

Central Questions to Ask about a LM: “ADMI”

- **Application:** Why do you need a LM? For what purpose?



Evaluation metric for a LM

Information Retrieval

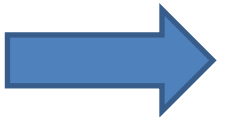
- **Data:** What kind of data do you want to model?



Data set for estimation & evaluation

Documents & Queries

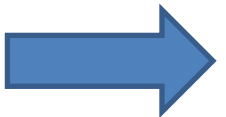
- **Model:** How do you define the model?



Assumptions to be made

Doc. vs. Query generation,
independence

- **Inference:** How do you infer/estimate the parameters?



Inference/Estimation algorithm

Smoothing methods,
Pseudo feedback