

Sujet de stage - 2015

Structuration de séries TV par les sous-titres

Encadrement : Xavier Bost et Georges Linarès

Les séries télévisées ont souvent des structures narratives assez contraintes, dans lesquelles plusieurs histoires sont développées en parallèle. Les outils de navigation dans ces documents doivent disposer de descripteurs caractérisant ces structures narratives. Les *histoires* peuvent être caractérisées par des indices visuels, auditifs, sémantiques, etc... Le sujet de ce stage concerne le désentrelacement des histoires à partir des contenus linguistiques. On analysera les sous-titres d'une série pour tenter de segmenter l'ensemble des épisodes en histoire.

Ce travail fait partie d'un projet plus global, soutenu par l'Université d'Avignon et des Pays de Vaucluse et par l'Agence Nationale de la Recherche, dont le sujet est le ré-éditorialisation de vidéos. Les expériences seront réalisées dans le cadre de ce projet, en particulier sur les séries *Games Of Throne*, *House of Cards* et *Breaking Bad*, dont les formats sont sensiblement différents.

Le travail consistera d'abord à étudier les travaux précédents publiés sur ce sujet. Des méthodes de segmentation thématique et de catégorisation de textes seront ensuite développées dans le cadre particulier de la segmentation en histoires.

[1] Xavier Bost, Georges Linarès, [Constrained speaker diarization of TV series based on visual patterns](#), IEEE/ISCA Speech and Language Technology (SLT'14), South Lake Tahoe, USA, 2014.

[2] Hervé Bredin, [Segmentation of TV shows into scenes us-ing speaker diarization and speech recognition](#) in Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on. IEEE, 2012, pp. 2377–2380.

[3] Philippe Ercolessi, [Extraction multimodale de la structure narrative des épisodes de séries télévisées](#). IRIT, Thèse de Université de Toulouse sous la direction de P. Joly, H. Bredin, C. Senac.