Aharon Shweka, Yaacov Choueka, Lior Wolf and Nachum Dershowitz, "Automatic Computerized Identification of Handwriting and Matching of Genizah Fragments", Ginzei Qedem 7 (2011), pp. 171-208. (Hebrew)

In this paper we present the problem of matching joins of genizah fragments from the perspective of Jewish studies scholars and genizah researchers. We start with remarks of scholars who where aware to this problem more than others, and we give a brief survey of few previous attempts to present in specific subject the genizah fragments joined into reconstructed copies, instead of simple list of shelfmarks. We than proceed to describe how our system can contribute to this task. We distinguish between "strong-join" – two fragments from the same codex, which should be and can be confirmed by other codicological features like page size and number of lines, together of course with correspondence in content, as opposed to "weak-join", e.g. a join of two fragments which are matching by handwriting only but originated from different codices. In the latter case, the certainty of the join is much less determined, as it should be assured by a through paleographic study done by an expert. We explain how the performance of the system is measured, and how we plan to improve it.
In the second part of the paper we introduce 113 join-sets, compound out of 320 genizah fragments, which were matched by our software. We limited the list to joins which were not matched before, and we focused on cases where the fragments were not identified, or were identified erroneously. Thus we demonstrate the importance of matching fragments, as this is an important tool for correct identification of them.