SOTON-WAIS @ DIV2013
THE PRE-FILTERS, THE RERANKER AND THE DIVERSIFIER

JONATHON HARE
OVERALL APPROACH

INCREASE PRECISION

INPUT RESULTS LIST

PRE-FILTER

RERANKER

OUTPUT RESULTS LIST

DIVERSIFIER

INCREASE DIVERSITY

BINARY PREDICATES:
FACES, TEXT, BLUR
GEOLOCATION, NUMBER OF VIEWS,
DESCRIPTION LENGTH

LUCENE-BASED PHRASE PROXIMITY SEARCH

GREEDY MIN-MAX SIMILARITY DIVERSIFIER:
CN-CORR, SIFT-LSH
TIMEUSER, MONTHDELTA
MAIN ACHIEVEMENTS

• A PIPELINE THAT INCREASES BOTH PRECISION AND DIVERSITY

• NOVEL DIVERSIFICATION FEATURES

• OPEN SOURCE
  – PARTS ALREADY INTEGRATED INTO OPENIMAJ. REMAINING BITS WILL BE SOON... WOO!