Diversity and Fairness From a Ranking Evaluation Perspective

Damiano Spina

@damiano10 / damiano.spina@rmit.edu.au

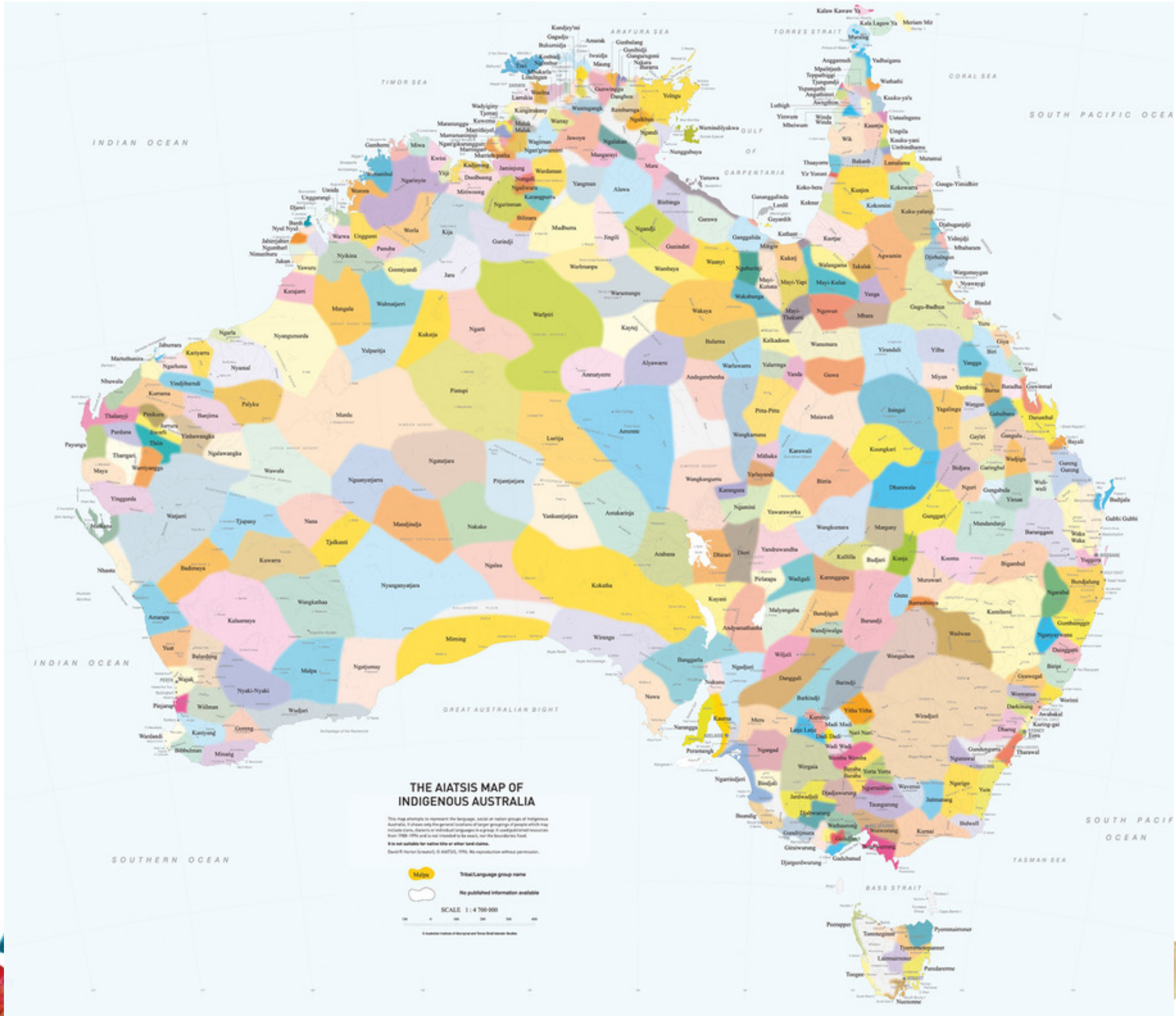
30 Nov 2021

What's next...

ADM +S ARC Centre of Excellence for Automated Decision-Making and Society

CiDDA

RMIT UNIVERSITY



Source: AIATSIS Map of Indigenous Australia.
https://aiatsis.gov.au/explore/map-indigenous-australia

Source: AIATSIS Map of Indigenous Australia. https://aiatsis.gov.au/explore/map-indigenous-australia

Acknowledgment of Country

I would like to acknowledge the people of the Woi wurrung and Boon wurrung language groups of the eastern Kulin Nation on whose unceded lands I live, teach, work, and learn.

I respectfully acknowledge Ancestors and Elders past and present.

I also acknowledge the Traditional Custodians and their Ancestors of the lands and waters across Australia where we conduct our business.

About Me

About Me

PhD in Computer Science (UNED, 2014)
Supervisors: Julio Gonzalo and Enrique Amigó

About Me

PhD in Computer Science
(UNED, 2014)
Supervisors: Julio Gonzalo
and Enrique Amigó

Postdoc at RMIT University
(2015-17, 2017-2019)
Mentors: Lawrence
Cavedon, Mark Sanderson,
Falk Scholer

About Me

PhD in Computer Science (UNED, 2014)
Supervisors: Julio Gonzalo and Enrique Amigó

Postdoc at RMIT University (2015-17, 2017-2019)
Mentors: Lawrence Cavedon, Mark Sanderson, Falk Scholer

and W. Bruce Croft

About Me

PhD in Computer Science
(UNED, 2014)
Supervisors: Julio Gonzalo
and Enrique Amigó

Postdoc at RMIT University
(2015-17, 2017-2019)
Mentors: Lawrence
Cavedon, Mark Sanderson,
Falk Scholer

and W. Bruce Croft

Lecturer (2019-) and
DECRA Fellow (2020-
2023) at RMIT

About RMIT CIDDA, CoE ADM+S & FactLab

About RMIT CIDDA, CoE ADM+S & FactLab

RMIT Research Centre for Information Discovery and Data Analytics

http://rmit.edu.au/cidda

About RMIT CIDDA, CoE ADM+S & FactLab

RMIT Research Centre for Information Discovery and Data Analytics

http://rmit.edu.au/cidda

ARC Centre of Excellence for Automated Decision-Making and Society

Bringing together universities, industry, government and the community to support the development of responsible, ethical and inclusive automated decision-making.

https://www.admscentre.org.au/

About RMIT CIDDA, CoE ADM+S & FactLab

RMIT Research Centre for Information Discovery and Data Analytics

http://rmit.edu.au/cidda

Bringing together universities, industry, government and the community to support the development of responsible, ethical and inclusive automated decision-making.

https://www.admscentre.org.au/

Research Hub (RMIT ABC Fact Check + CIDDA + ADM+S)

https://www.rmit.edu.au/about/schools-colleges/media-and-communication/industry/factlab

