

PROGETTO M.U.S.E

IL PROGETTO

Il progetto di ricerca M.U.S.E. è frutto della collaborazione fra BNova e il Laboratorio di Linguistica Computazionale del Dipartimento di Filologia Letteratura e Linguistica dell'Università di Pisa e prevede l'analisi semantica automatica di materiali digitali multimediali utilizzando tecniche di **Natural Language Processing (NLP)**, **Computer Vision (CV)** e **semantica distribuzionale** per estrarre informazione da testi e immagini e consentire processi di indicizzazione e mining cross-modale.

In una prima fase, il progetto si focalizzerà sull'analisi semantica di testi e di immagini utilizzando principalmente tag, didascalie e sottotitoli delle immagini stesse, mentre in una seconda fase il progetto utilizzerà features estratte dalle immagini che verranno combinate con features distribuzionali estratte da dati testuali attraverso Multimodal Distributional Semantic Models (MDSM), utilizzando anche algoritmi di Deep Learning. ***Questa tecnologia innovativa rappresenta un punto chiave per descrivere le immagini con etichette linguistiche, e in generale per operare collegamenti cross-modali tra dati visuali e testuali.***

Lo studio porterà all'implementazione di un framework per l'analisi integrata di informazioni provenienti da fonti multimodali che potrà essere applicato a qualsiasi settore lo richieda.

MOTIVAZIONI DELLA RICERCA

Un accesso completo a grandi archivi di immagini può essere garantito soltanto studiando in modo contestuale le immagini e le informazioni testuali ad esse collegate. Tipicamente lo studio delle immagini afferisce all'area di ricerca della Computer Vision (CV), mentre l'analisi dei testi avviene nell'ambito del NLP. Con l'analisi semantica multimodale, d'altro lato, è possibile non solo migliorare la qualità delle rappresentazioni semantiche, ma anche collegare immagini e testi attraverso un processo di associazione tra rappresentazioni vettoriali nello spazio delle immagini e rappresentazioni vettoriali di loro descrittori linguistici nello spazio testuale. La comunicazione è infatti intrinsecamente multimodale.

OBIETTIVO GENERALE

Obiettivo generale del progetto è definire una metodologia per sfruttare le sinergie tra NLP e CV dal momento che da un lato i testi posti a corredo delle immagini possono essere molto utili per interpretazione di queste ultime e, dall'altro, le informazioni estratte dalle immagini possono migliorare le performance di task di NLP, come per esempio l'information extraction.