

Similarity matrices with 5% of full-rank of document-token matrices