Purpose of Today/Tonight

Diversity

and

Fairness

from a

Ranking Evaluation

perspective

Diversity

and

Fairness

from a

Ranking (Offline) Evaluation

perspective

Grounding Question

Source: Vicki Smith / Getty Images

Go to **www.menti.com** and use the code **3168 7558**

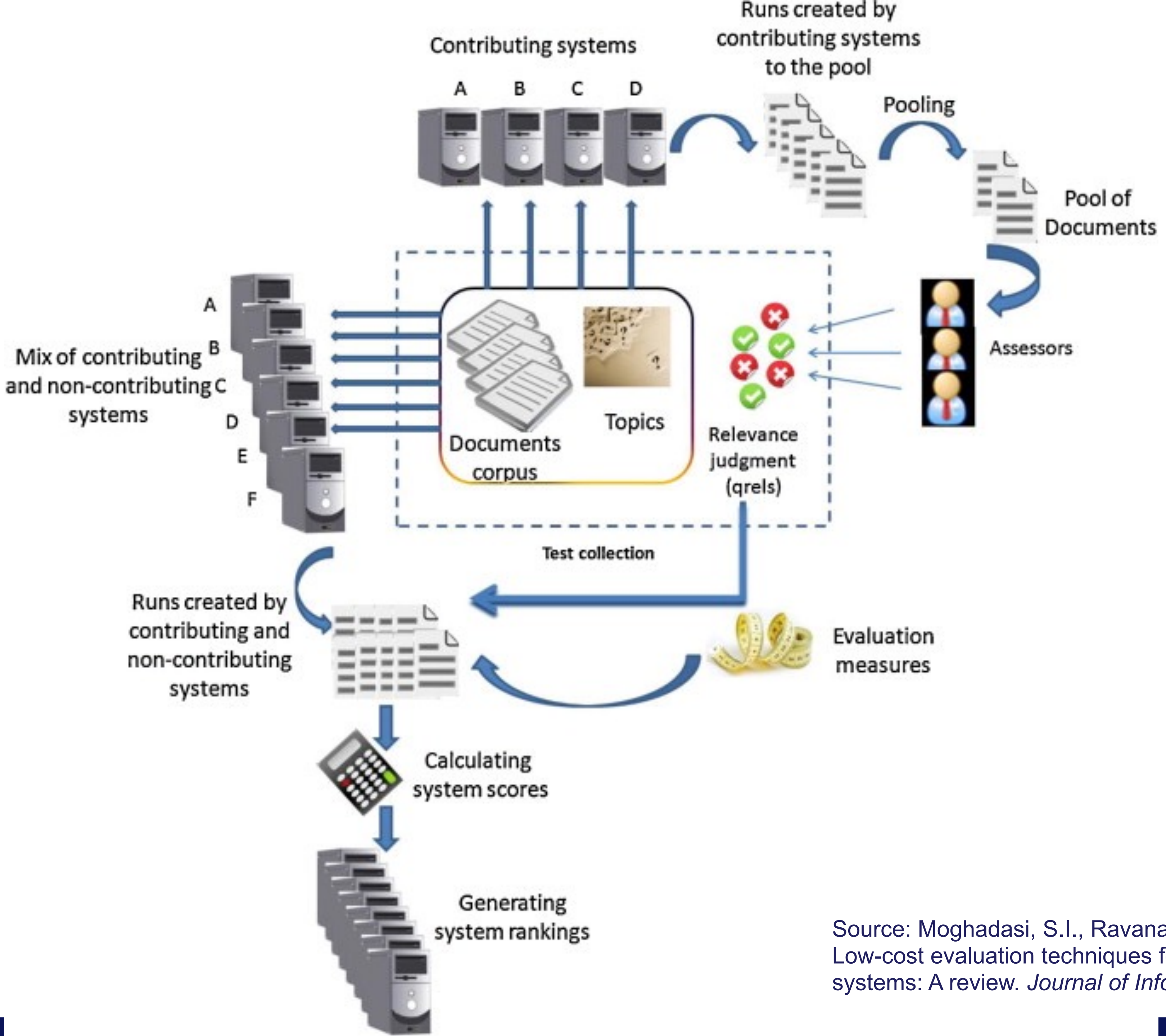
https://www.menti.com/eh8eb77k8w

List elements used to characterize IR evaluation

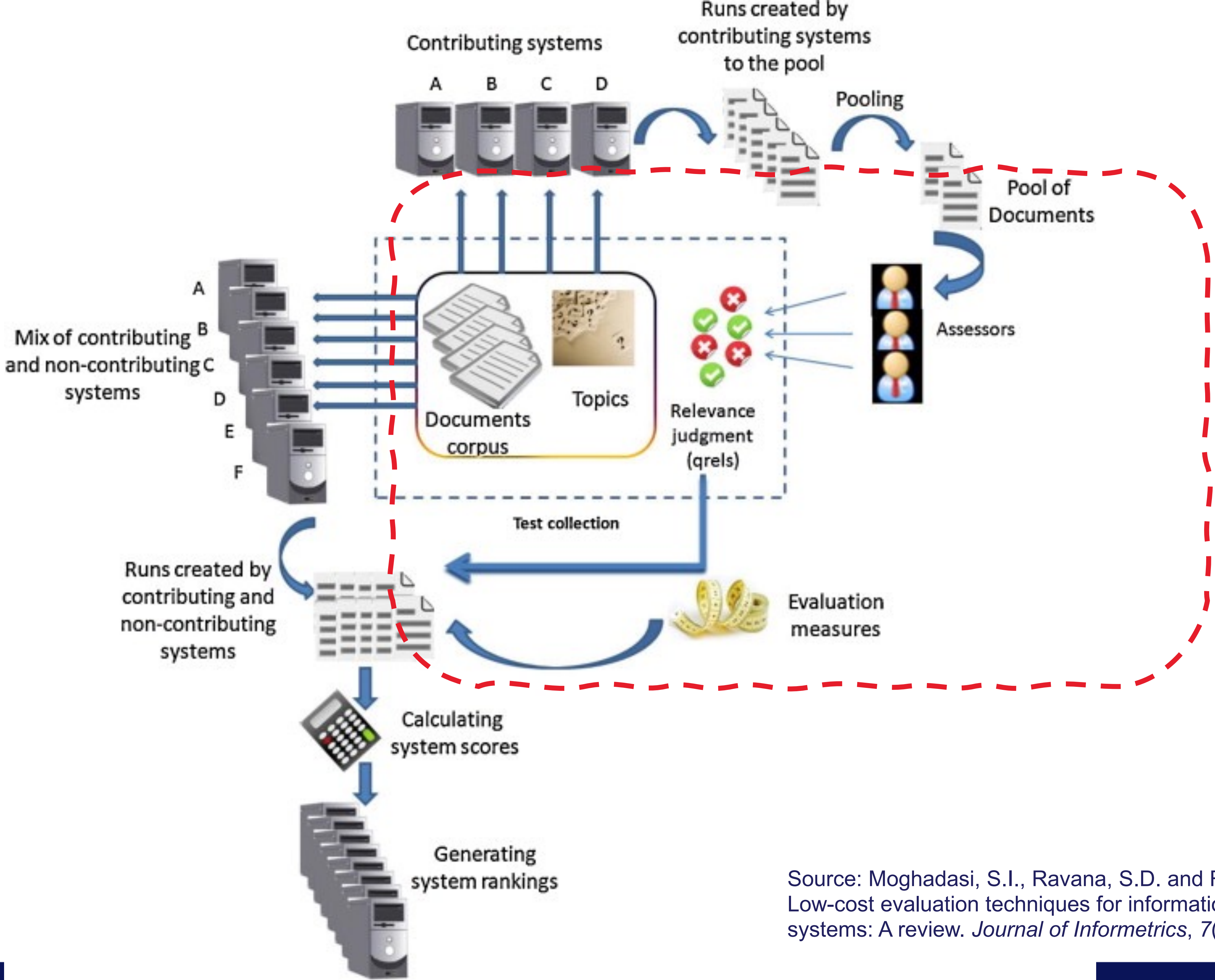
Mentimeter

ranking based evaluation
sentences
expensive
effectiveness
click through rate
efficient
mmr
efficiënt
cursor location
relevance
f-score
query
ndcg
cheese
document score
bias
bm25
posting
no ground truth
recall
sounds good
metrics
f1-score
mrr
user needs
formulas
k-precision
words
precision
ranking
usability
f score
dgc
changing
position bias
quality
relevancy
dcg
efficiency
satisfaction
features
uniqueness
metric
map
f-measure
reciprocal rank
set based evaluation
relevance judgments

60



Source: Moghadasi, S.I., Ravana, S.D. and Raman, S.N., 2013. Low-cost evaluation techniques for information retrieval systems: A review. *Journal of Informetrics*, *7*(2), pp.301-312.



