

Cutting-Edge Computer Applications for Jewish Heritage Research

In this session three sophisticated and ground-braking projects will be presented. All of the three projects are using cutting-edge computer technologies and algorithms to process Jewish manuscripts, from the Dead Sea Scrolls to medieval manuscripts in the Genizah and elsewhere.

Nachum Dershowitz and Lior Wolf

From Caves to Cyberspace: AI Aids in the Study of the Dead Sea Scrolls

We have developed several computerized tools for the analysis of the new multispectral images of the Dead Sea Scrolls. Using multiple wavelengths from old and new images allows us to produce a more informative binary image . Combining results of several different binarization methods helps build a more accurate separator of text from background. We adapted a keypoint detection method based on the black regions in the binarized image to locate letters and then character spotting is used to search for similar letters to a given character image. Our approach is capable of distinguishing between different scripts in the scrolls and may help identify fragments written in the same hand.

Another important challenge is that of aligning transcript letters to their coordinates in manuscript images. We directly match the historical image with a synthetic one created from a transcript, rather than attempting to recognize individual letters using OCR. This method is robust with respect to document degradation, variations between script styles and non-linear image transformations.

Yaacov Choueka

The Friedberg Project for Variant Readings of the Babylonian Talmud

Establishing and effectively displaying the reading variants of the Babylonian Talmud is the mission of the project presented in this talk. The aim is to build a freely accessible website where high quality digital images, as well as their carefully checked transcriptions, of all manuscripts, fragments and first editions of the Talmud, available anywhere, as well as various and novel ways for presenting these variants, including new formats of synopsis that can be dynamically tailored to the needs of every user, will be available in the website.

Roni Shweka

Joining the Cairo Genizah Fragments: Project Report

One of the main obstacles which stand in the way of every Genizah researcher is the fragmentary state of the manuscripts. A typical Genizah fragment is no more than a single leaf, usually torn and damaged. Moreover, the scattering of more than 300,000 Genizah fragments in some 60 different libraries around the world adds another significant difficulty. A researcher studying a given fragment can never tell if there some other fragments in the same library or in another collection, which either complements the missing parts of this fragment or represent other folios from the original manuscript .

In the last three years a research group composed of researchers from Genazim, the Computerization Unit of the Friedberg Genizah Project, cooperating with colleagues from the School of Computer Science of Tel Aviv University, has been involved in a research which is addressing this problem of finding "joins" by computational means. A complex computer program was developed which compares pairs of fragments and mark every pair with a similarity grade based on the fragments' handwriting. In this presentation I shall report on the results of running this system on 160,000 fragments that were found suitable for such a comparison, comparing every possible pair of them, altogether about 12.4 billion comparisons. I believe that the results we have obtained will open a new era in the field of Genizah study, and may help us recovering the original state of the "Genizah Archive".