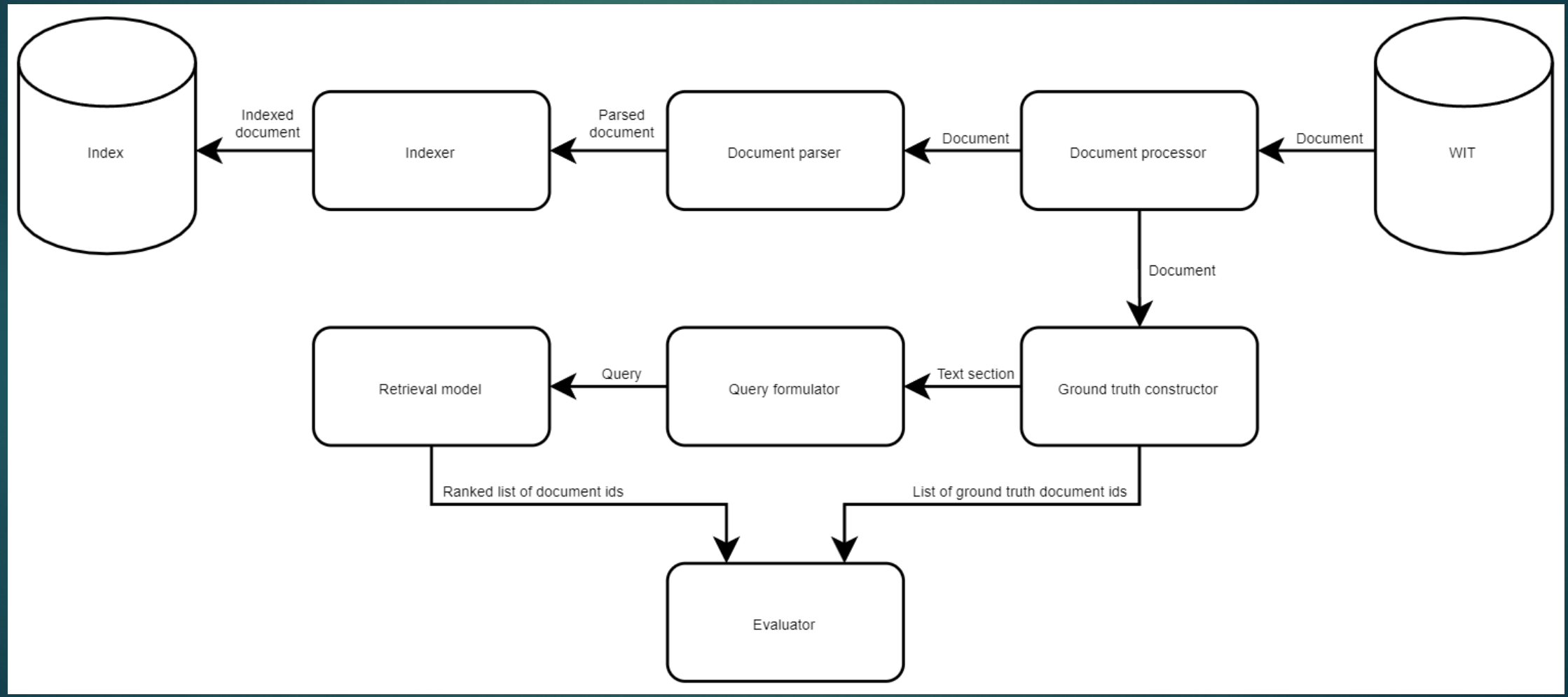
“A Picture is Worth a Thousand Words”: Automatic Illustration of Text via Multimodal Interaction

Introduction

- Motivation
- Problem statement
  - Build the index
  - Select query terms
  - Obtain relevant results
  - Evaluate model
- Contributions
- Research questions
  - Given some context, how effectively can we extract the information need from that context?
  - Given the information need, how effectively can we retrieve the content of different modalities?

Methodology



Experiments

A stairway, staircase, stairwell, flight of stairs, or simply stairs, is a construction designed to bridge a large vertical distance by dividing it into smaller vertical distances, called steps. Stairs may be straight, round, or may consist of two or more straight pieces connected at angles. Special types of stairs include escalators and ladders. Some alternatives to stairs are elevators, stairlifts and inclined moving walkways.

stair, straight, stairlift, vertic, distanc, stairwel, escal, stairwai, inclin, ladder