Source: Moghadasi, S.I., Ravana, S.D. and Raman, S.N., 2013. Low-cost evaluation techniques for information retrieval systems: A review. *Journal of Informetrics*, *7*(2), pp.301-312.

What is a topic? How are topics related to queries and intents?

What is a topic? How are topics related to queries and intents?

What does it make a document relevant to a topic?

What is a topic? How are topics related to queries and intents?

What does it make a document relevant to a topic?

Which evaluation measure should we use?

What is a topic? How are topics related to queries and intents?

What does it make a document relevant to a topic?

Which evaluation measure should we use?

What does the evaluation score tell us about the quality of systems?

Diversity

and

Fairness

from a

Ranking Evaluation

perspective

Topics, Intents, Queries, and Aspects

User's
information need

Topics, Intents, Queries, and Aspects

User's information need → Topic

Topics, Intents, Queries, and Aspects

$\underbrace{\text{User's information need} \longrightarrow \text{Topic}}_{\text{Intent}}$

Topics, Intents, Queries, and Aspects

$$\underbrace{\text{User's information need} \rightarrow \text{Topic}}_{\text{Intent}} \xrightarrow{1..m} \textbf{Query}$$

Topics, Intents, Queries, and Aspects

User's information need → Topic **1..n** 1..m → **Query**

Intent

Topics, Intents, Queries, and Aspects

Source: https://plg.uwaterloo.ca/~trecweb/2012.html

```
<topic number="6" type="ambiguous">
  <query>kcs</query>
  <description>Find information on the Kansas City Southern railroad.
  </description>
  <subtopic number="1" type="nav">
    Find the homepage for the Kansas City Southern railroad.
  </subtopic>
  <subtopic number="2" type="inf">
    I'm looking for a job with the Kansas City Southern railroad.
  </subtopic>
  <subtopic number="3" type="nav">
    Find the homepage for Kanawha County Schools in West Virginia.
  </subtopic>
  <subtopic number="4" type="nav">
    Find the homepage for the Knox County School system in Tennessee.
  </subtopic>
  <subtopic number="5" type="inf">
    Find information on KCS Energy, Inc., and their merger with
    Petrohawk Energy Corporation.
  </subtopic>
</topic>

<topic number="16" type="faceted">
  <query>arizona game and fish</query>
  <description>I'm looking for information about fishing and hunting
  in Arizona.
  </description>
  <subtopic number="1" type="nav">
    Take me to the Arizona Game and Fish Department homepage.
  </subtopic>
  <subtopic number="2" type="inf">
    What are the regulations for hunting and fishing in Arizona?
  </subtopic>
  <subtopic number="3" type="nav">
    I'm looking for the Arizona Fishing Report site.
  </subtopic>
  <subtopic number="4" type="inf">
    I'd like to find guides and outfitters for hunting trips in Arizona.
  </subtopic>
</topic>
```

Initial topic release will include only the *query* field.

As shown in these examples, topics are categorized as either "ambiguous" or "faceted". Ambiguous queries are those that have multiple distinct interpretations. We assume that a user interested in one interpretation would not be interested in the others. On the other hand, facets reflect underspecified queries, with different aspects covered by the subtopics. We assume that a user interested in one aspect may still be interested in others.

```
151 1 clueweb09-en0005-98-00099 1
151 2 clueweb09-en0005-98-00099 0
151 3 clueweb09-en0005-98-00099 0
151 4 clueweb09-en0005-98-00099 0
151 5 clueweb09-en0005-98-00099 0
151 1 clueweb09-en0005-98-00107 2
151 2 clueweb09-en0005-98-00107 2
151 3 clueweb09-en0005-98-00107 2
151 4 clueweb09-en0005-98-00107 2
151 5 clueweb09-en0005-98-00107 4
151 1 clueweb09-en0005-98-00108 1
151 2 clueweb09-en0005-98-00108 0
151 3 clueweb09-en0005-98-00108 0
151 4 clueweb09-en0005-98-00108 0
151 5 clueweb09-en0005-98-00108 0
151 1 clueweb09-en0005-98-00110 1
151 2 clueweb09-en0005-98-00110 0
```

Qrels

```
151 1 clueweb09-en0005-98-00099 1
151 2 clueweb09-en0005-98-00099 0
151 3 clueweb09-en0005-98-00099 0
151 4 clueweb09-en0005-98-00099 0
151 5 clueweb09-en0005-98-00099 0
151 1 clueweb09-en0005-98-00107 2
151 2 clueweb09-en0005-98-00107 2
151 3 clueweb09-en0005-98-00107 2
151 4 clueweb09-en0005-98-00107 2
151 5 clueweb09-en0005-98-00107 4
151 1 clueweb09-en0005-98-00108 1
151 2 clueweb09-en0005-98-00108 0
151 3 clueweb09-en0005-98-00108 0
151 4 clueweb09-en0005-98-00108 0
151 5 clueweb09-en0005-98-00108 0
151 1 clueweb09-en0005-98-00110 1
151 2 clueweb09-en0005-98-00110 0
```

Qrels

Topic

```
151 1 clueweb09-en0005-98-00099 1
151 2 clueweb09-en0005-98-00099 0
151 3 clueweb09-en0005-98-00099 0
151 4 clueweb09-en0005-98-00099 0
151 5 clueweb09-en0005-98-00099 0
151 1 clueweb09-en0005-98-00107 2
151 2 clueweb09-en0005-98-00107 2
151 3 clueweb09-en0005-98-00107 2
151 4 clueweb09-en0005-98-00107 2
151 5 clueweb09-en0005-98-00107 4
151 1 clueweb09-en0005-98-00108 1
151 2 clueweb09-en0005-98-00108 0
151 3 clueweb09-en0005-98-00108 0
151 4 clueweb09-en0005-98-00108 0
151 5 clueweb09-en0005-98-00108 0
151 1 clueweb09-en0005-98-00110 1
151 2 clueweb09-en0005-98-00110 0
```

Qrels

Topic

Document Id

```
151 1 clueweb09-en0005-98-00099 1
151 2 clueweb09-en0005-98-00099 0
151 3 clueweb09-en0005-98-00099 0
151 4 clueweb09-en0005-98-00099 0
151 5 clueweb09-en0005-98-00099 0
151 1 clueweb09-en0005-98-00107 2
151 2 clueweb09-en0005-98-00107 2
151 3 clueweb09-en0005-98-00107 2
151 4 clueweb09-en0005-98-00107 2
151 5 clueweb09-en0005-98-00107 4
151 1 clueweb09-en0005-98-00108 1
151 2 clueweb09-en0005-98-00108 0
151 3 clueweb09-en0005-98-00108 0
151 4 clueweb09-en0005-98-00108 0
151 5 clueweb09-en0005-98-00108 0
151 1 clueweb09-en0005-98-00110 1
151 2 clueweb09-en0005-98-00110 0
```

Qrels

Topic

Document Id

Relevance Judgment

```
151 1 clueweb09-en0005-98-00099 1
151 2 clueweb09-en0005-98-00099 0
151 3 clueweb09-en0005-98-00099 0
151 4 clueweb09-en0005-98-00099 0
151 5 clueweb09-en0005-98-00099 0
151 1 clueweb09-en0005-98-00107 2
151 2 clueweb09-en0005-98-00107 2
151 3 clueweb09-en0005-98-00107 2
151 4 clueweb09-en0005-98-00107 2
151 5 clueweb09-en0005-98-00107 4
151 1 clueweb09-en0005-98-00108 1
151 2 clueweb09-en0005-98-00108 0
151 3 clueweb09-en0005-98-00108 0
151 4 clueweb09-en0005-98-00108 0
151 5 clueweb09-en0005-98-00108 0
151 1 clueweb09-en0005-98-00110 1
151 2 clueweb09-en0005-98-00110 0
```

