Università degli Studi di Padova

Department of Information Engineering
University of Padova

A Progressive Visual Analytics Tool for Incremental Experimental Evaluation

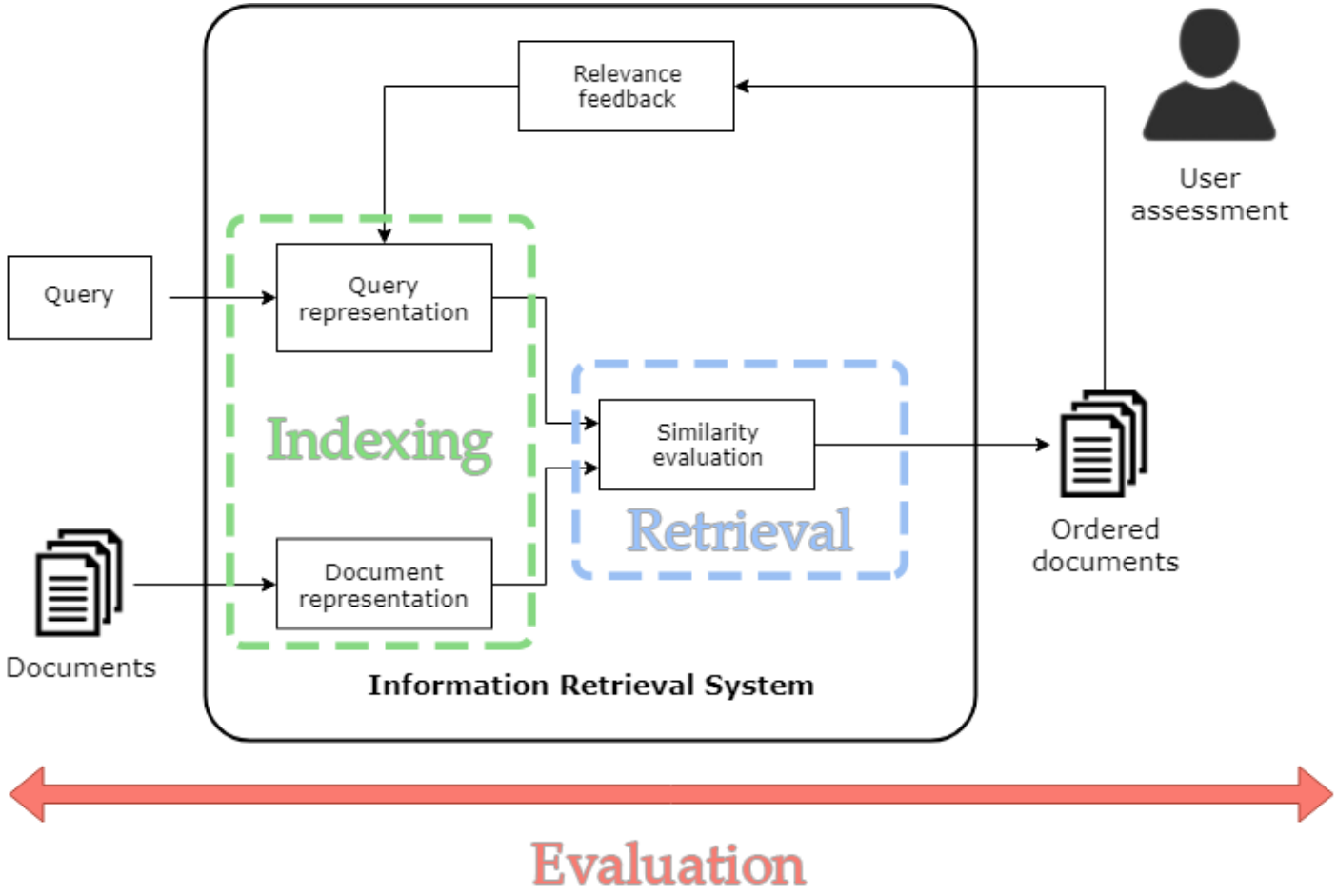
Fabio Giachelle, Gianmaria Silvello

16 September 2019

Contents

Introduction

In the field of Information Retrieval (IR), evaluating Information Retrieval Systems (IRS), it is of fundamental importance. The evaluation process allows to assess and compare the performances of different IR systems.



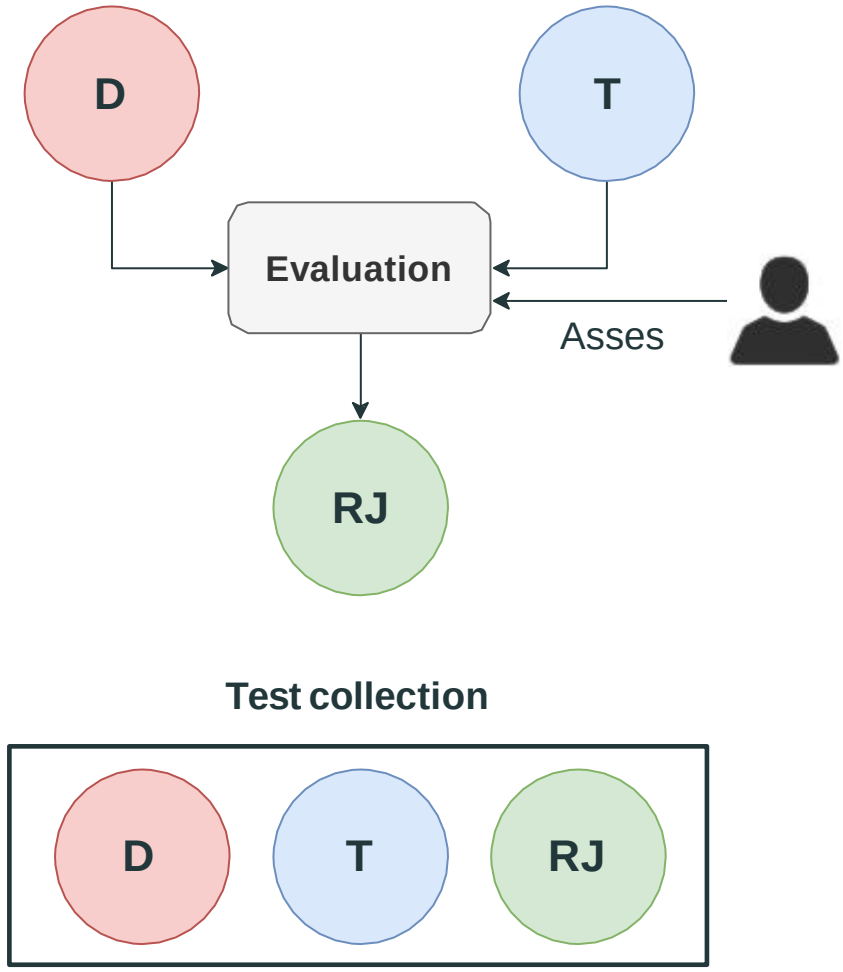


Figure 2: Test collection $C = \{ D, T, RJ \}$ and Evaluation.

