

Summary

- VSM instantiation: dimension, vector placement, similarity
- Simplest VSM
 - Dimension = word
 - Vector = 0-1 bit vector (word presence/absence)
 - Similarity = dot product
 - $f(q,d)$ = number of **distinct** query words matched in d