Qrels

Topic

Subtopic/Aspect

Document Id

Relevance Judgment

```
151 1 clueweb09-en0005-98-00099 1
151 2 clueweb09-en0005-98-00099 0
151 3 clueweb09-en0005-98-00099 0
151 4 clueweb09-en0005-98-00099 0
151 5 clueweb09-en0005-98-00099 0
151 1 clueweb09-en0005-98-00107 2
151 2 clueweb09-en0005-98-00107 2
151 3 clueweb09-en0005-98-00107 2
151 4 clueweb09-en0005-98-00107 2
151 5 clueweb09-en0005-98-00107 4
151 1 clueweb09-en0005-98-00108 1
151 2 clueweb09-en0005-98-00108 0
151 3 clueweb09-en0005-98-00108 0
151 4 clueweb09-en0005-98-00108 0
151 5 clueweb09-en0005-98-00108 0
151 1 clueweb09-en0005-98-00110 1
151 2 clueweb09-en0005-98-00110 0
```

Qrels

Topic

Subtopic/Aspect

Document Id

Relevance Judgment

```
151 1 clueweb09-en0005-98-00099 1
151 2 clueweb09-en0005-98-00099 0
151 3 clueweb09-en0005-98-00099 0
151 4 clueweb09-en0005-98-00099 0
151 5 clueweb09-en0005-98-00099 0
151 1 clueweb09-en0005-98-00107 2
151 2 clueweb09-en0005-98-00107 2
151 3 clueweb09-en0005-98-00107 2
151 4 clueweb09-en0005-98-00107 2
151 5 clueweb09-en0005-98-00107 4
151 1 clueweb09-en0005-98-00108 1
151 2 clueweb09-en0005-98-00108 0
151 3 clueweb09-en0005
151 4 clueweb09-en0005
151 5 clueweb09-en0005
151 1 clueweb09-en0005
151 2 clueweb09-en0005
```

Ranking

```
151 Q0 clueweb09-enwp02-06-01125 1 32.38 example2012
151 Q0 clueweb09-en0011-25-31331 2 29.73 example2012
151 Q0 clueweb09-en0006-97-08104 3 21.93 example2012
151 Q0 clueweb09-en0009-82-23589 4 21.34 example2012
151 Q0 clueweb09-en0001-51-20258 5 21.06 example2012
151 Q0 clueweb09-en0002-99-12860 6 13.00 example2012
151 Q0 clueweb09-en0003-08-08637 7 12.87 example2012
151 Q0 clueweb09-en0004-79-18096 8 11.13 example2012
151 Q0 clueweb09-en0008-90-04729 9 10.72 example2012
```

Qrels

Topic | Subtopic/Aspect | Document Id | Relevance Judgment

```
151 1 clueweb09-en0005-98-00099 1
151 2 clueweb09-en0005-98-00099 0
151 3 clueweb09-en0005-98-00099 0
151 4 clueweb09-en0005-98-00099 0
151 5 clueweb09-en0005-98-00099 0
151 1 clueweb09-en0005-98-00107 2
151 2 clueweb09-en0005-98-00107 2
151 3 clueweb09-en0005-98-00107 2
151 4 clueweb09-en0005-98-00107 2
151 5 clueweb09-en0005-98-00107 4
151 1 clueweb09-en0005-98-00108 1
151 2 clueweb09-en0005-98-00108 0
151 3 clueweb09-en0005
151 4 clueweb09-en0005
151 5 clueweb09-en0005
151 1 clueweb09-en0005
151 2 clueweb09-en0005
```

Ranking

```
151 Q0 clueweb09-enwp02-06-01125 1 32.38 example2012
151 Q0 clueweb09-en0011-25-31331 2 29.73 example2012
151 Q0 clueweb09-en0006-97-08104 3 21.93 example2012
151 Q0 clueweb09-en0009-82-23589 4 21.34 example2012
151 Q0 clueweb09-en0001-51-20258 5 21.06 example2012
151 Q0 clueweb09-en0002-99-12860 6 13.00 example2012
151 Q0 clueweb09-en0003-08-08637 7 12.87 example2012
151 Q0 clueweb09-en0004-79-18096 8 11.13 example2012
151 Q0 clueweb09-en0008-90-04729 9 10.72 example2012
```

Evaluation Measure(Ranking, qrels) = Score

About Metrics

Go to **www.menti.com** and use the code **3168 7558**

https://www.menti.com/eh8eb77k8w

Effectiveness evaluation metrics for diversity

Mentimeter

exploration vs exploitati

inverse document frequenc

nr of subtopic in top x

rank biased overlap

distance unrelated feats

common results list

no idea

overlap

relevance

income

alpha-ndcg

map

cheese

ndcg

no

ab testing

take into account user

item dissimilarity

most clicks

cosine similarity

vector dissimilarity

some vector distance

relative representation

23

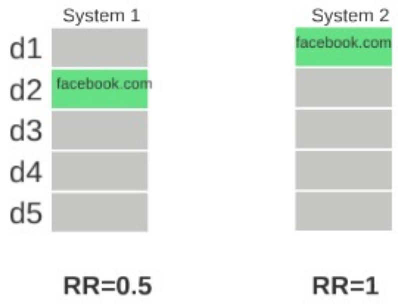
How Do We Pick The Right Metric?

Example

Metric to analyze which system performs better for **navigational queries**?

query: facebook

Reciprocal Rank (RR)



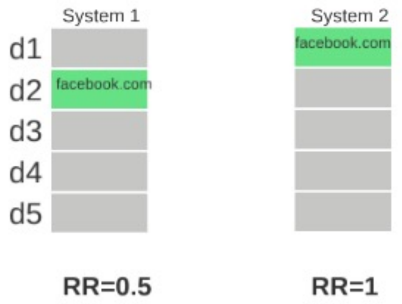
How Do We Pick The Right Metric?

Example

Metric to analyze which system performs better for **navigational queries**?

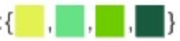
query: facebook

Reciprocal Rank (RR)



and for **Search Result Diversification**?

query: things to do in Nijmegen

query aspects:{  }

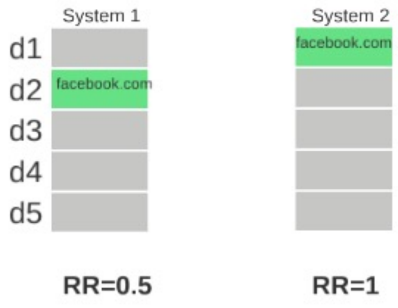
How Do We Pick The Right Metric?

Example

Metric to analyze which system performs better for **navigational queries**?

query: facebook

Reciprocal Rank (RR)



and for **Search Result Diversification**?

query: things to do in Nijmegen

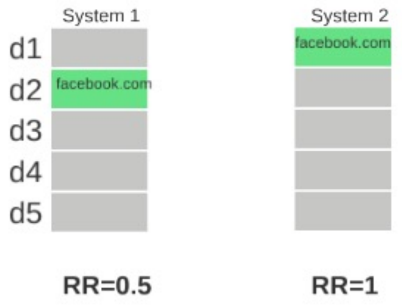
How Do We Pick The Right Metric?

Example

Metric to analyze which system performs better for **navigational queries**?

query: facebook

Reciprocal Rank (RR)



and for **Search Result Diversification**?

query: things to do in Nijmegen

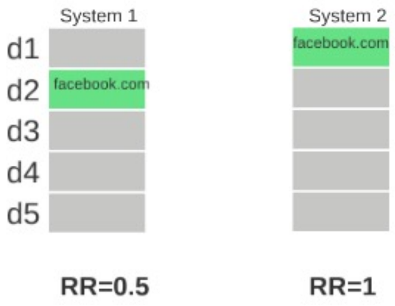
How Do We Pick The Right Metric?

