

“Mining Chinese Historical Sources at Scale: A Machine Learning Approach to Qing State Capacity”

Speaker: Prof. Carol SHIUE

Abstract

To understand the ways in which social unrest was reported by Qing officials, we train different language models based on nearly one-thousand existing samples describing different kinds of protests in the *Qing Shilu*, the Veritable Records, which is based on court documents of the Qing Dynasty. We then apply the model to the full text of the *Qing Shilu* to extract all the records about unrest in the empire. Distinguishing different types of unrest (peasant, militia, or secret society) allows us to have a more precise understanding of what the Qing rulers saw as a threat to the legitimacy of the dynasty, and provides new evidence on elements of state capacity.