Title: Toward concept visualization through image generation

Author: Tien Dat Nguyen

Department: Institute of Formal and Applied Linguistics

Supervisors: Pavel Pecina (Charles University in Prague), Angeliki Lazaridou, Raffaella Bernardi, Marco Baroni (University of Trento),

Abstract: Computational linguistic and computer vision have a common way to embed the semantics of linguistic/visual units through vector representation. In addition, high-quality semantic representations can be effectively constructed thanks to recent advances in neural network methods. Nevertheless, the understanding of these representations remains limited, so they need to be assessed in an intuitive way. Cross-modal mapping is mapping between vector semantic embedding of words and the visual representations of the corresponding objects from images. Inverting image representation involves learning an image inversion of visual vectors (SIFT, HOG and CNN features) to reconstruct the original one. The goal of this project is to build a complete pipeline, in which word representations are transformed into image vectors using cross modal mapping and these vectors are projected to pixel space using inversion. This suggests that there might be a groundbreaking way to inspect and evaluate the semantics encoded in word representations by generating pictures that represent it.

Keywords: text2image, Cross-model Mapping, Distributed Semantics, Convolutional Neural Networks, Visual Feature Inversion.