

Des folksonomies à la sémantisation des annotations vidéo pour l'analyse filmique

Olivier Aubert*

*Université de Nantes
olivier.aubert@univ-nantes.fr,
<https://www.olivieraubert.net/>

Résumé. L'annotation vidéo est un des outils essentiels pour l'analyse filmique dans le contexte des humanités numériques. Nous décrivons ici comment l'outil Advene a été mobilisé pour soutenir une analyse filmique via la production de données sémantisées.

1 Introduction

Le projet AdA¹ mené au sein de l'équipe Cinepoetics de la Freie Universität de Berlin a pour objectif d'étudier les représentations filmiques de la crise financière, en s'intéressant plus particulièrement aux aspects esthétiques et à l'expression des affects. Pour cela, un corpus de films de fiction, d'actualités et de documentaires traitant du sujet va être analysé de manière systématique et empirique, en utilisant une approche quantitative.

L'objectif final est d'étudier ce que des méthodes automatisées de classification peuvent apporter à l'objet de l'étude. Mais dans un premier temps, il est nécessaire de constituer une base de référence pour l'analyse du corpus filmique. Pour cela, l'ensemble des films du corpus est annoté manuellement et finement, avec un vocabulaire contraint issu d'une ontologie. L'objectif du présent article est de décrire la mise en oeuvre du processus d'annotation manuelle au sein de l'application d'annotation vidéo Advene. Nous verrons les adaptations et extensions apportées au logiciel pour permettre d'un côté de constituer une interface de saisie adaptée aux besoins utilisateurs (analystes de film) et d'un autre côté d'obtenir des données aptes à être mobilisées dans une base de données sémantiques, afin de permettre l'étude de méthodes automatisées. L'enjeu est ici en termes de processus de création de métadonnées et d'Interface Homme-Machine, ainsi que de visualisation des données.

Le logiciel Advene² est une plate-forme d'annotation vidéo, permettant d'annoter des vidéos, puis d'utiliser ces annotations pour naviguer au sein des vidéos et pour générer de nouveaux documents, appelés hypervidéos. Parmi les logiciels d'annotation vidéo (ELAN, Anvil, Transana, Lignes de Temps...), il a pour spécificités son adaptabilité à un grand nombre de contextes, les structures manipulées étant conçues pour être génériques, ainsi que sa capacité à générer des hypervidéos. Ses fonctionnalités correspondaient déjà en grande partie aux besoins du projet. Il a été étendu pour faciliter l'annotation manuelle, via la possibilité d'annoter rapidement par des mots-clés prédéfinis et l'ajout de plusieurs modules de validation des données

1. *Affektrhetoriken des Audiovisuellen* <http://www.ada.cinepoetics.fu-berlin.de/>

2. <https://www.advene.org/>

Des folksonomies à la sémantisation des annotations vidéo pour l'analyse filmique

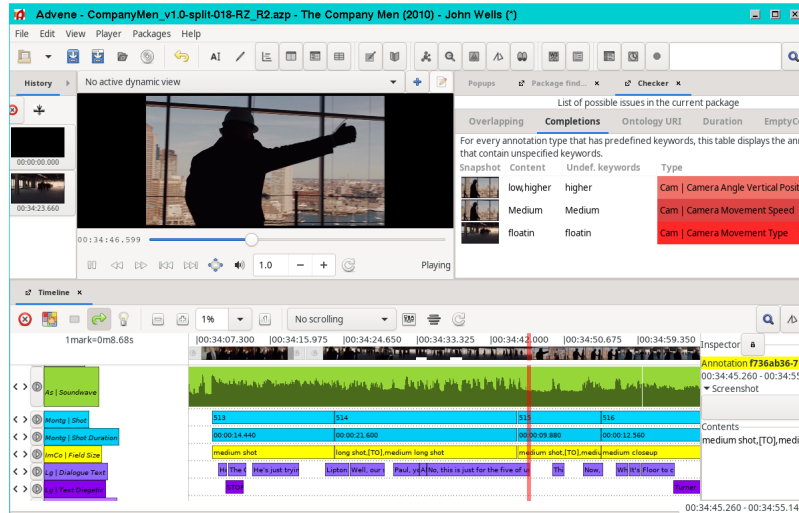


FIG. 1 – Analyse vidéo avec Advene, montrant une représentation sous forme de timeline ainsi qu'un module de validation des données.

(validation à posteriori et non impérative, qui permet de déroger au modèle lors des phases transitoires d'élaboration des schémas d'annotation). D'autre part, l'interaction avec une base de données sémantique a été assurée par l'ajout de fonctionnalités d'import/export de données RDF.

Les mots-clés prédéfinis, au sein de chaque type d'annotation, constituent une folksonomie sous la forme d'un vocabulaire contraint, accessible aux utilisateurs. Des métadonnées spécifiques à chaque type permettent d'associer une URI à chaque mot-clé, afin de générer des documents dans un format RDF WebAnnotation étendu par des propriétés spécifiques au projet. La possibilité d'avoir des valideurs souples au sein d'Advene a permis l'élaboration progressive par enrichissement et affinement des informations de l'ontologie.

2 Conclusion

Le logiciel Advene a été utilisé et étendu au cours du projet pour faire le lien entre des utilisateurs non-experts (analystes) et l'équipe chargée de travailler sur la sémantisation des données et l'extraction automatique d'indicateurs. Le travail va se poursuivre notamment autour de la visualisation des données à des fins d'analyse et de communication, particulièrement sur la notion de rythme.

Summary

We describe how the Advene video annotation platform has been used to produce semantic metadata that will be used for movie analysis.