Figure 3: Visual Analytics workflow.

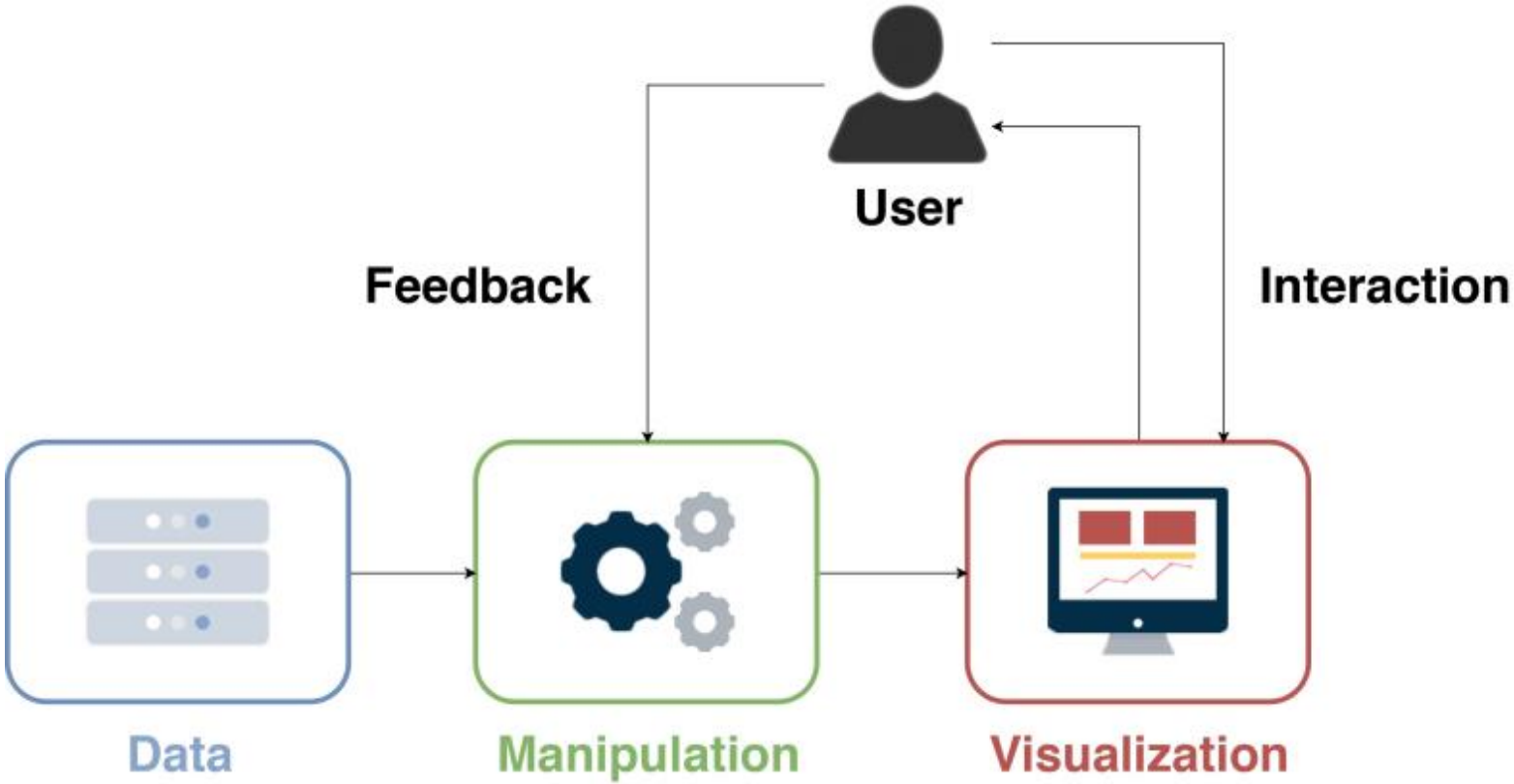


Figure 4: Visual Analytics workflow and Human-in-the-loop model.

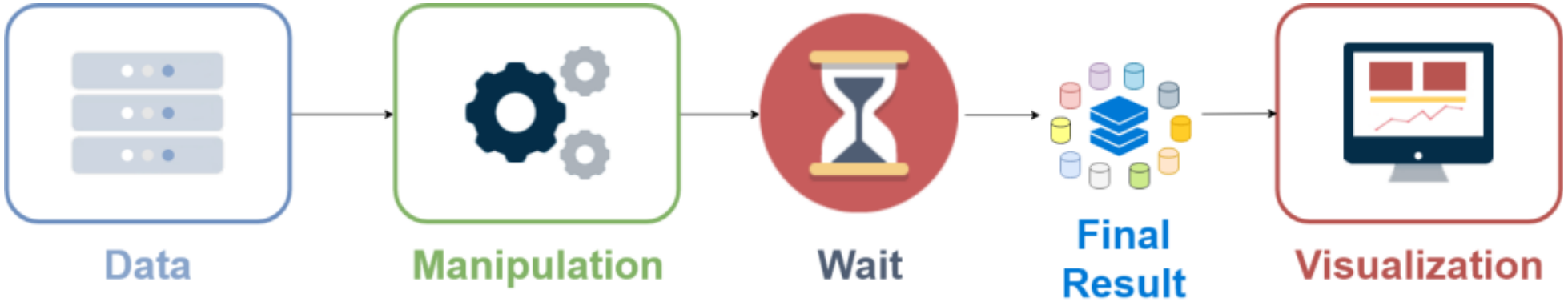


Figure 5: Visual Analytics workflow and time issue.