Evaluation Metric	Value
P@5	0.15
Recall	0.21
MRR	0.28
MAP	0.16

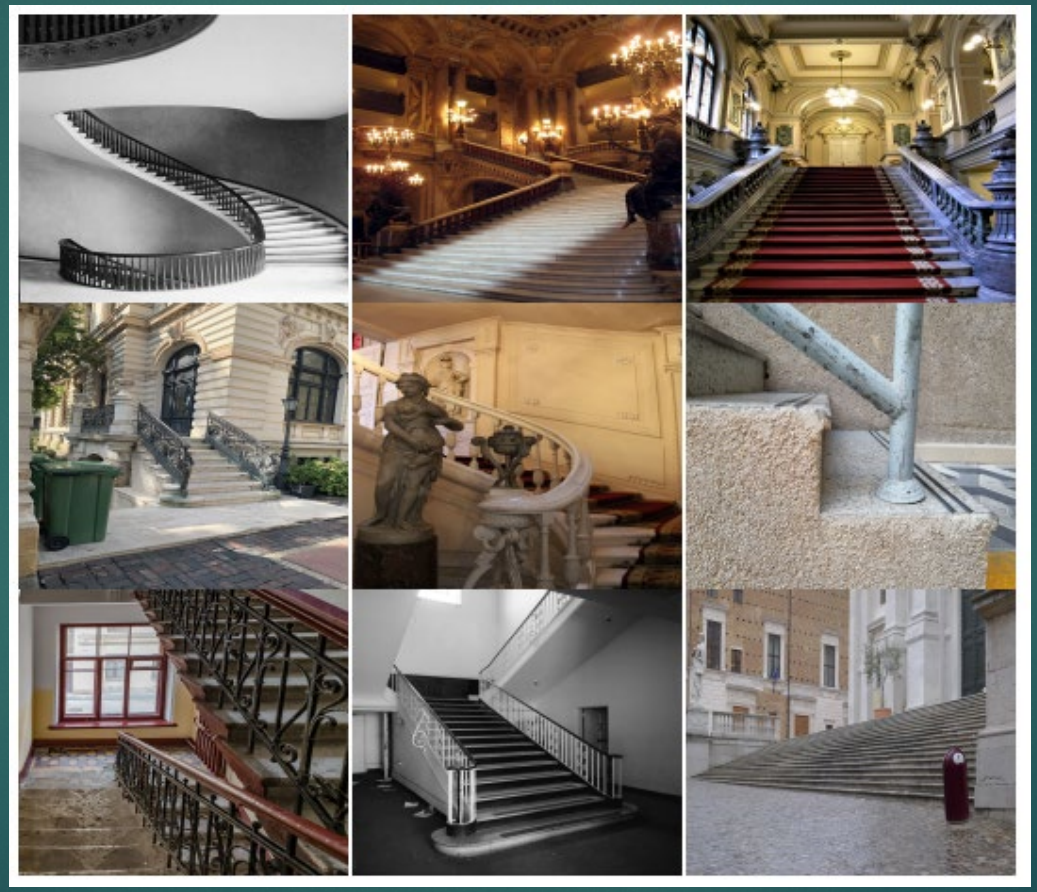
Evaluation Metric	Value
P@5	0.00
Recall	0.04
MRR	0.02
MAP	0.02

Evaluation Metric	Value
P@5	0.30
Recall	0.00
MRR	0.58
MAP	0.27
NDCG	0.12

Evaluation Metric	Value
P@5	0.00
Recall	0.00
MRR	0.10
MAP	0.10
NDCG	0.01

Experiments

Experiments



Experiments

- BM25 parameters
  - k – document term frequency scaling
  - b – term weights' scaling by document length
- Percentage of Query Words

