This course is a project-based introduction for CS / software engineering students to the field of digital humanities. Digital humanities is a developing field of knowledge that deals with the
intersection of methodologies in humanities and computing to describe and study digital collections of cultural data. The development of digital humanities has been triggered by the rising amount of digital texts and multimedia artifacts annotated by historical information, the growth of the Internet and the Semantic Web, and the establishment of their concepts.

In this course, students will read critical and technical papers about methodological questions, such as ‘what is the difference between research in the humanities and research in exact sciences’ and will learn about traditional tools and concepts developed in each field to describe cultural works - curation, preservation, classifying, authorship, genre.

Students will design and work on software projects that utilize and implement tools for solving specific research questions or address need in the humanities.

Students will design and work on software projects that utilize and implement tools for solving specific research questions or address need in the humanities. The course will cover topics about attributes of texts, annotation of cultural information, classification, as well as integration and visualization of information from multiple knowledge sources. Tools in Machine Learning, Text Mining and Natural Language Processing will be taught. All participants in the class will be expected to produce a software project based on the use of digital tools and techniques by the end of the semester.

The course will cover topics about attributes of texts, annotation of cultural information, classification, as well as integration and visualization of information from multiple knowledge sources. Tools in Machine Learning, Text Mining and Natural Language Processing will be taught. All participants in the class will be expected to produce a software project based on the use of digital tools and techniques by the end of the semester.

The course will focus on four major axes:

- Humanities issues (main concepts in literature theory: narrative, authorship, genre)
- Digital Humanities (archiving metadata, digital libraries and curation, distant reading).
- Basic computing skills including quantitative textual data analysis, scientific visualization, basic statistical models - multinomial distributions, generative models, topic models and classification models.
- Programming Skills (using Data analysis tools in Python).

The focus in this course will be on Hebrew literary texts such as those found in benyehuda.org project and additional sources such as the Talmud.
The course will focus on working with sources and sources of information in Hebrew, such as texts that exist in the project by Yehuda Ben.

www.benyehuda.org

or the Talmud/Mishna.

The course contents—
- Introducing to Digital Humanities: platforms, suggestions, the movement of the humanities
- Working on a free platform (cluster, color, and source code)
- Comparing to the models (sample, catalog, etc.)
- Digital humanities and social networks
- Methods for visualization, network analysis
- Representation of information (geographic)
- Requirements of the course

Participation in the course is recommended and is a part of the calculation of the grade.

Class lecture, reading of articles, completion of programming tasks and a final project (which can be done in groups).

The course evaluation criteria are:
- Participation - 10% of the course
- Exercises (writing/reading) - 40%, the final project evaluation up to 50%.

The main references for the course (partially):


Digital Humanities 2.0: A Report on Knowledge, Todd Presner, http://cnx.org/content/m34246/latest/