Example

Metric to analyze which system performs better for **navigational queries**?

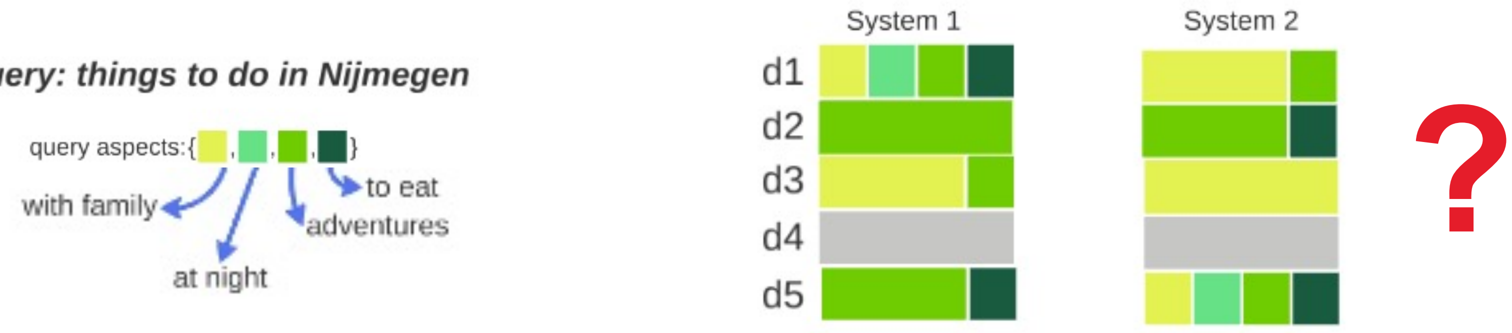
query: facebook

Reciprocal Rank (RR)



and for **Search Result Diversification**?

query: things to do in Nijmegen



Which Metric Would You Pick?

Go to **www.menti.com** and use the code **3168 7558**

https://www.menti.com/eh8eb77k8w

Which Approach Would You Follow To Pick The Right Evaluation Measure?

I.' Mentimeter

Study metric behaviour in relation to desirable properties wins!

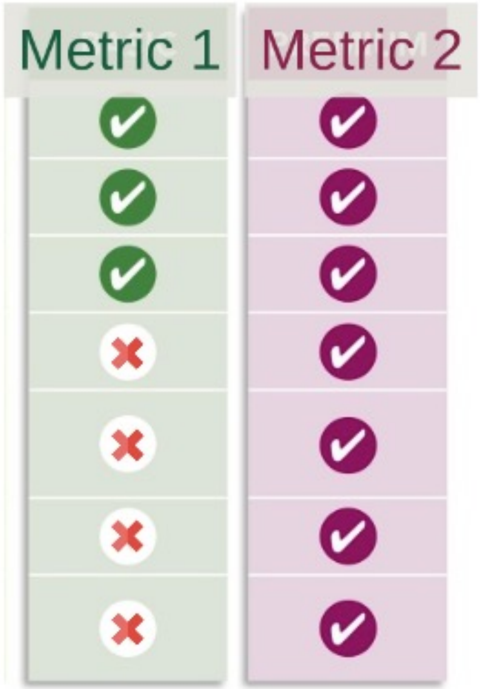
44

Axiomatic Analysis

Metric 1	Metric 2
✓	✓
✓	✓
✓	✓
✗	✓
✗	✓
✗	✓
✗	✓

Axiomatic Analysis

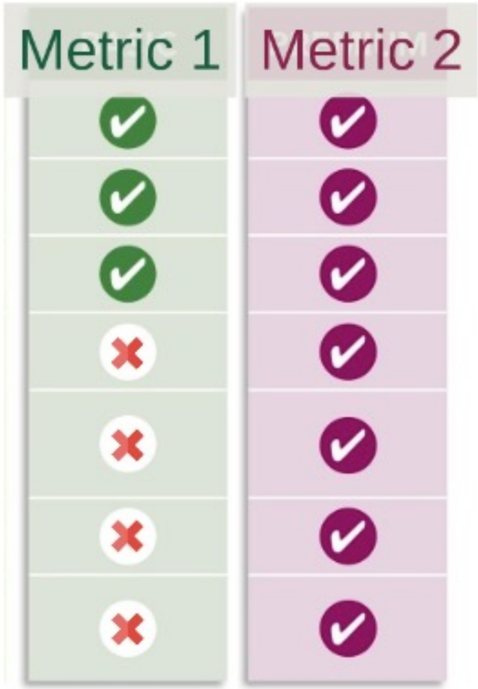
We may not be able to prescribe how to design the evaluation metric for our task...



Axiomatic Analysis

We may not be able to prescribe how to design the evaluation metric for our task...

...but we know certain **boundary conditions** (or constraints) of our problem

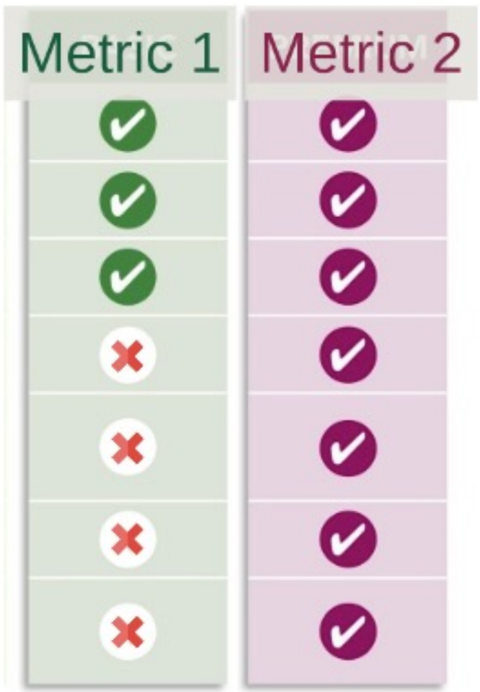


Axiomatic Analysis

We may not be able to prescribe how to design the evaluation metric for our task...

...but we know certain **boundary conditions** (or constraints) of our problem

desirable properties of a metric for our problem (e.g., search result diversification)



An Axiomatic Analysis of Diversity Evaluation Metrics: Introducing the Rank-Biased Utility Metric

Enrique Amigó, Damiano Spina, and Jorge Carrillo-de-Albornoz. 2018. An Axiomatic Analysis of Diversity Evaluation Metrics: Introducing the Rank-Biased Utility Metric. In *Proceedings of SIGIR'18.* DOI:https://doi.org/10.1145/3209978.3210024

P@k
RR
AP
nDCG@k
ERR@k
RBP
P-IA@k
RR-IA@k
AP-IA
nDCG-IA@k
ERR-IA@k
RBP-IA
S-Recall@k
S-RR@100%
NRBP
D#-Measure@k
α-nDCG@k
EU
CT@k

	Priority	Deepness	Deepness Threshold	Closeness	Confidence
P@k	○	○	●	●	○
RR	○	○	●	○	○
AP	○	○	○	●	○
nDCG@k	●	●	○	●	○
ERR@k	●	●	●	○	○
RBP	●	●	●	●	○
P-IA@k	○	○	●	●	○
RR-IA@k	○	○	●	○	○
AP-IA	●	●	○	●	○
nDCG-IA@k	●	●	○	●	○
ERR-IA@k	●	●	●	○	○
RBP-IA	●	●	●	●	○
S-Recall@k	○	○	●	○	○
S-RR@100%	○	○	●	○	○
NRBP	●	●	●	●	○
D#-Measure@k	●	●	○	●	○
α-nDCG@k	●	●	●	●	○
EU	●	●	●	●	●
CT@k	●	●	●	○	○

