

Candidat au prix Pierre-Ardouin 2018

GLO-3002 - Projet en génie logiciel

Par Vincent Aubé, Magalie Caouette et Isabelle Chum-Chhin

Description du projet

En partenariat avec l'entreprise Poka, notre projet est un service qui permet de **générer des sous-titres automatiquement** à partir d'une vidéo et de les **traduire**. Notre client offre une plateforme « sociale » au monde industriel permettant le partage de connaissance, ce qui facilite la formation des employés.

En plus de faciliter la communication, il est également possible de traduire les sous-titres pour rendre la vidéo **accessible à une plus grande diversité de personnes**. Par exemple, un vidéo filmé en anglais pourrait être traduit en russe et utilisé de nouveau pour la formation de travailleurs en Russie. C'est une solution qui **promouvoit le partage de connaissances** de manière optimisée et efficace pour les manufacturiers ainsi que pour Poka.

Résultats

Génération de sous-titres

Les résultats que nous avons obtenus à partir des différentes plateformes (Google, Bing, IBM, Amazon) nous montrent que les sous-titres générés sont similaires, mais que chacun a leurs forces et faiblesses.

Google est en général excellent et est apte dans plusieurs langues (du moins celles que nous avons testées). **Watson**, malgré certaines lacunes en français, est capable de détecter les hésitations dans le texte. Quant à **Amazon**, celui-ci ne permet que l'anglais et l'espagnol, mais ajoute de la ponctuation, ce qui pourrait être intéressant pour les algorithmes de génération de sous-titres. Malgré de bons résultats de la part de **Bing**, nous avons dû le retirer étant donné le manque de *timestamp* (quand nous l'avons testé). Cependant, il serait possible de faire un client dans le futur ou de développer des algorithmes *NLP* plus sophistiqués afin de combiner plusieurs résultats (ce n'était cependant pas dans la portée du projet).

Traduction de sous-titres

Les résultats des traductions de sous-titres sont fortement influencés par la qualité des sous-titres obtenus lors de l'étape de génération.

Nous constatons dans tous les cas à quel point la traduction manque de contexte. Même si elle peut faire du sens, beaucoup de mots sont mal traduits. Elle nous montre l'importance de bien corriger, à la main, les sous-titres générés dans la langue d'origine avant d'essayer de traduire ceux-ci.