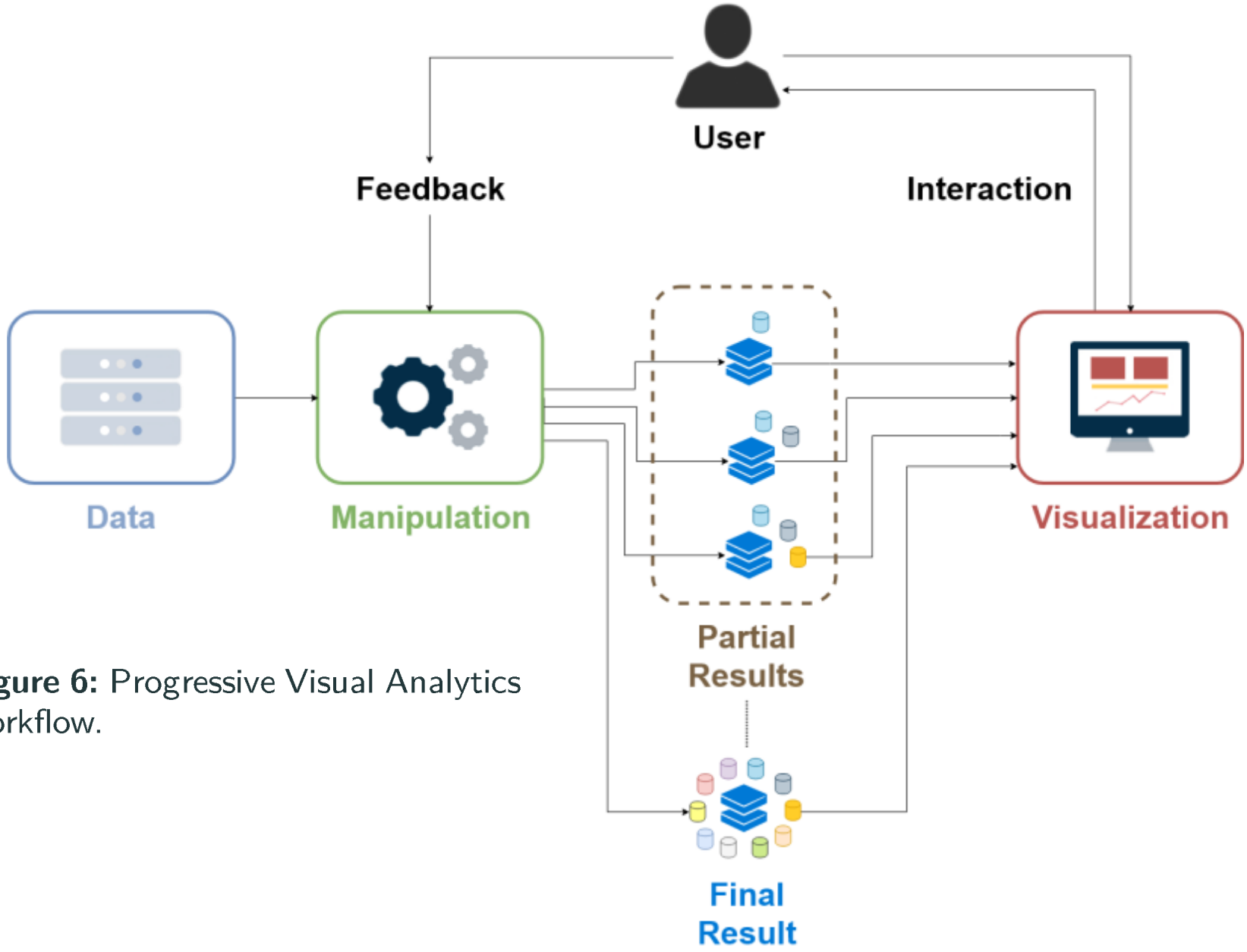


Figure 6: Progressive Visual Analytics workflow.

The Problem

The major **issues** regarding a generic IR process are:

1. Test collections can be made of millions of documents.
2. Indexing Very Large Collections (VLC) requires a lot of time.
3. Many types of experiments require to re-index the same test collection several times.

Work Objectives

Development of a progressive visual analytics tool, for the evaluation of IR systems, capable of incrementally indexing a given test collection.

Advantages:

- Progressive evaluation of IR systems.
- Visual exploration of evaluation data.
- Automatic indexing, retrieval and evaluation.

Aviator

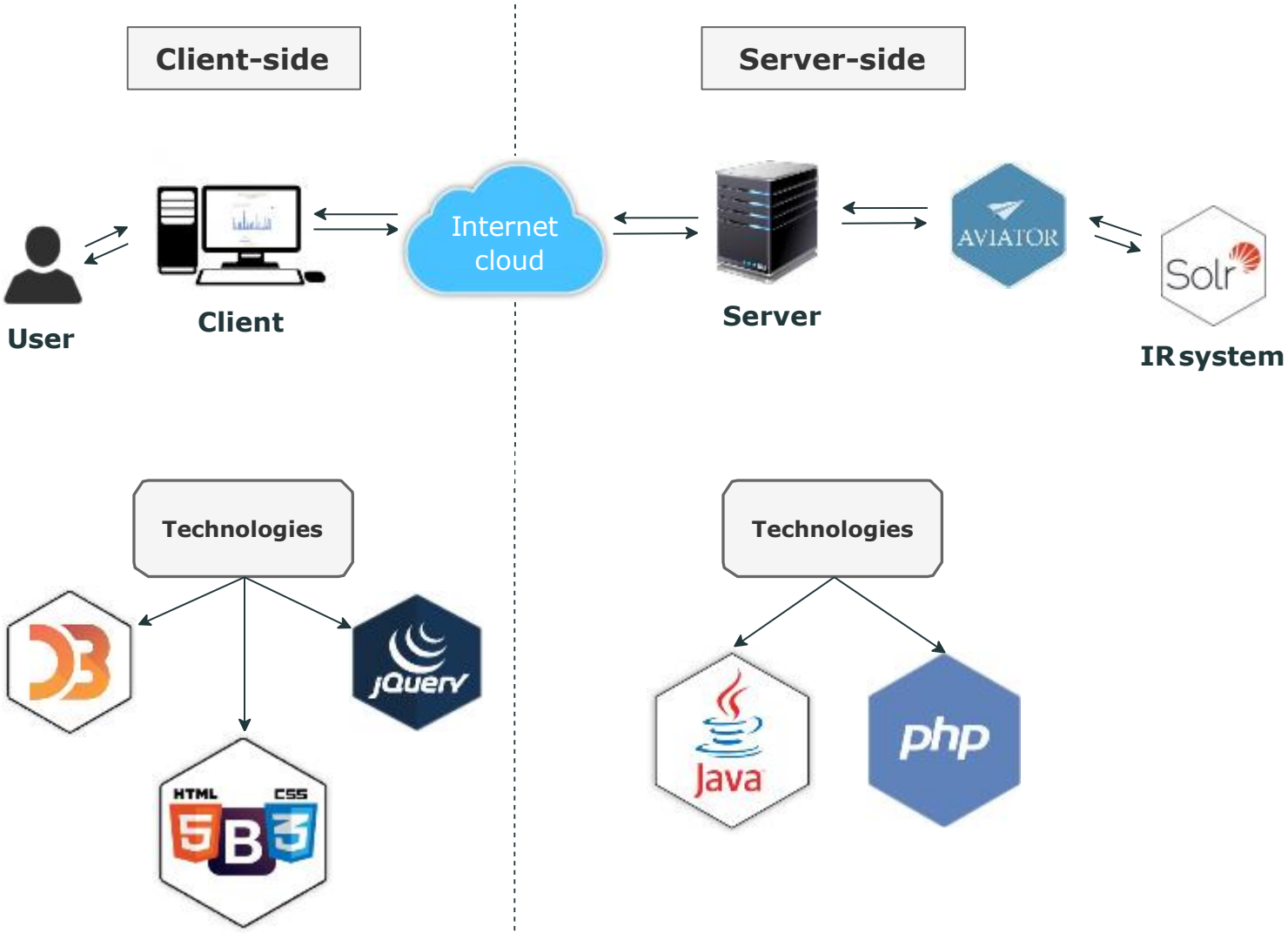


Figure 7: High level architecture: components and interactions.

