Abstract:
What are the ethical implications of carrying out digital humanities research in so-called “Islamic languages”? Funding for topics like sentiment analysis, natural language processing, and machine translation has been abundant—especially from state organizations. How does one avoid the instrumentalization of this research when even apparently innocent projects can have applications whose implications would be impossible to imagine in contexts involving European languages? This paper will use several case studies from natural language processing and affiliated fields to evaluate whether such work is primarily concerned with fostering broader cross-cultural communication or domain-specific, and even invasive, forms of knowledge. How can digital humanities research enable a more ample range of expression and construct more “authentic” representations of South Asian emotions? Examples will be drawn from a now-defunct music recommendation service, an Urdu sentiment-extraction system for social media, and a transliteration project with which the author is currently involved.