

Summary of Query Likelihood Probabilistic Model

- Effective ranking functions obtained using pure probabilistic modeling
 - Assumption 1: $\text{Relevance}(q,d) = p(R=1 | q,d) \approx p(q | d, R=1) \approx \mathbf{p(q | d)}$
 - Assumption 2: Query words are generated independently
 - Assumption 3: Smoothing with $p(w | C)$
 - Assumption 4: JM **or** Dirichlet prior smoothing
- Less heuristic compared with VSM
- Many extensions have been made [Zhai 08]