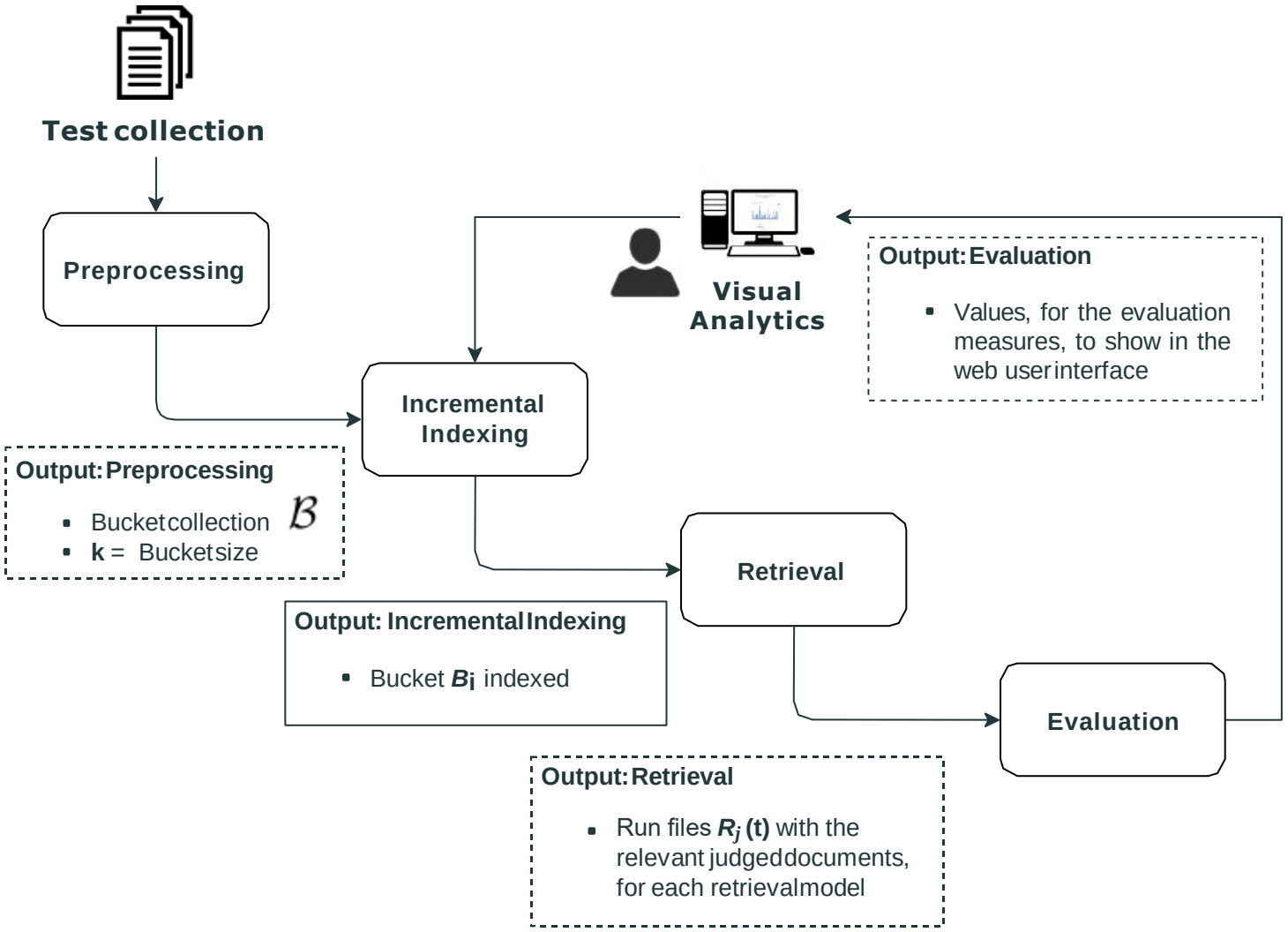


Figure 8: Aviator process: step by step.

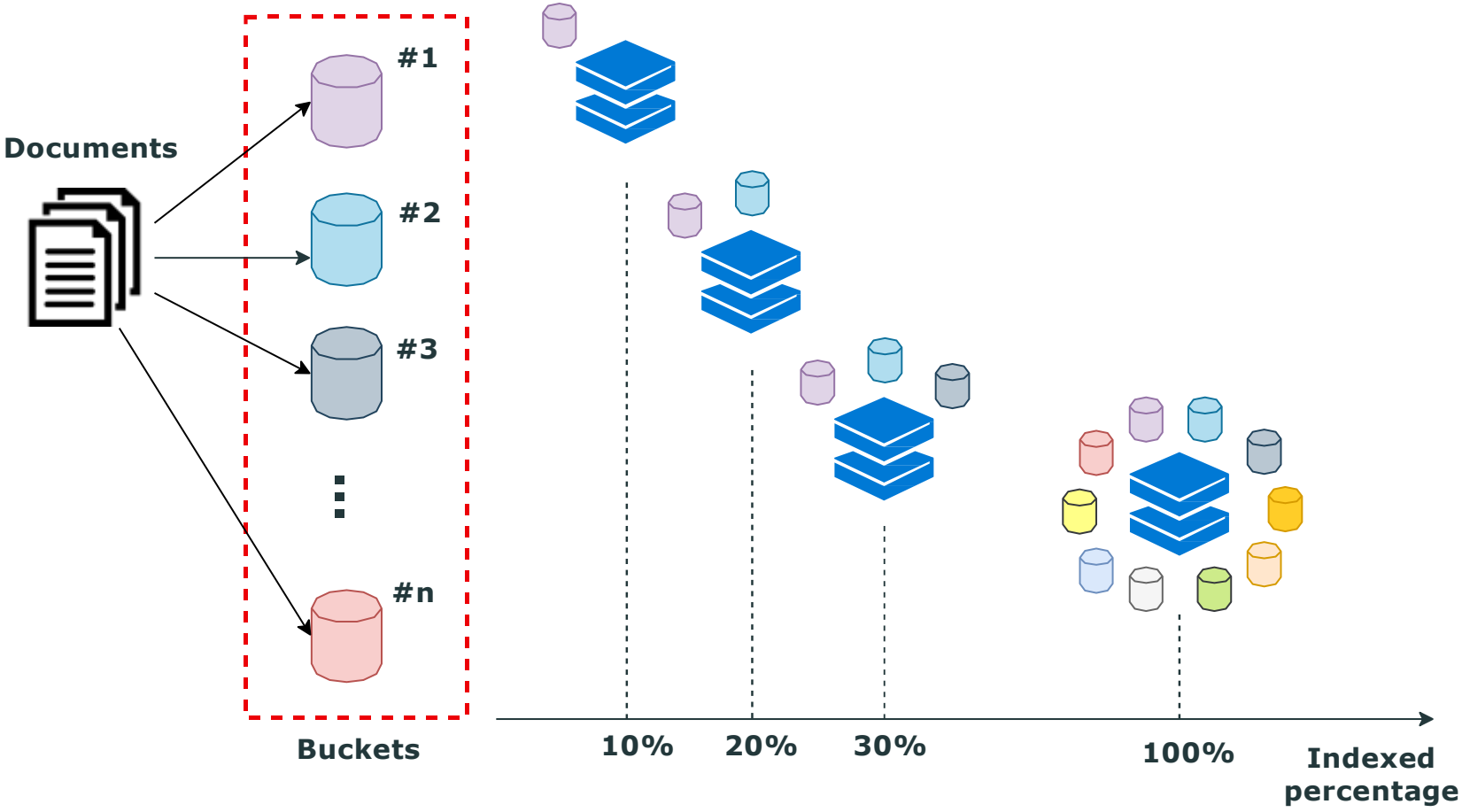


Figure 9: Bucket partitioning and Incremental Indexing.

