



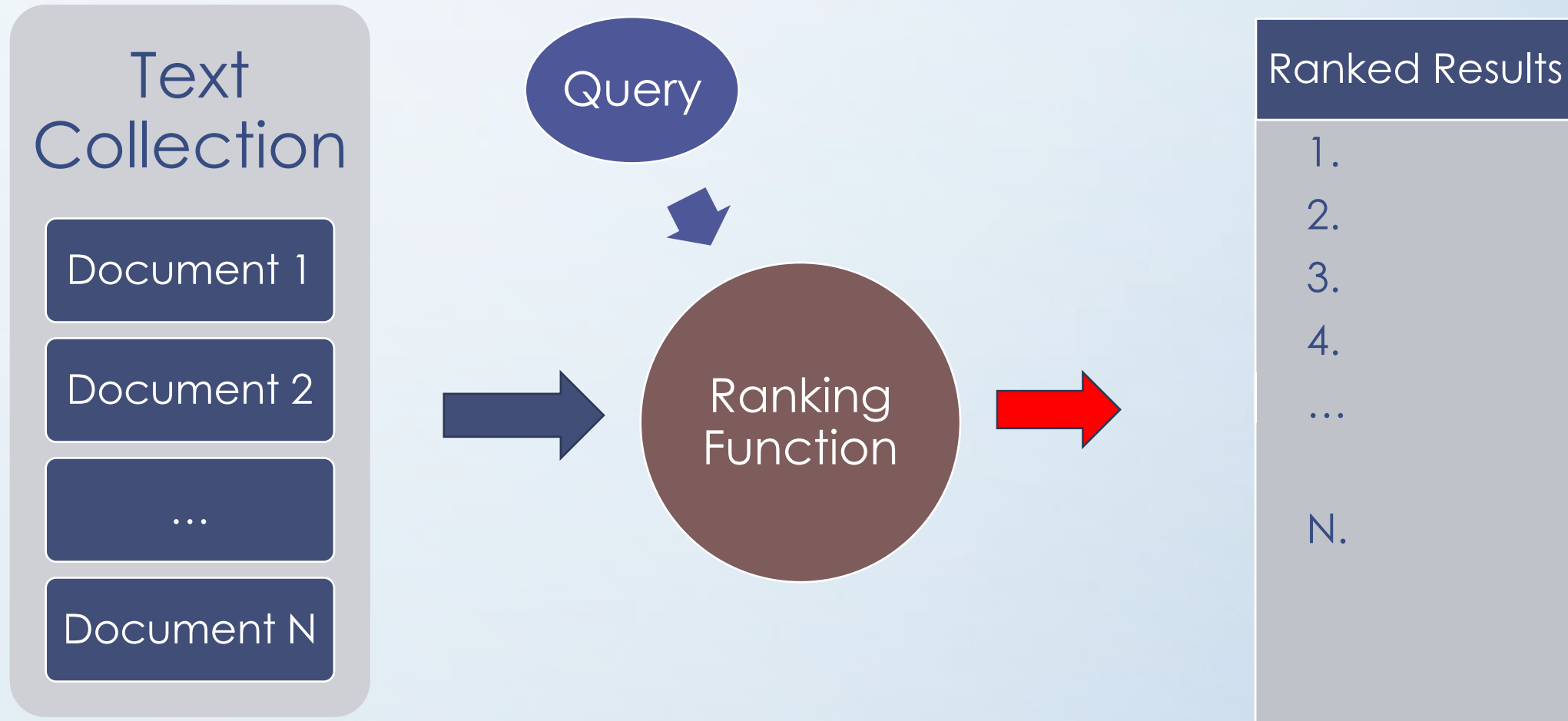
Ranking Segmented documents using Data Fusion

by Hamed Rezanejad

Outline

- Description of the problem
- Motivation/Importance
- Methodology
- Expected results

Description



Description

- **Order** of retrieved documents is very important
- Generally, **Size** of documents differs compare to each other.
- Each document has **different segments** discussing **different issues**
- Using these segments can help us to have **better order** of retrieved documents

Motivation/Importance

- Passage Retrieval
 - ✓ Unit of retrieval is blocks of text from the stored document
 - Current IR systems are used for indexing a great **variety of documents**.
 - For **big size documents**, standard ranking is not of value.
 - Tracking topics in **information feeds**, is a case that standard ranking has nothing to do.

