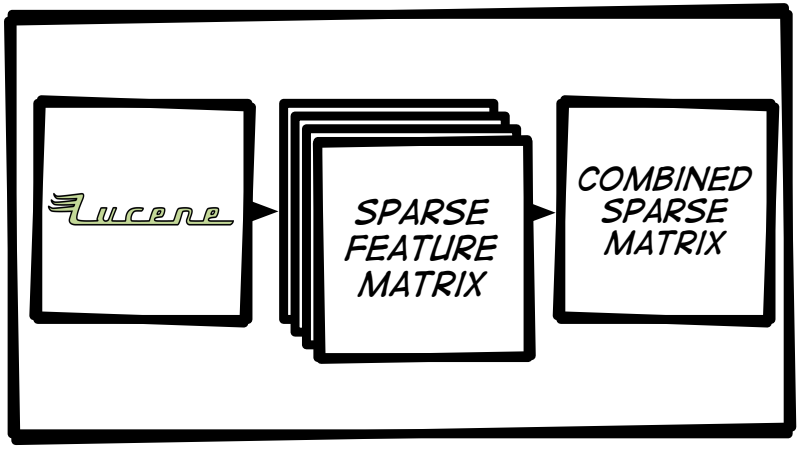
SOTON-WAIS @ SED2013

SPARSE FEATURES AND INCREMENTAL DENSITY BASED CLUSTERING

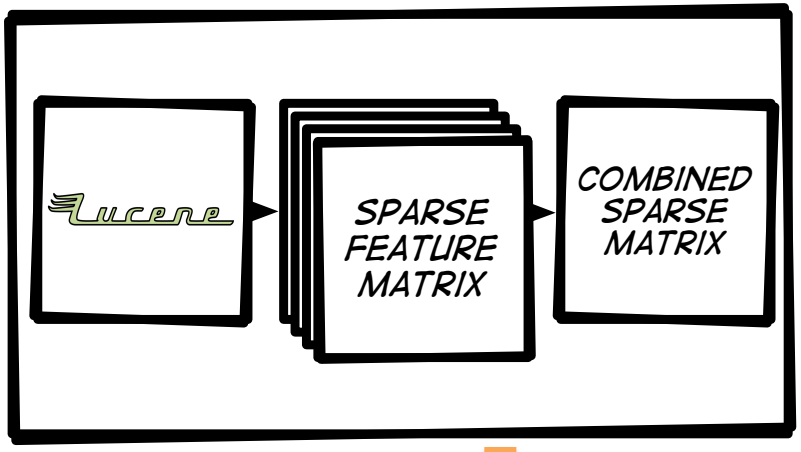
SINA SAMANGOOEI

TEAM SOTON-WAIS (UNIVERSITY OF SOUTHAMPTON)
POWERED BY Lucene OpenIMAJ
SPONSORS University of Southampton arcomem TrendMiner SEVENTH FRAMEWORK PROGRAMME

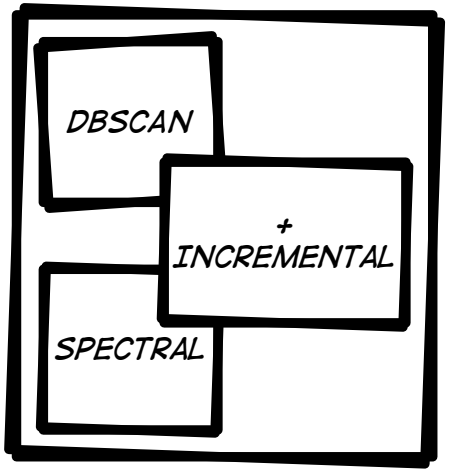
BUILD SPARSE MATRIX



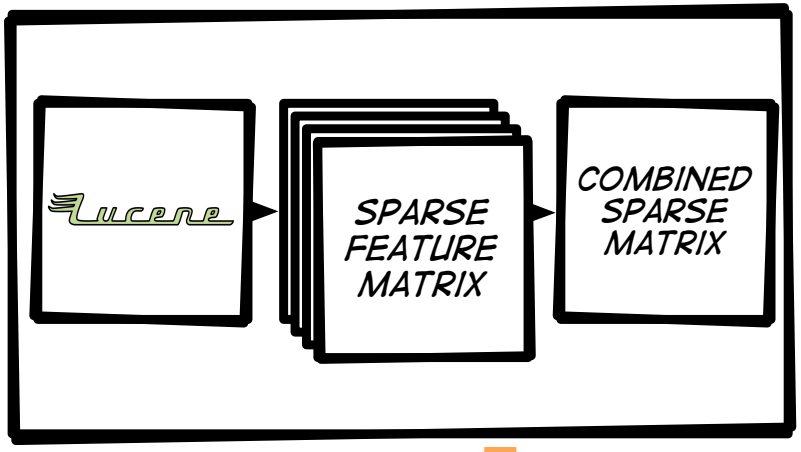
BUILD SPARSE MATRIX



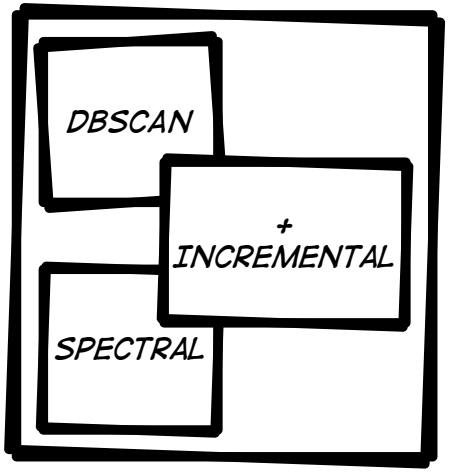
CLUSTER



BUILD SPARSE MATRIX



CLUSTER



MAIN ACHIEVEMENTS

- CREATE ***BIG*** SPARSE SIMILARITY MATRICES
  - (QUICKLY)
- ***OPEN SOURCE***
  - CLUSTERING, EXPERIMENTS… ***THE LOT***
  - ***…OPENIMAJ…*** WOO!
- ***INCREMENTAL*** CLUSTERING
  - SPECTRAL CLUSTERING WITH 300K ITEMS, 0.01 DENSITY.
  - HARD TO DO NAIVELY, TRY IN MATLAB OR PYTHON:

```
eigsh(srand(300000,300000,0.001),1000)
```

ASSUMING YOU HAVE ***128GB RAM*** IT'LL BE DONE SOON(ISH)