Incremental Indexing

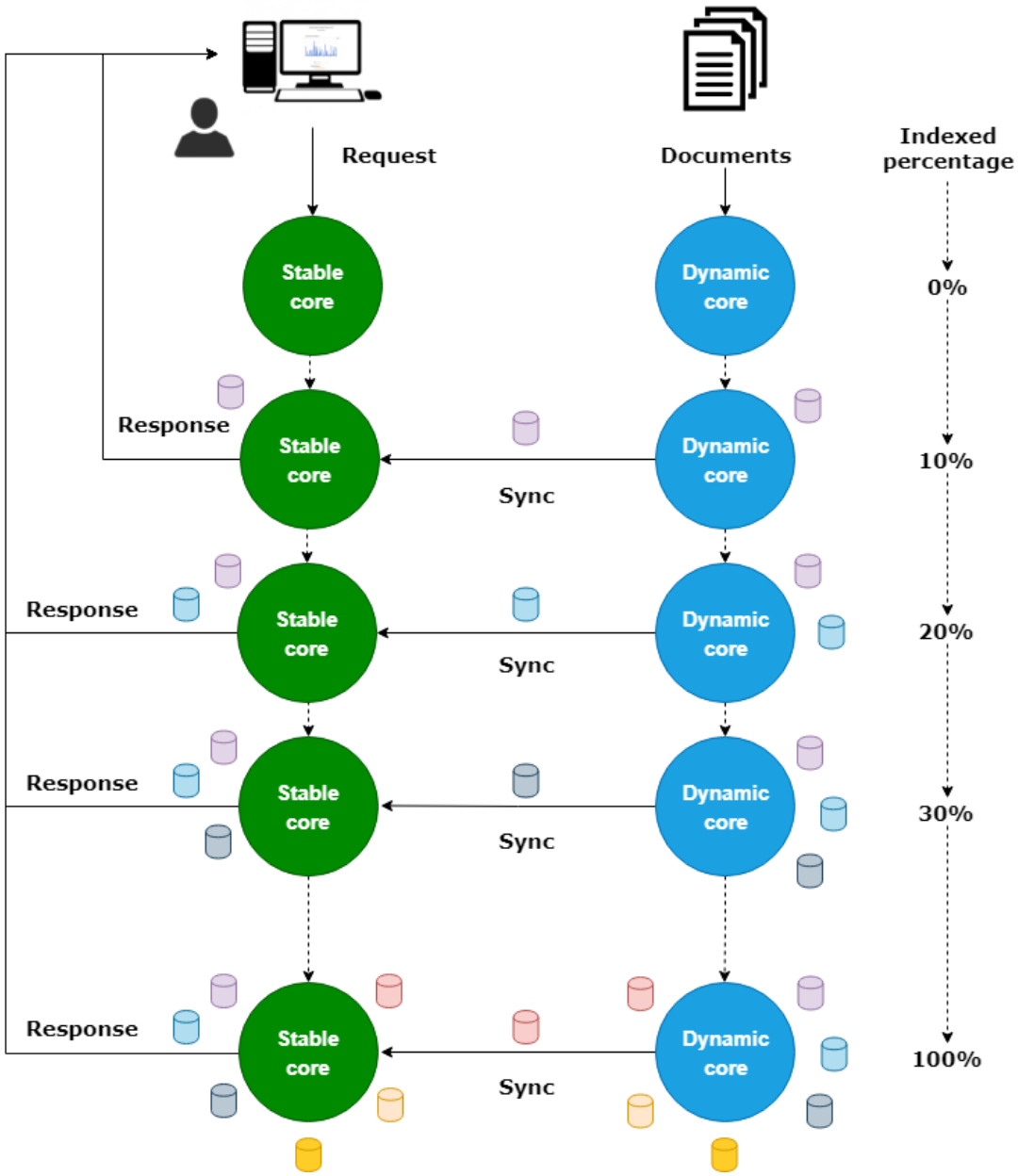


Figure 11: Aviator web user interface: homepage.

