Digital Apprehensions of Poetics

Professor A. Sean Pue, Michigan State University

SPARC Course

Course Description

It is clear that knowledge of Indian poetics—the theory of literary forms and discourse that allows for understanding of poetry—is dwindling in our contemporary age. At the same time, poems themselves are circulating more widely than ever before, primarily on digital platforms. This course, "Digital Apprehensions of Poetics," will explore how we can use digital technologies to understand, interpret, and annotate the poetics of Indian literatures, which circulate in digital texts, in manuscript, and as oral or musical performance.

The course begins with a general overview of digital texts and media on the internet, then proceeds to a discussion of scholarly critical editions, in general, and of digital critical editions, in particular. We then cover different types of annotation—metadata used by libraries and the semantic web, manuscript annotation that can be used to train optical character recognition, and audio and video annotation, including that used in linguistics. Building on this general background, we then turn our attention to Indian poetics, providing an overview of poetic forms across Indian languages, and discussing key issues, such as poetic meter. We then consider how computational modeling and annotation can be applied to Indian poetries. The course concludes with a discussion of user experience design and a group project in which students both design an interface to present and model poetics and conduct experiments to see how their designed tools can enhance the apprehension of poetics.

Prerequisites:

The course assumes no prior knowledge, and is intended for graduate students and advanced undergraduates. Students can come from humanities, social sciences, media and the arts.

Classroom Requirements

An internet-connected laptop or desktop computer (Apple, Windows, or Linux—no Chromebooks) will be required during class lectures and labs. We will exclusively use free and/or open source software in the class.

Course Contents:

Week 1

Day 1. Digital Texts and Media on the Internet: An Overview

- What is Digital Humanities?
- What is a digital text?
- How can we search texts using regular expressions?
- What is a web server and how can we connect to one?
- How is a web page rendered?
- What is a stylesheet?
- How is media stored?

Day 2. Digital Critical Editions (2 hours)
- What is the scientific method used in the study of manuscripts?
- How does it relate to Indian traditions of textual scholarship?
- What is the Textual Encoding Initiative (TEI)?
- How is TEI used to create digital surrogates of printed books?
- How does TEI relate to Extensible Markup Language (XML) and Extensible Stylesheet Language Transformations (XSLT)?
- How does one collaboratively create a digital edition?

Day 3. Metadata and Digital Annotation (2 hours)
- What is metadata (data about data)?
- What is linked data?
- What is the Resource Data Framework (RDF)?
- What is an Authority File, and how do international libraries share information?
- What is an RDF triple (involving a subject, predicate, and object)?
- What is the semantic web?
- How can we query RDF triples using the SPARQL query language?

Day 4. Syntactic and Semantic Annotation of Digital Text (2 hours)
- What is a treebank (parsed text corpus)?
- What is a lemma (dictionary form of a word)?
- What forms of treebanks have been applied to Indian languages (Paninian and Universal Dependencies)?
- What is semantic annotation?

Week 2

Day 5. Manuscript Annotation (2 hours)
- How do we annotate scanned digital versions of manuscripts and books?
- What is keyword spotting?
- How can manuscript transcription be used to train optical character recognition (OCR) software?
- How can these annotations be converted into TEI?
- What are examples of successfully crowdsourced annotation projects?

Day 6. Audio and Visual Annotation (2 hours)
- What are the physical components represented by an audio or visual recording?
- How do we annotate audio and video recordings?
- How do we align transcriptions of audio and visual recordings with digital texts?
- What features of linguistic annotation (fundamental frequency, phoneme boundaries, intonation and stress markings) can be useful for understanding poetics?

Day 7. Indian Poetics and its Computational Modeling (2 hours)
- What are the key features of Indian poetics?
- What forms of poetry are found across Indian languages?
- What is poetic meter?
- How does poetic meter relate to the linguistic syllable?

- What systems are used to denote poetic meter?
- What is a computational model?
- How can we model poetry?
- How can we represent poetic meter as a graph?
- How can we visualize poetic meter?

Day 8. User Experience Design and Group Project Discussion (2 hours)
- What is user experience design?
- How can we use experiments to design user interfaces?

Group Project (Lab Component), 20 hours

Design a user-interface to present and model poetics. Conduct experiments, using audio or video recording, to see how graphical interfaces can enhance the apprehension of poetics.