Motivation/Importance

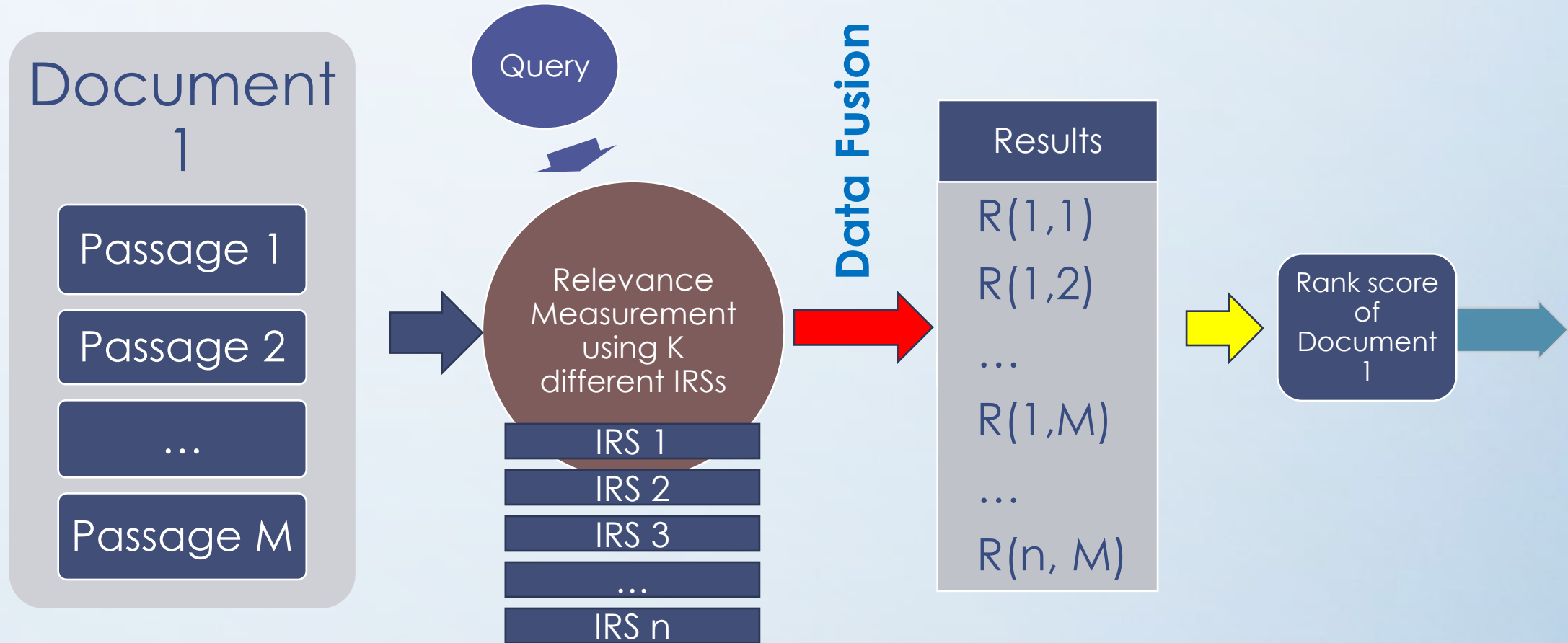
- Data Fusion

- ✓ Accepts two or more ranked lists and **merges** these lists into a single ranked list

Aim of data fusion:

1. Providing **a better effectiveness** than all systems used for data fusion.
2. Grouping existing search services under **one umbrella**.

Methodology



Expected Results

- There is some works about using data fusion on document ranking
- We are going to compare our works with them
- We expect **improvement** in ranking measurements

