# Psychological Research with Social Media Posts and Computational Text Analysis
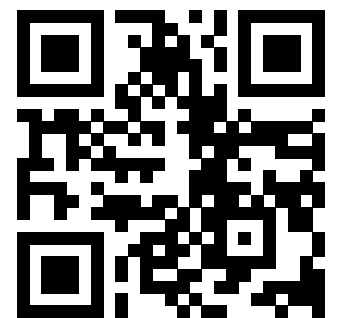
## Prof. Lee-Xing YANG
### National Chengchi University, Taiwan

**Workshop 3**

**Date:** August 12, 2021 (Thursday)
**Time:** 1:00pm - 2:30pm
**Venue:** *Seminar will be conduced via ZOOM

**Join the Zoom Meeting**

## Abstract

Nowadays people are used to sharing their lives, thinks, and emotions on social media, which in turn becomes a gigantic online database of human behaviors. A number of recent studies have shown that the posts on social media can be used to predict psychological attributes of humans (e.g., personality). Although these studies were conducted for different goals, the procedures of them can have a common three-stage form. The main purpose of this workshop is to introduce this three-stage procedure for studying the posts on social media: data collection (e.g., using the web scraper to collect the posts of interest), data preprocessing (e.g., removing irrelevant HTML codes and extracting keywords or topics), and data analysis (e.g., providing interpretations to the observed data with statistical models). Specifically, as the posts are all textual, the techniques of computational text analysis will be introduced. Text analysis can be very different in the fields of psychology and machine learning. In the first example, I will compare these two types of text analysis on detecting the gender differences in the relationship breakup posts. The results suggest that these two approaches of text analysis can be integrated to gain a more complete understanding of gender differences. In addition to extracting key words, text analysis can also be used to generate the topics of texts. In the second example, I will show that topic generation is a better solution to detect the gender differences in the relationship breakup posts than keyword extraction.

## Speaker

My name is Lee-Xieng Yang and I am now an associate professor in the department of psychology of National Chengchi University in Taiwan. My research basically can be summarized as two parts. The first part includes cognitive psychology and mathematic psychology. Specifically, I am particularly interested in understanding how people learn categories and form concepts with laboratory experiments and mathematical models, such as neural network models and Bayesian models. In addition to the on-going studies on human categorization, recently, I also start to examine how to attenuate or eliminate the continued influence effect of misinformation (or fake news) in respect of the encoding and retrieval of memorized materials. The second part is more relevant to developing research methods for social media study. As most of the contents on social media are textual, developing the methods of text analysis becomes my main concern. Specifically, I put emphasis on applying the Dirichlet process in Bayesian methods to develop mixture models to summarize texts by representative topics. Recently, I also try to improve the sentiment analysis for texts via examining the possibility of using verbs together with adjectives to identify the sentiment of a text.

### All are Welcome