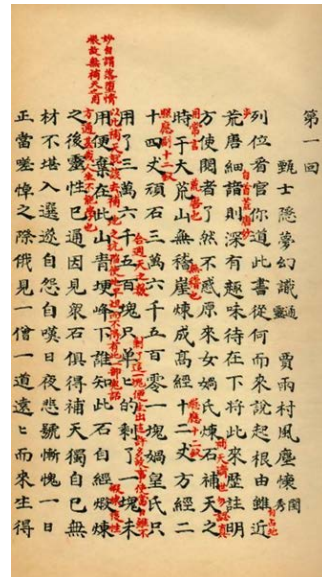# "Some Applications of Text Mining for Historical and Literary Texts in Chinese"

by

Prof. Chao-Lin Liu,
National Chengchi University,Taipei, Taiwan
Fulbright and TUSA Scholar, Institute for Quantitative Social Science, Harvard University

April 26, 2017
Noon, Cohen House

Abstract

Techniques for text mining are essential for the research of digital humanities. In this brief presentation, I review some applications of text mining for the analysis of the Chinese poetry and the enrichment of the China Biographical Database, which is hosted at the Harvard University. I discuss the applications of named-entity recognition and a basic form of grammar induction for extracting biographical information from classical Chinese texts, including local gazetteers and tomb biographies. Biographical information about historical figures proves to be instrumental for the studies of the history of words and their patterns in Chinese poetry. I will also extend our adventure into the Chinese poetry to explore the issues of stylometry, imageries, social network analysis based on some statistical analysis and string comparison methods.