WING.NUS’s Muthu Kumar and previous research assistant Ankur Khanna have been collaborating with Dr Kokil Jaidka, alumnus of Nanyang Technological University (NTU), to annotate and release a development corpus for scientific document summarization. The dataset consists of 10 ACL Anthology papers and the neighborhood of papers that cite them. The annotations indicate the provenance (source) of a citing sentence’s claim in the original paper and what facet it belongs to (from a fixed set) -- following the annotation guidelines for summarization set out by the TAC BioMedSumm task. You can find the corpus and more details on GitHub: https://github.com/WING-NUS/scisumm-corpus