	Priority	Deepness	Deepness Threshold	Closeness	Confidence	Query Aspect Diversity	Redundancy	Monotonic Redundancy	Aspect Relevance Saturation	Aspect Relevance
P@k	○	○	●	●	○	○	○	○	○	○
RR	○	○	●	○	○	○	○	○	○	○
AP	○	○	○	●	○	○	○	○	○	○
nDCG@k	●	●	○	●	○	○	○	○	○	○
ERR@k	●	●	●	○	○	○	○	○	○	○
RBP	●	●	●	●	○	○	○	○	○	○
P-IA@k	○	○	●	●	○	○	○	○	○	●
RR-IA@k	○	○	●	○	○	○	○	○	●	●
AP-IA	●	●	○	●	○	○	○	○	○	●
nDCG-IA@k	●	●	○	●	○	●	○	○	○	●
ERR-IA@k	●	●	●	○	○	●	●	●	●	●
RBP-IA	●	●	●	●	○	●	○	○	○	●
S-Recall@k	○	○	●	○	○	○	○	○	●	○
S-RR@100%	○	○	●	○	○	○	○	○	●	○
NRBP	●	●	●	●	○	●	●	○	○	○
D#-Measure@k	●	●	○	●	○	●	○	○	○	●
α-nDCG@k	●	●	●	●	○	●	●	○	○	○
EU	●	●	●	●	●	●	●	○	○	●
CT@k	●	●	●	○	○	●	●	○	●	●

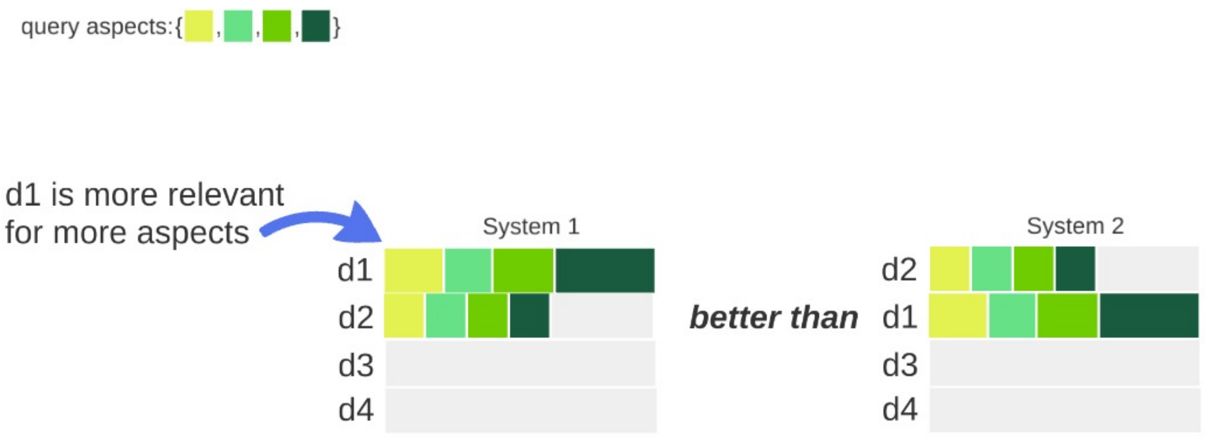
	Priority	Deepness	Deepness Threshold	Closeness	Confidence	Query Aspect Diversity	Redundancy	Monotonic Redundancy	Aspect Relevance Saturation	Aspect Relevance
P@k	○	○	●	●	○	○	○	○	○	○
RR	○	○	●	○	○	○	○	○	○	○
AP	○	○	○	●	○	○	○	○	○	○
nDCG@k	●	●	○	●	○	○	○	○	○	○
ERR@k	●	●	●	○	○	○	○	○	○	○
RBP	●	●	●	●	○	○	○	○	○	○
P-IA@k	○	○	●	●	○	○	○	○	○	●
RR-IA@k	○	○	●	○	○	○	○	○	●	●
AP-IA	●	●	○	●	○	○	○	○	○	●
nDCG-IA@k	●	●	○	●	○	●	○	○	○	●
ERR-IA@k	●	●	●	○	○	●	●	●	●	●
RBP-IA	●	●	●	●	○	●	○	○	○	●
S-Recall@k	○	○	●	○	○	○	○	○	●	○
S-RR@100%	○	○	●	○	○	○	○	○	●	○
NRBP	●	●	●	●	○	●	●	○	○	○
D#-Measure@k	●	●	○	●	○	●	○	○	○	●
α-nDCG@k	●	●	●	●	○	●	●	○	○	○
EU	●	●	●	●	●	●	●	○	○	●
CT@k	●	●	●	○	○	●	●	○	●	●
RBU@k	●	●	●	●	●	●	●	●	●	●

	Priority	Deepness	Deepness Threshold	Closeness	Confidence	Query Aspect Diversity	Redundancy	Monotonic Redundancy	Aspect Relevance Saturation	Aspect Relevance
P@k	○	○	●	●	○	○	○	○	○	○
RR	○	○	●	○	○	○	○	○	○	○
AP	○	○	○	●	○	○	○	○	○	○
nDCG@k	●	●	○	●	○	○	○	○	○	○
ERR@k	●	●	●	○	○	○	○	○	○	○
RBP	●	●	●	●	○	○	○	○	○	○
P-IA@k	○	○	●	●	○	○	○	○	○	●
RR-IA@k	○	○	●	○	○	○	○	○	●	●
AP-IA	●	●	○	●	○	○	○	○	○	●
nDCG-IA@k	●	●	○	●	○	●	○	○	○	●
ERR-IA@k	●	●	●	○	○	●	●	●	●	●
RBP-IA	●	●	●	●	○	●	○	○	○	●
S-Recall@k	○	○	●	○	○	○	○	○	●	○
S-RR@100%	○	○	●	○	○	○	○	○	●	○
NRBP	●	●	●	●	○	●	●	○	○	○
D#-Measure@k	●	●	○	●	○	●	○	○	○	●
α-nDCG@k	●	●	●	●	○	●	●	○	○	○
EU	●	●	●	●	●	●	●	○	○	●
CT@k	●	●	●	○	○	●	●	○	●	●
RBU@k	●	●	●	●	●	●	●	●	●	●

Formal Properties (Diversity)

Query Aspect Diversity

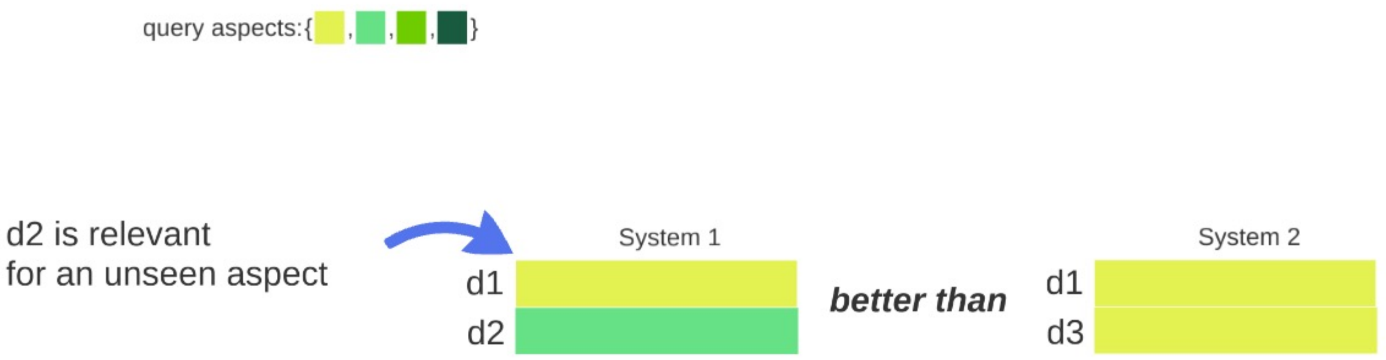
Covering more about the aspects in the same document (i.e., without additional effort of inspecting more documents) increases the score.