Evaluation Metric	Value
P@5	0.15
Recall	0.21
MRR	0.28
MAP	0.16

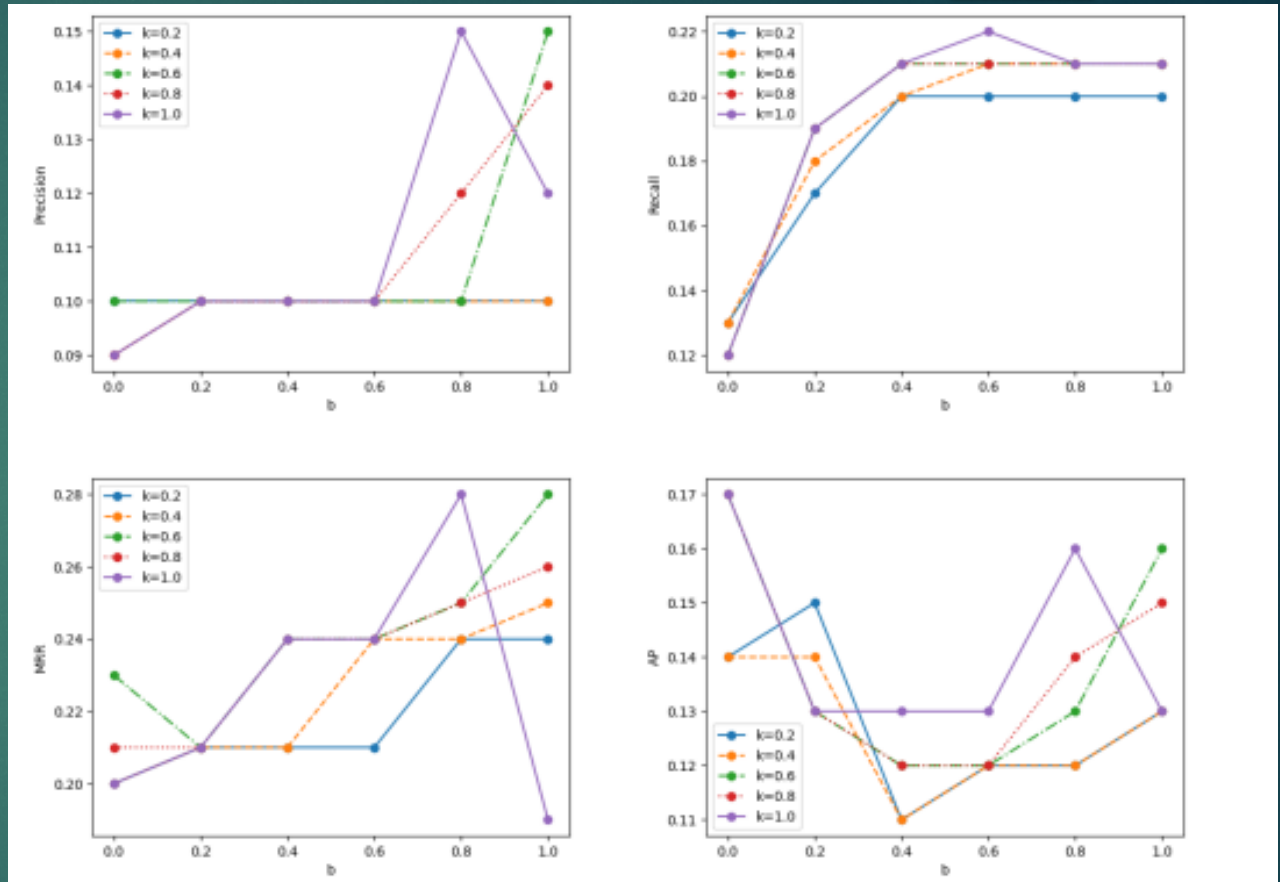


Figure 6.7: Graphs displaying how Precision, Recall, MRR, and AP values vary based on BM25 similarity parameters k and b.