Visual Analytics Interface: Measure selection

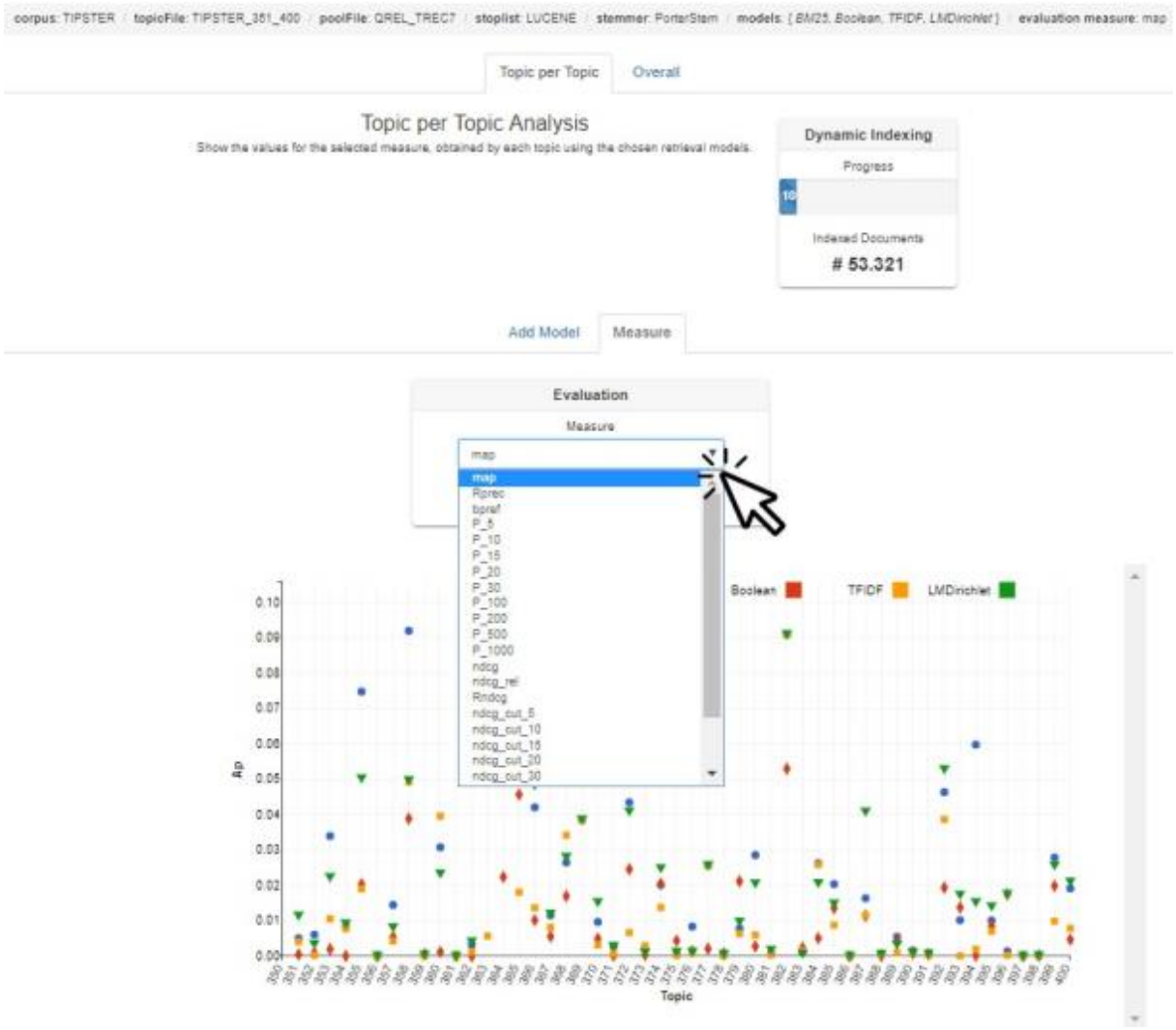


Figure 12: Aviator visual analytics UI: topic per topic, measure selection, progress 10%.

Visual Analytics Interface: Model selection

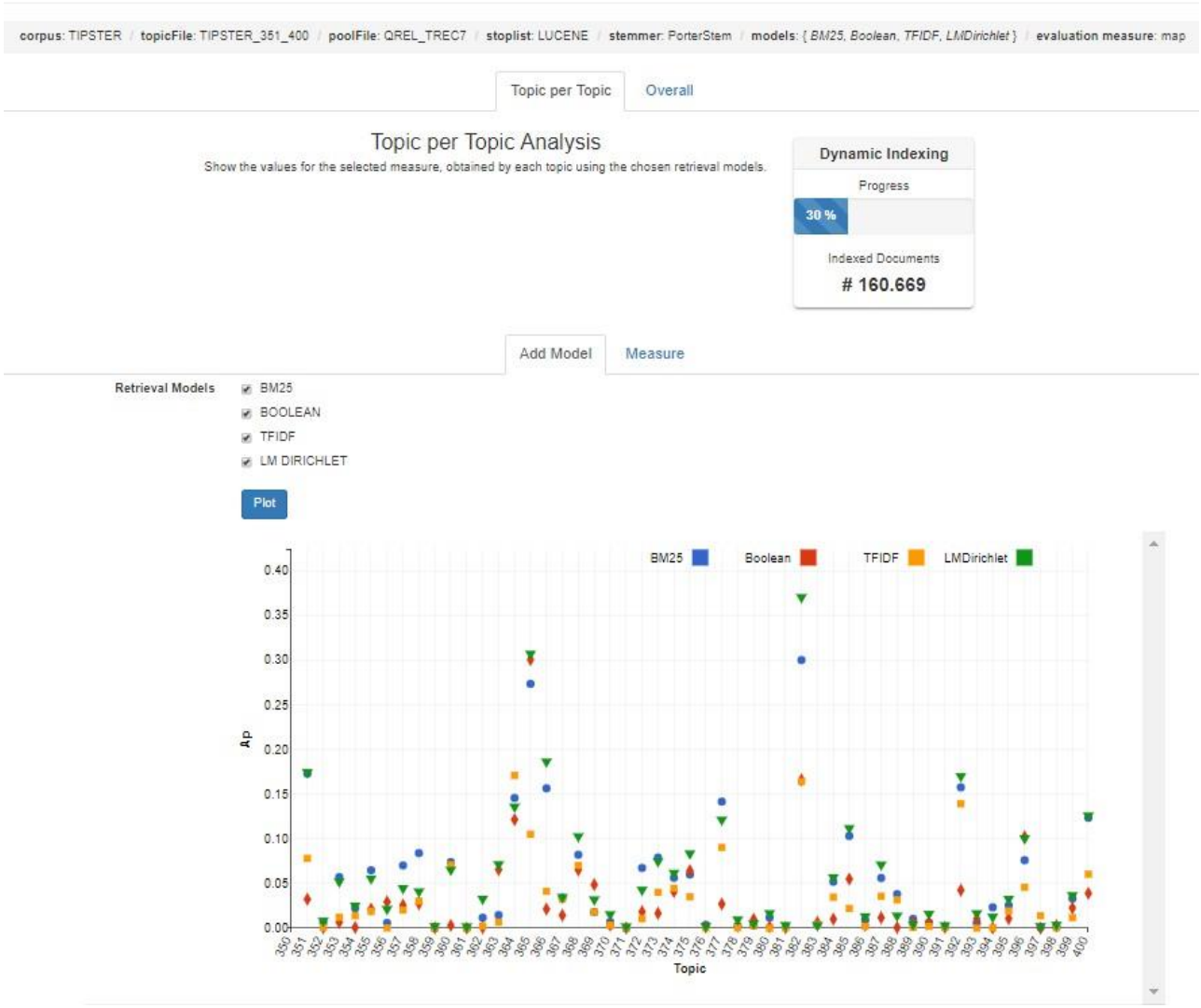


Figure 13: Aviator visual analytics UI: topic per topic, IR model selection, progress 30%.

