Last Code Cell Output

```
Loading and preprocessing documents...
Extracting text from file:
                            bbc articles/innovation/innovation news 10.pdf
Extracting text from file:
                            bbc articles/innovation/innovation news 8.pdf
Extracting text from file:
                            bbc articles/innovation/innovation news 9.pdf
Extracting text from file:
                            bbc articles/innovation/innovation news 4.pdf
Extracting text from file:
                            bbc articles/innovation/innovation news 5.pdf
                            bbc articles/innovation/innovation_news_7.pdf
Extracting text from file:
Extracting text from file:
                            bbc articles/innovation/innovation news 6.pdf
Extracting text from file:
                            bbc articles/innovation/innovation news 2.pdf
Extracting text from file:
                            bbc articles/innovation/innovation news 3.pdf
Extracting text from file:
                            bbc articles/innovation/innovation news 1.pdf
Extracting text from file:
                            bbc articles/innovation/innovation news 0.pdf
Extracting text from file:
                            bbc articles/arts/arts news 0.pdf
Extracting text from file:
                            bbc articles/arts/arts news 1.pdf
Extracting text from file:
                            bbc articles/arts/arts news 3.pdf
Extracting text from file:
                            bbc articles/arts/arts_news_2.pdf
Extracting text from file:
                            bbc articles/arts/arts news 6.pdf
Extracting text from file:
                            bbc articles/arts/arts news 7.pdf
Extracting text from file:
                            bbc articles/arts/arts news 5.pdf
Extracting text from file:
                            bbc articles/arts/arts news 4.pdf
Extracting text from file:
                            bbc articles/arts/arts news 9.pdf
Extracting text from file:
                            bbc articles/arts/arts news 8.pdf
Extracting text from file:
                            bbc articles/technology/technology news 15.pdf
Extracting text from file:
                            bbc articles/technology/technology news 5.pdf
Extracting text from file:
                            bbc articles/technology/technology news 4.pdf
                            bbc articles/technology/technology news 14.pdf
Extracting text from file:
Extracting text from file:
                            bbc_articles/technology/technology_news_6.pdf
Extracting text from file:
                            bbc articles/technology/technology news 16.pdf
Extracting text from file:
                            bbc articles/technology/technology news 7.pdf
Extracting text from file:
                            bbc_articles/technology/technology_news_3.pdf
Extracting text from file:
                            bbc articles/technology/technology news 13.pdf
Extracting text from file:
                            bbc articles/technology/technology news 12.pdf
Extracting text from file:
                            bbc articles/technology/technology news 2.pdf
Extracting text from file:
                            bbc_articles/technology/technology_news_10.pdf
Extracting text from file:
                            bbc articles/technology/technology news_0.pdf
Extracting text from file:
                            bbc articles/technology/technology news 1.pdf
Extracting text from file:
                            bbc articles/technology/technology news 11.pdf
Extracting text from file:
                            bbc_articles/technology/technology_news_9.pdf
Extracting text from file:
                            bbc articles/technology/technology_news_8.pdf
Extracting text from file:
                            bbc articles/business/business news 9.pdf
Extracting text from file:
                            bbc articles/business/business news 8.pdf
Extracting text from file:
                            bbc articles/business/business news 3.pdf
Extracting text from file:
                            bbc articles/business/business news 2.pdf
Extracting text from file:
                            bbc_articles/business/business_news_0.pdf
Extracting text from file:
                            bbc articles/business/business news 1.pdf
Extracting text from file:
                            bbc articles/business/business news 5.pdf
Extracting text from file:
                            bbc articles/business/business news 4.pdf
                            bbc articles/business/business_news_6.pdf
Extracting text from file:
Extracting text from file:
                            bbc articles/business/business news 7.pdf
Extracting text from file:
                            bbc articles/business/business news 13.pdf
```

```
bbc articles/business/business news 12.pdf
Extracting text from file:
Extracting text from file:
                            bbc articles/business/business news 10.pdf
Extracting text from file:
                            bbc_articles/business/business_news_11.pdf
Extracting text from file:
                            bbc articles/business/business news 14.pdf
Extracting text from file:
                            bbc articles/travel/travel_news_0.pdf
Extracting text from file:
                            bbc articles/travel/travel news 1.pdf
Extracting text from file:
                            bbc articles/travel/travel news 3.pdf
Extracting text from file:
                            bbc_articles/travel/travel_news_2.pdf
                            bbc articles/travel/travel news 6.pdf
Extracting text from file:
Extracting text from file:
                            bbc articles/travel/travel news 5.pdf
Extracting text from file:
                            bbc articles/travel/travel news 4.pdf
Loaded 60 documents from bbc articles
```

Building inverted index and calculating TF-IDF scores... Top 10 terms in the inverted index:

Term	Doc Count	Documents
follow sell respons advertis copyright share technic link	60 60 60 60 60 60 60 60	[0, 1, 2, 3, 4] [0, 1, 2, 3, 4]
help	60	[0, 1, 2, 3, 4]
pm	60	[0, 1, 2, 3, 4]

Creating test queries and relevance judgments...

Performing comprehensive evaluation...

Visualizing evaluation results...

```
====== SUMMARY METRICS ======
```

Mean Precision: 0.4622 Mean Recall: 0.6916 Mean F1 Score: 0.5288

Mean Ap: 0.4525

Mean R Precision: 0.5677
Mean Precision@5: 0.5600
Mean Recall@5: 0.1834
Mean Precision@10: 0.5400
Mean Recall@10: 0.3203
Mean Precision@20: 0.5000
Mean Recall@20: 0.5198

Simulating user profile recommendations...

```
Generating recommendations for user01...

Top 5 recommendations for user01:

1. arts/arts_news_3.pdf (Score: 0.1967)

2. arts/arts_news_8.pdf (Score: 0.1577)
```

- 3. arts/arts news 0.pdf (Score: 0.1073)
- 4. travel/travel_news_3.pdf (Score: 0.0677)
- 5. arts/arts_news_4.pdf (Score: 0.0522)

Generating recommendations for user02...

Top 5 recommendations for user02:

- 1. innovation/innovation_news_9.pdf (Score: 0.0797)
- 2. arts/arts_news_8.pdf (Score: 0.0737)
- 3. technology/technology_news_16.pdf (Score: 0.0619)
- 4. innovation/innovation_news_0.pdf (Score: 0.0545)
- 5. technology/technology_news_6.pdf (Score: 0.0545)

Evaluating user recommendations...

Evaluation for user01:

Precision (category match): 0.7000 Average relevance score: 0.5100

Evaluation for user02:

Precision (category match): 0.9000 Average relevance score: 0.5600