Formal Properties (Diversity)

Redundancy

Adding a document from a less present (less redundant) aspect, increases the score



	Priority	Deepness	Deepness Threshold	Closeness	Confidence	Query Aspect Diversity	Redundancy	Monotonic Redundancy	Aspect Relevance Saturation	Aspect Relevance
P@k	○	○	●	●	○	○	○	○	○	○
RR	○	○	●	○	○	○	○	○	○	○
AP	○	○	○	●	○	○	○	○	○	○
nDCG@k	●	●	○	●	○	○	○	○	○	○
ERR@k	●	●	●	○	○	○	○	○	○	○
RBP	●	●	●	●	○	○	○	○	○	○
P-IA@k	○	○	●	●	○	○	○	○	○	●
RR-IA@k	○	○	●	○	○	○	○	○	●	●
AP-IA	●	●	○	●	○	○	○	○	○	●
nDCG-IA@k	●	●	○	●	○	●	○	○	○	●
ERR-IA@k	●	●	●	○	○	●	●	●	●	●
RBP-IA	●	●	●	●	○	●	○	○	○	●
S-Recall@k	○	○	●	○	○	○	○	○	●	○
S-RR@100%	○	○	●	○	○	○	○	○	●	○
NRBP	●	●	●	●	○	●	●	○	○	○
D#-Measure@k	●	●	○	●	○	●	○	○	○	●
α-nDCG@k	●	●	●	●	○	●	●	○	○	○
EU	●	●	●	●	●	●	●	○	○	●
CT@k	●	●	●	○	○	●	●	○	●	●
RBU@k	●	●	●	●	●	●	●	●	●	●

NEW!

RBU: Rank-Biased Utility

aspect weight (*-IA)

relevance

redundancy (ERR)

$$\mathrm{RBU@k}(\vec{d}) = \sum_{i=1}^{k} p^i \left(\sum_{t \in \mathcal{T}} \left(w(t) r(d_i, t) \prod_{j=1}^{i-1} \left(1 - r(d_j, t)\right) \right) - e \right)$$

patience parameter (RBP)

effort (EU, Expected Utility)

ranking $\vec{d}$

k = *min(cutoff, $|\vec{d}|$)*

set of aspects $\mathcal{T}$
each aspect *t* has a weight *w(t)*

https://github.com/rmit-ir/RBU and also available in EvALL: http://evall.uned.es/

Diversity

and

Fairness

from a

Ranking Evaluation

perspective

Diversity

and

Fairness

from a

Ranking Evaluation

perspective

Is Diversity Enough?

Go to **www.menti.com** and use the code **3168 7558**

https://www.menti.com/eh8eb77k8w

Djoerd Hiemstra
Research, Teaching and More

Home Contact Publications

Fairness in Information Retrieval

Djoerd Hiemstra
June 22, 2021
Fair Machine Learning

Was fairness in IR discussed by Cooper and Robertson in the 1970's?

https://djoerdhiemstra.com/2021/fairness-in-information-retrieval/

Describe a scenario on which rankings mav lead to **discriminative or unfair outcomes**

Mentimeter

HR/ job search	Automatic candidate selection for jobs	Smaller parties can be left out
if the system was evaluated by a non diverse group of people	Ranking candidates for job vacancies, or ranking job vacancies for candidates.	If all top rankings have a certain characteristic but that doesn't actually cause the increased relevancy
Job offers	Fraud	Candidate ranking for job vacancies

30

Describe a scenario on which rankings mav lead to **discriminative or unfair outcomes**

bank loan discrimination

Elections

Scholarship grants

Ranking might put fake news higher than other real news.

Ranking or recommendation of online (fake) news articles.

A fake story about a person may be ranked higher then a true story about the person because the true stray is less interesting. Hence the ranking helps to spread fake news

Information about political decisions, public opinions and facts.

Face recognition

Web stores where users can market their own products

Describe a scenario on which rankings mav lead to **discriminative or unfair outcomes**

Mentimeter

Image retrieval.

(Face recognition applied in surveillance - I'd like to be discriminated against here)

Etsy

Mortgages

Finding relevant previous court cases to use as basis for decision in new case

The system is trained on what a specific majority finds relevant.

election voting compass

Recommender systems that do not explore new recommendations, and only recommend the most popular items.

Finding a doctor (e.g. male doctors might rank above female doctors)

Vacancies

cheeseballz

30

Fairness

Fairness

Individual Fairness

Similar individuals should be treated similarly.

Two candidates with the same skills and experience should receive the same treatment (e.g., positioned similarly in rankings).

Fairness

Individual Fairness

Similar individuals should be treated similarly.

Two candidates with the same skills and experience should receive the same treatment (e.g., positioned similarly in rankings).

Group Fairness

Each salient group should be treated comparably.

Female candidates should not be less likely to get shortlisted than male candidates, and vice versa.

Evaluating Fairness in Argument Retrieval

Sachin Pathiyan Cherumanal, Damiano Spina, Falk Scholer, and W. Bruce Croft. 2021. Evaluating Fairness in Argument Retrieval. In *Proceedings of CIKM'21.* DOI:https://doi.org/10.1145/3459637.3482099

args uniforms in school

All Discussions News People

Topic space view 1242 arguments retrieved in 0.0 ms

PRO 2) The Rationale Behind School Uniforms School uniforms...

▶ Show full argument

(2) The Rationale Behind **School Uniforms School uniforms** are useful because of the fact that they not only restrict students" clothing options, but they prevent the problems that come from "individualized" dress. ... Because ...

https://www.debate.org/debates/school-uniforms/20/ score

PRO Those statistics are getting better because of school...

▶ Show full argument

Those statistics are getting better because of **school uniforms**. ... I believe if schools can adapt good **school uniforms**, Like the **uniforms** at Long Beach California, Then the **school** will be a more professional and a more safe ...

https://www.debate.org/debates/School-Uniforms/85/ score

PRO Although most say that parents wont be able to afford...

▶ Show full argument

Although most say that parents wont be able to afford these **uniforms**, it is also known that **uniforms** help stop bullying. ... I've heard and seen many humans being treated wrongly because clothing, but that can stop thanks to ...

https://www.debate.org/debates/School-Uniforms/17/ score

CON 2) By using a school uniform, it is promoted that a...

▶ Show full argument

(2) By using a **school** uniform, it is promoted that a student has to conform to society and not individuality. ... It is shown to be a contradictory method when **school uniforms** are put into place as it contradicts the lesson that ...

https://www.debate.org/debates/School-Uniforms/80/ score

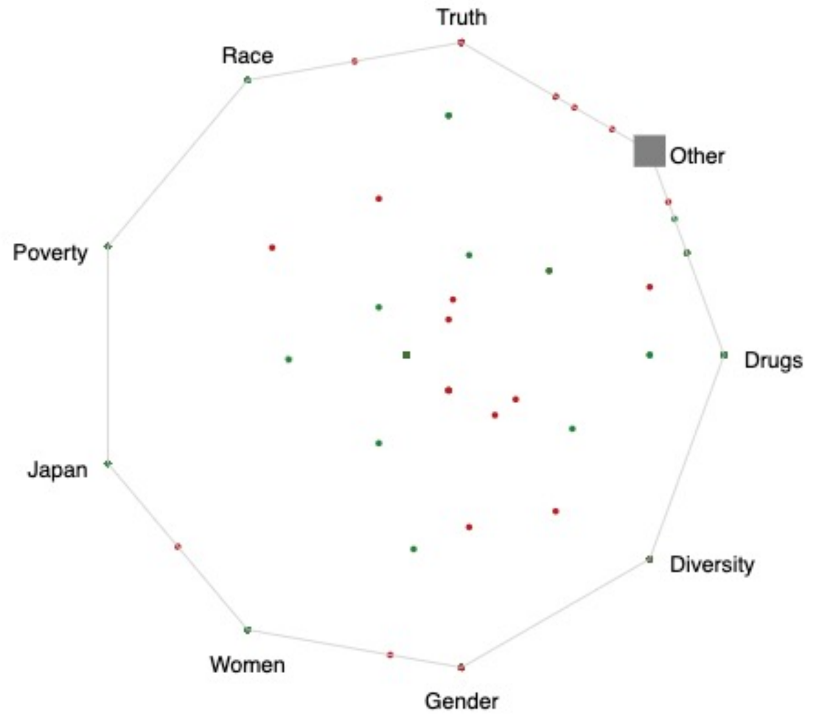
PRO I understand that there is always a good chance that...