Visual Analytics Interface: Pan and zoom

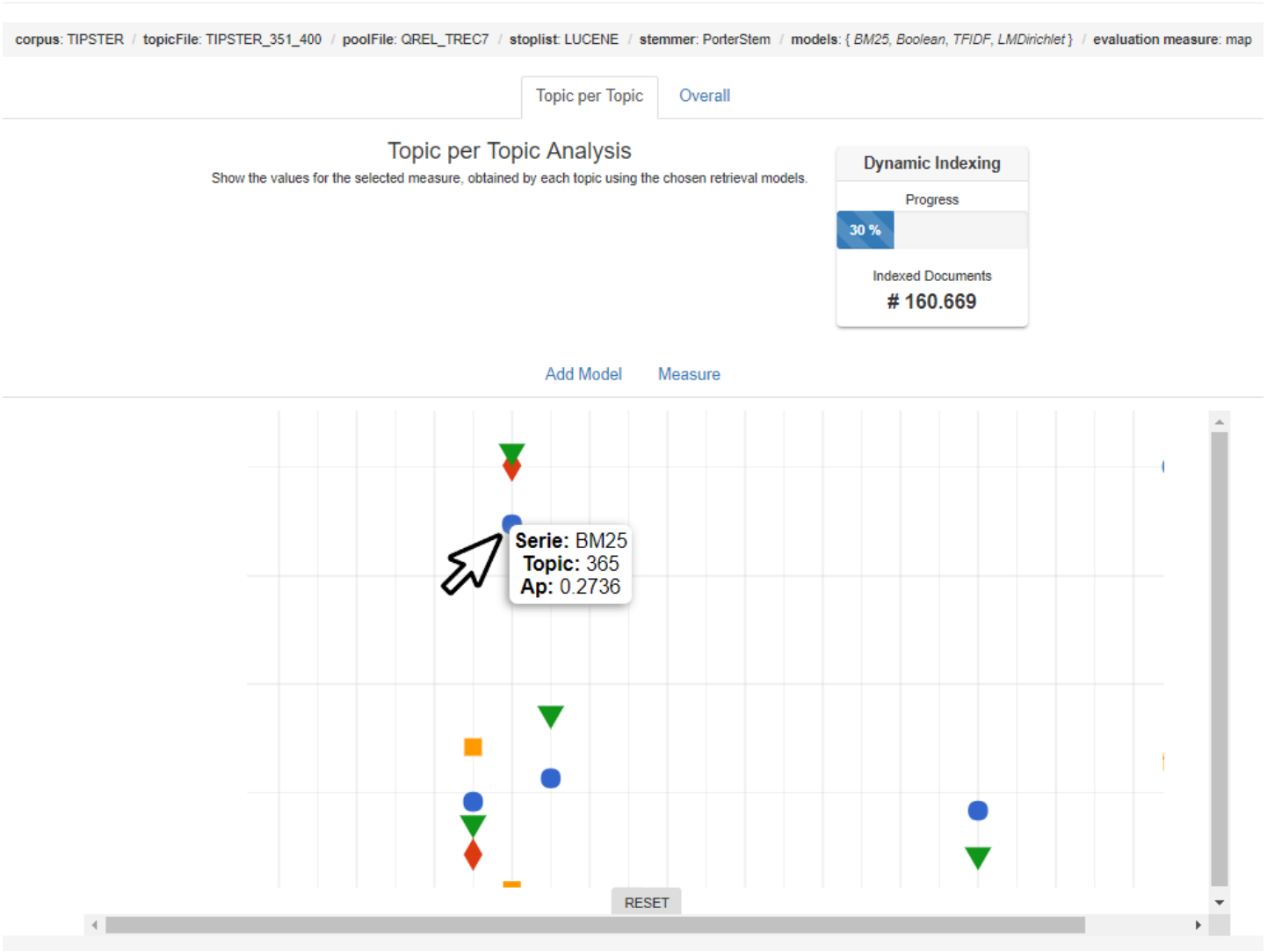


Figure 14: Aviator visual analytics UI: topic per topic, zoom and inspect, progress 30%.

Visual Analytics Interface: Upgrade

Figure 15: Aviator visual analytics UI: topic per topic, progress 90%.

Topic per topic analysis: Progress 100%

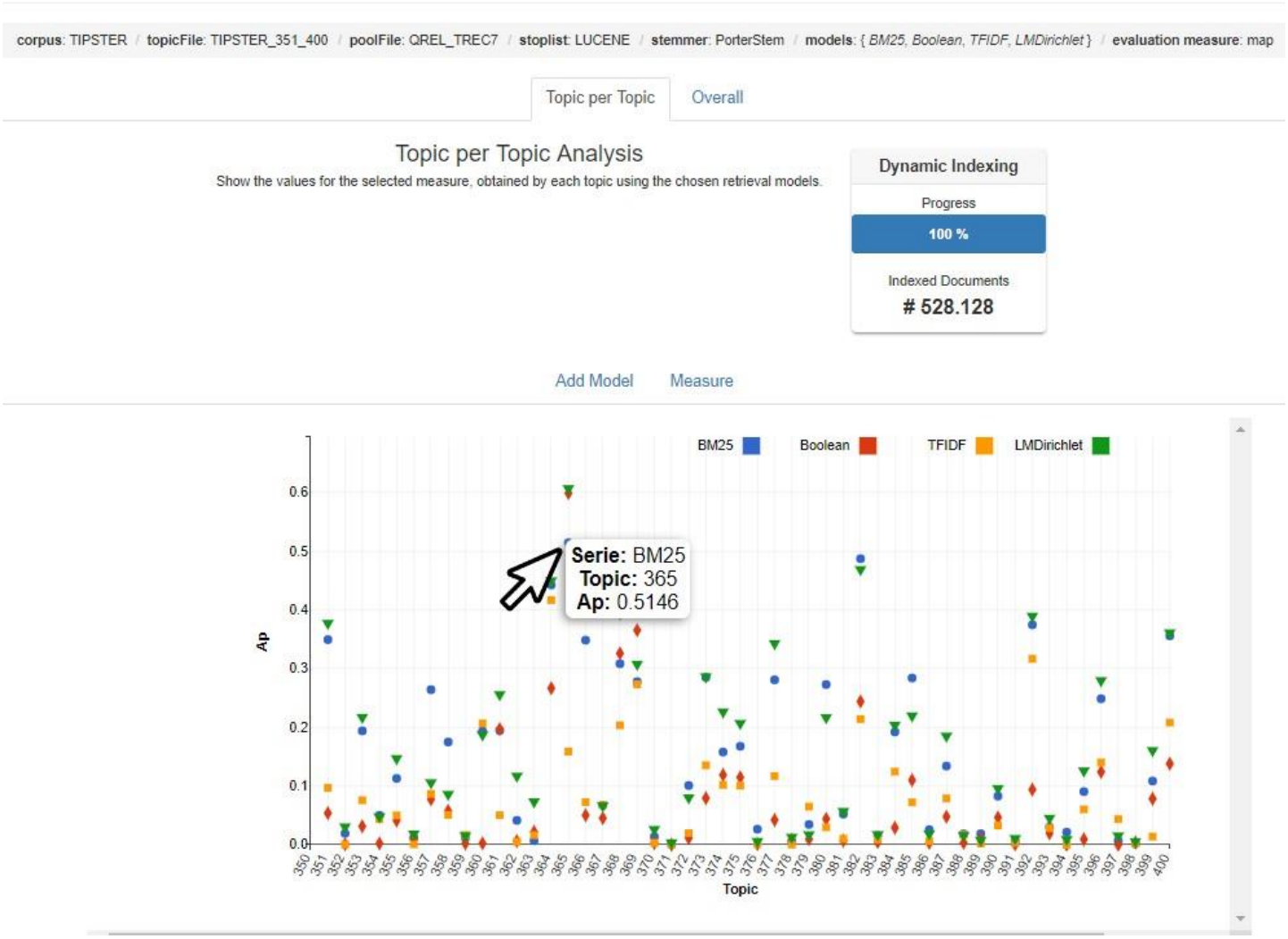


Figure 16: Aviator visual analytics UI: topic per topic, progress 100%.

