

# KEYWORD EXTRACTION AND IMAGE ANNOTATION GAMES TO ENHANCE THE CULTURAL DATABASE CREATION

*Virach Sornlertlamvanich and Thatsanee Charoenporn*

National Electronics and Computer Technology Center, Thailand  
virach.sornlertlamvanich@nectec.or.th, thatsanee.charoenporn@nectec.or.th

ESP and Peekaboom games are successfully done to engage annotators in labeling the images and locating the objects in the images. The advantage of these games is making annotators feel like enjoying a game rather than working hard to label the ton of images. We extended Peekaboom game by introducing Asian WordNet (AWN) to expand the guessed label in terms of its synonym, hypernym, hyponym, meronym, and holonym. The keyword for labeling is also extracted from its title and description according to the term frequency TF-IDF weighting manner. We found that the languages such as Thai, which has no word boundary, yield difficulty in making the candidate term list. Therefore, we additionally need word segmentation and keyword chunking model to generate the possible terms. In our preliminary experiment, the result yields in a higher matched ratio up to 22%. Moreover, AWN will also provide the semantic link between 13 Asian languages and the English language. Our method also shows a potential in creating cross-language image labeling game.