

Linking Media: adopting Semantic Technologies for multimodal media connection

Dèlia Fernández-Canellas^{1,2}, Elisenda Bou-Balust¹, Xavier Giró-i-Nieto², Juan Carlos Riveiro¹, Joan Espadaler¹, David Rodríguez¹, Aleix Colom¹, Joan Marco Rimmek¹, David Varas¹, Issey Massuda¹, and Carlos Roig¹

¹ Vilynx, Inc.

² Universitat Politècnica de Catalunya (UPC)

Abstract. Today’s media and news organizations are constantly generating large amounts of multimedia content, majorly delivered online. As the online media market grows, the management and delivery of contents is becoming a challenge. Computational approaches can help to overcome this challenge by governing different applications such as content creation, production, search, and its promotion and distribution to different audiences.

In this abstract we present a success story of the adoption of semantic technologies on the aforementioned applications, which are built on top of a semantic tagging framework, based on a Knowledge Graph (KG). The presented pipeline combines multimodal inputs into a contextual entity linking module, which indexes documents and links them to trends and stories developing on the news. We will describe how documents are linked and provided to media producers through Vilynx’s platform, which is currently indexing over 20k media documents a day.

Keywords: semantic web · knowledge graph · linked data · multimedia

1 Motivation

Media producers publish large amounts of multimedia content online - both text, audio and video. In order to be able to explore all this information we need methods to connect multimodal documents. Integrating and linking media documents requires the understanding and extraction of semantics which describe its content with a universal representation. Labels could be used to describe document contents. However, most of the times this data is not labeled or when labeled it does not use standards. Moreover, manually labeling data is extremely expensive and unfeasible in practice, therefore automatic methods for tagging are needed.

Vilynx provides a media platform with semantic solutions to automatically index multimedia documents from a library and generates search and recommendation engines by linking them to other contents, trends and to stories developing in the news. The user interface displays in an intuitive manner the links between media documents and stories and allows navigation through related content by using associated semantic tags. This interface is a powerful industrial tool for publishers to index, retrieve and visualize their contents. It helps them identify

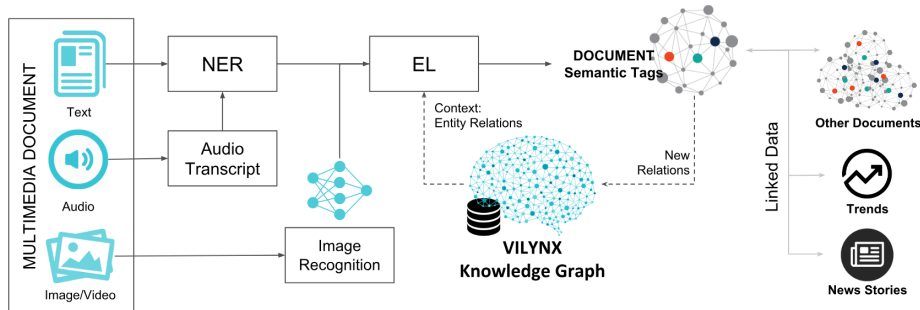


Fig. 1. Scheme of Vilynx’s multimodal document tagging framework.

which topics require more attention, or retrieve related content that has already been published about the stories. Moreover, recommendation and search tools are build on top of the detected semantic entities and integrated on customer’s web pages.

2 System Overview

In this section we give a brief explanation of the framework that powers this media linking platform (shown in Figure 1). The media documents are indexed with a rich collection of tags associated to KG entities. The system provides tags from three different sources: visual, audio and associated text. The visual tagging algorithm provides detection of the people and places appearing on the video using deep learning analytics. The tags from the text block are generated by parsing document web pages and applying a Name Entity Recognition (NER) module to extract mentions. The audio transcript from videos is obtained through speech to text algorithms, and the mentions in the audio are extracted using NER, as for the text source. Finally, mentions and entities extracted from the multi-modal sources are linked to KG entities by the Entity Linking (EL) module, using the entity’s relations as context. Moreover, entities’s relations are updated on the KG when new relations are found on tagged documents.

Once semantic tags are assigned, the document becomes part of the linked data space. This allows us to relate it to other documents, trends and associated stories developing on the news. Due to space considerations, we have omitted the details of our system. For more information on the KG, the text tagging algorithm and the tags fusion see [1].

3 The Media Platform

Vilynx’s platform displays contents from the customer library of multimedia documents, the top trending entities from social networks, and the top trending stories on the news. This platform allows navigation through the costumer contents using semantic links. Also, search tools with filters are developed for fast access to information. For more information on the platform, see [2].

References

1. Fernández, Dèlia, et al. "ViTS: Video tagging system from massive web multimedia collections." Proceedings of the 5th Workshop on Web-scale Vision and Social Media (VSM). IEEE Press, 2017.
2. Fernández, Dèlia, et al. "What is going on in the world? A display platform for media understanding. Proceedings of the 1st International Conference on Multimedia Information Processing and Retrieval (MIPR). IEEE Press, 2018.