Overall analysis: Progress 100%

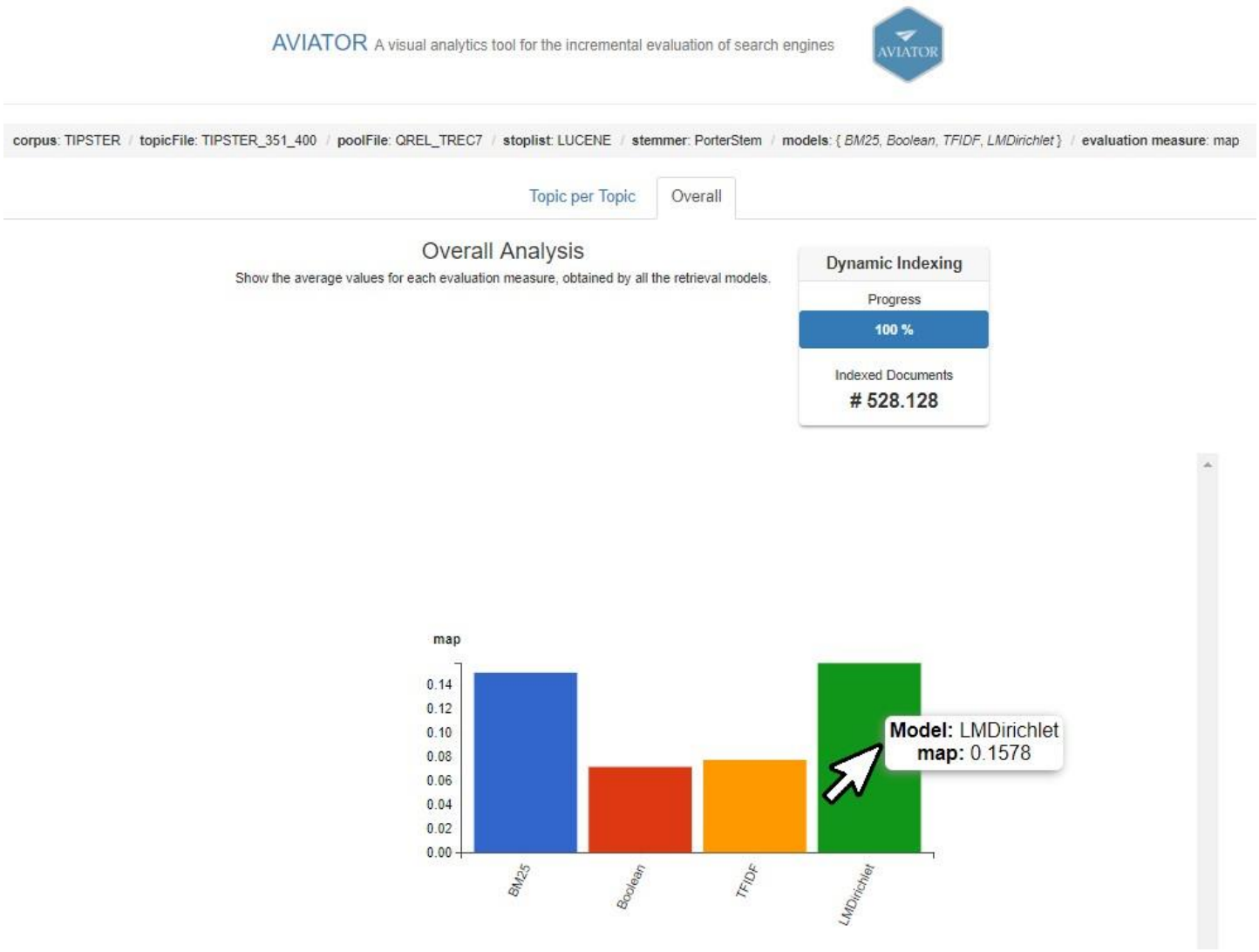


Figure 17: Aviator visual analytics UI: overall, progress 100%.

Experimental Results

Description of the systems

System	Description
S_1	(Indri, Porter stemmer, BM25)
S_2	(Indri, Porter stemmer, TFIDF)
S_3	(Indri, No stemmer, BM25)
S_4	(Indri, No stemmer, TFIDF)
S_5	(No stoplist, Porter stemmer, BM25)
S_6	(No stoplist, Porter stemmer, TFIDF)
S_7	(No stoplist, No stemmer, BM25)
S_8	(No stoplist, No stemmer, TFIDF)

Table 1: Description of each system S_j in terms of stoplist, stemmer and retrieval model.

MAP Results

System \ Measure	MAP_1	MAP_2	MAP_3	MAP_4	MAP_5	MAP_6	MAP_7	MAP_8	MAP_9	MAP_{10}
S_1	0.0207 (-87%)	0.0394 (-75%)	0.0532 (-66%)	0.0670 (-58%)	0.0810 (-49%)	0.0959 (-39%)	0.1060 (-33%)	0.1212 (-24%)	0.1394 (-12%)	0.1585 (0%)
S_2	0.0124 (-86%)	0.0232 (-73%)	0.0314 (-64%)	0.0376 (-57%)	0.0465 (-46%)	0.0561 (-35%)	0.0595 (-31%)	0.0664 (-24%)	0.0766 (-12%)	0.0868 (0%)
S_3	0.0216 (-86%)	0.0369 (-76%)	0.0511 (-67%)	0.0660 (-58%)	0.0805 (-48%)	0.0938 (-40%)	0.1034 (-33%)	0.1174 (-24%)	0.1345 (-13%)	0.1553 (0%)
S_4	0.0130 (-85%)	0.0209 (-76%)	0.0291 (-66%)	0.0356 (-59%)	0.0452 (-48%)	0.0532 (-38%)	0.0575 (-33%)	0.0642 (-26%)	0.0743 (-14%)	0.0863 (0%)
S_5	0.0215 (-87%)	0.0402 (-75%)	0.0550 (-66%)	0.0702 (-57%)	0.0847 (-48%)	0.0997 (-39%)	0.1093 (-33%)	0.1252 (-23%)	0.1436 (-12%)	0.1631 (0%)
S_6	0.0117 (-86%)	0.0213 (-74%)	0.0296 (-64%)	0.0360 (-56%)	0.0439 (-46%)	0.0530 (-35%)	0.0559 (-31%)	0.0626 (-23%)	0.0721 (-11%)	0.0811 (0%)
S_7	0.0218 (-86%)	0.0372 (-76%)	0.0519 (-66%)	0.0664 (-57%)	0.0807 (-47%)	0.0940 (-39%)	0.1024 (-33%)	0.1167 (-24%)	0.1331 (-13%)	0.1533 (0%)
S_8	0.0124 (-84%)	0.0192 (-76%)	0.0273 (-65%)	0.0335 (-58%)	0.0417 (-47%)	0.0494 (-38%)	0.0531 (-33%)	0.0592 (-25%)	0.0686 (-13%)	0.0791 (0%)

Figure 18: MAP for each system S_j and bucket B_i .

Conclusions

Achieved goals

- Estimation of final results for each evaluation measure.
- Dynamic control of IR processes.
- Visual exploration of evaluation data.
- Automatic indexing, retrieval and evaluation.

Thank you for your kind attention.

> *Remember to look up at the stars and not down at your feet. Be curious. And however difficult life may seem, there is always something you can do and succeed at. It matters that you don't just give up.*
>
> *Stephen Hawking*