

# Text Processing II

## Exercises

Consider these documents:

**Doc 1** breakthrough drug for schizophrenia

**Doc 2** new schizophrenia drug

**Doc 3** new approach for treatment of schizophrenia

**Doc 4** new hopes for schizophrenia patients

1. Compute the term-document incidence matrix for this document collection.
2. Compute the inverted index representation for this collection
3. Compute Merge Algorithm for the query “new schizophrenia”