Experiments

Evaluation Metric	Value
P@5	0.15
Recall	0.21
MRR	0.28
MAP	0.16

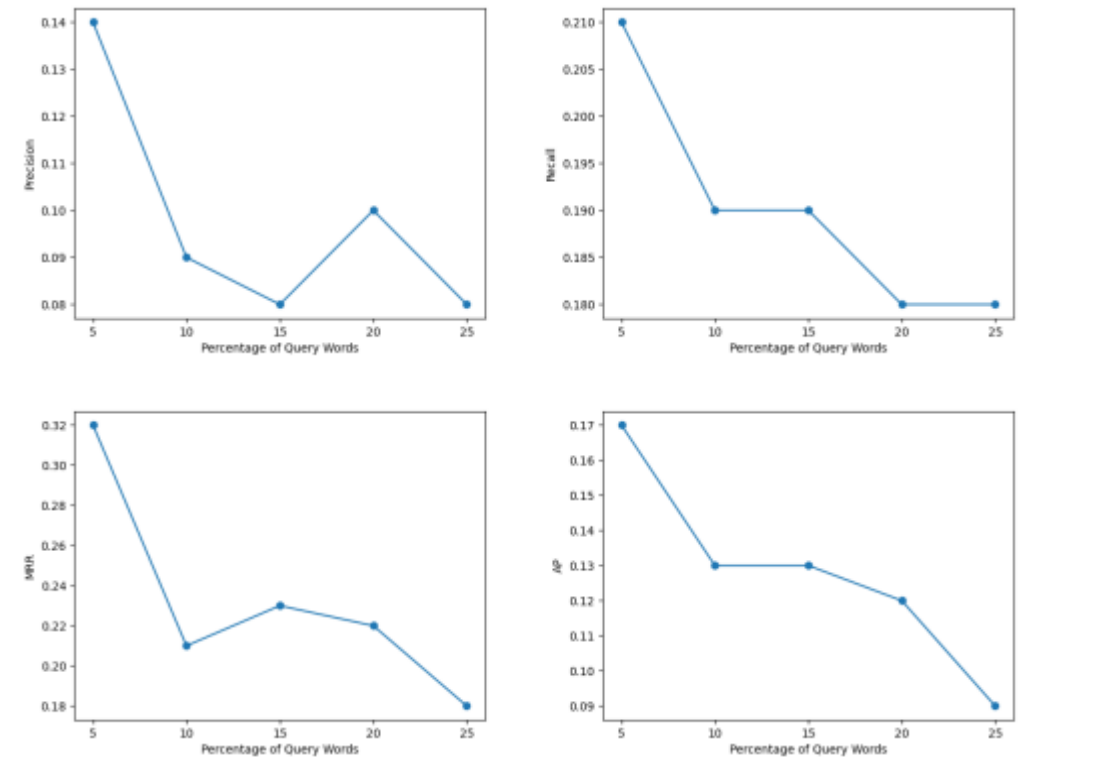


Figure 6.5: Graphs displaying evaluation metrics (Precision, Recall, MRR and AP) values based on Percentage of Query words (from 5% to 25% of text section length)

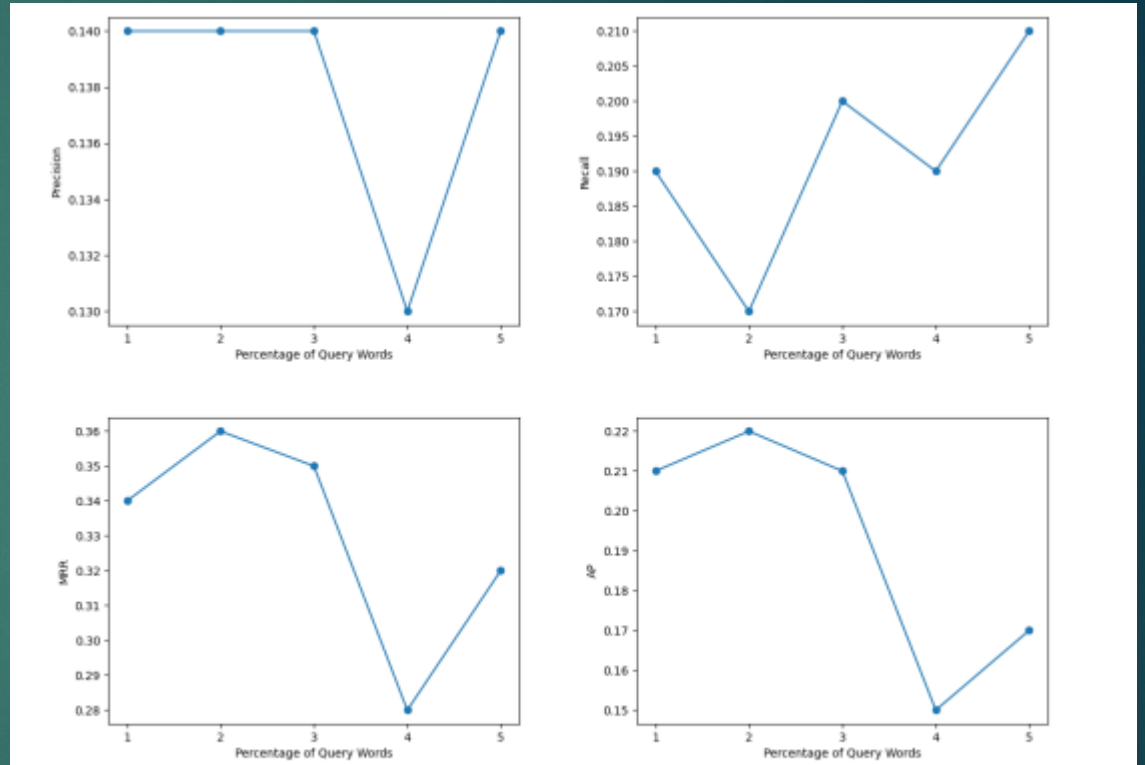


Figure 6.6: Graphs displaying evaluation metrics (Precision, Recall, MRR and AP) values based on Percentage of Query words (from 1% to 5% of text section length)

Summary

- Research questions
  - Given some context, how effectively can we extract the information need from that context?
  - Given the information need, how effectively can we retrieve the content of different modalities?
- Future work
  - Multiple language implementation
  - Query window size
  - GUI
  - Multimodal embedding space