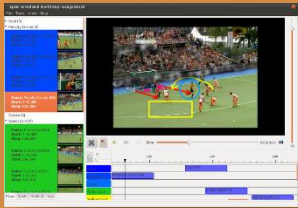
PLATEFORME



MERCI !

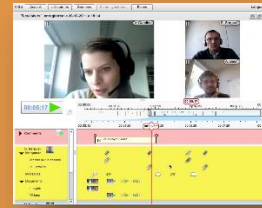
This work has received a French government support granted to the COMIN Labs excellence laboratory and managed by the National Research Agency in the "Investing for the Futures" program ANR-JO-LABX-07-0J.

CRÉDITS



LongoMatch

Capture d'écran de l'outil
<http://longomatch.org/features.php>



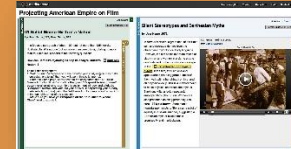
Visu

Capture d'écran de l'outil.
<http://visu.advene.org/>



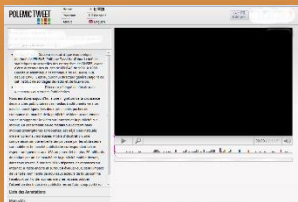
VideoNot.es

Capture d'écran de l'outil avec le film "Of Human Bondage" (domaine public), via [YouTube](https://www.youtube.com/watch?v=...).



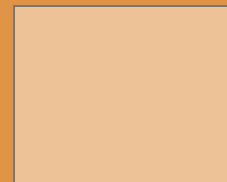
Mediathread

Capture d'écran de l'outil
http://ccnmtl.columbia.edu/our_services/tools/mediathread/example_film_study.html



PolemicTweet

Capture d'écran de l'outil
<http://polemictweet.com/attention-1314-05-marche-attention-avenement-publicite/polemicaltimeline.php>



Autres

Toutes les autres images/figures sont soit des travaux personnels soit appartiennent au domaine public.

LICENCE



Except where otherwise noted, this work is licensed under a Creative Commons Attribution 4.0 International License.

<http://creativecommons.org/licenses/by/4.0/>