▶ Show full argument

I understand that there is always a good chance that parents might teach students how to dress appropriately but this doesn't always happen. ... Sources: http://www.angelfire.com... http://www.ehow.com...

https://www.debate.org/debates/Uniforms-at-School/1/ score

Stance
PRO
CON



Source: https://www.args.me/search.html?query=school%20uniforms

(Un)fairness Metrics

Ke Yang and Julia Stoyanovich. 2017. Measuring Fairness in Ranked Outputs. In Proceedings of the 29th International Conference on Scientific and Statistical Database Management (SSDBM '17). Association for Computing Machinery, New York, NY, USA, Article 22, 1–6. DOI:https://doi.org/10.1145/3085504.3085526

(Un)fairness Metrics

Statistical Parity: proportion in various cut-offs of the ranking is similar to the proportion in the population

Ke Yang and Julia Stoyanovich. 2017. Measuring Fairness in Ranked Outputs. In Proceedings of the 29th International Conference on Scientific and Statistical Database Management (SSDBM '17). Association for Computing Machinery, New York, NY, USA, Article 22, 1–6. DOI:https://doi.org/10.1145/3085504.3085526

(Un)fairness Metrics

Statistical Parity: proportion in various cut-offs of the ranking is similar to the proportion in the population

Different ways of comparing distribution in the sample (ranking) with distribution in population (ground truth)

Ke Yang and Julia Stoyanovich. 2017. Measuring Fairness in Ranked Outputs. In Proceedings of the 29th International Conference on Scientific and Statistical Database Management (SSDBM '17). Association for Computing Machinery, New York, NY, USA, Article 22, 1–6. DOI:https://doi.org/10.1145/3085504.3085526

(Un)fairness Metrics

Statistical Parity: proportion in various cut-offs of the ranking is similar to the proportion in the population

Different ways of comparing distribution in the sample (ranking) with distribution in population (ground truth)

- Normalized Discounted Difference (**rND**)

Ke Yang and Julia Stoyanovich. 2017. Measuring Fairness in Ranked Outputs. In Proceedings of the 29th International Conference on Scientific and Statistical Database Management (SSDBM '17). Association for Computing Machinery, New York, NY, USA, Article 22, 1–6. DOI:https://doi.org/10.1145/3085504.3085526

(Un)fairness Metrics

Statistical Parity: proportion in various cut-offs of the ranking is similar to the proportion in the population

Different ways of comparing distribution in the sample (ranking) with distribution in population (ground truth)

- Normalized Discounted Difference (**rND**)
- Normalized Discounted K-L Divergence (**rKL**)

Ke Yang and Julia Stoyanovich. 2017. Measuring Fairness in Ranked Outputs. In Proceedings of the 29th International Conference on Scientific and Statistical Database Management (SSDBM '17). Association for Computing Machinery, New York, NY, USA, Article 22, 1–6. DOI:https://doi.org/10.1145/3085504.3085526

(Un)fairness Metrics

Statistical Parity: proportion in various cut-offs of the ranking is similar to the proportion in the population

Different ways of comparing distribution in the sample (ranking) with distribution in population (ground truth)

- Normalized Discounted Difference (**rND**)
- Normalized Discounted K-L Divergence (**rKL**)
- Normalized Discounted Ratio (**rRD**)

Ke Yang and Julia Stoyanovich. 2017. Measuring Fairness in Ranked Outputs. In Proceedings of the 29th International Conference on Scientific and Statistical Database Management (SSDBM '17). Association for Computing Machinery, New York, NY, USA, Article 22, 1–6. DOI:https://doi.org/10.1145/3085504.3085526

Experimental Setup

Code: https://github.com/rmit-ir/fair-arguments

Experimental Setup

- System runs from Touché Lab at CLEF 2020.

Code: https://github.com/rmit-ir/fair-arguments

Experimental Setup

- System runs from Touché Lab at CLEF 2020.
- Relevance (NDCG), (Un)fairness metrics,

Code: https://github.com/rmit-ir/fair-arguments

Experimental Setup

- System runs from Touché Lab at CLEF 2020.
- Relevance (NDCG), (Un)fairness metrics, and diversity (α-NDCG)

Code: https://github.com/rmit-ir/fair-arguments

Experimental Setup

- System runs from Touché Lab at CLEF 2020.
- Relevance (NDCG), (Un)fairness metrics, and diversity (α-NDCG)

Code: https://github.com/rmit-ir/fair-arguments

Experimental Setup

- System runs from Touché Lab at CLEF 2020.
- Relevance (NDCG), (Un)fairness metrics, and diversity (α-NDCG)

Topic: Universal Basic Income	
Argument	**Stance**
...lifting society out of poverty….	PRO
UBI is individually destructive ...	CON

Code: https://github.com/rmit-ir/fair-arguments

Experimental Setup

- System runs from Touché Lab at CLEF 2020.
- Relevance (NDCG), (Un)fairness metrics, and diversity (α-NDCG)
- **Test Collection:** Ground truth of 2,964 relevance judgments across 49 topics.

Topic: Universal Basic Income	
Argument	**Stance**
...lifting society out of poverty….	PRO
UBI is individually destructive ...	CON

Code: https://github.com/rmit-ir/fair-arguments

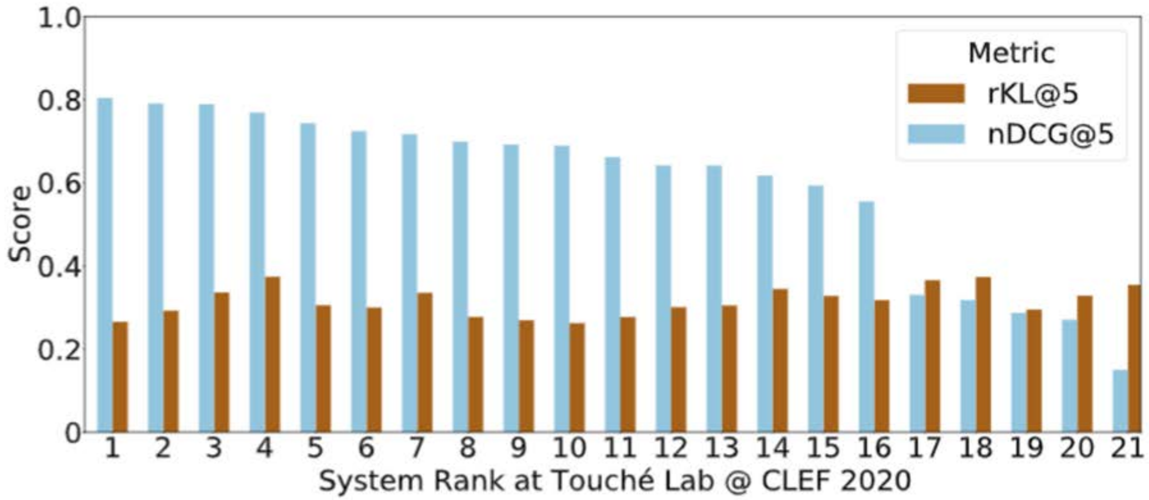
Experimental Setup

- System runs from Touché Lab at CLEF 2020.
- Relevance (NDCG), (Un)fairness metrics, and diversity (α-NDCG)
- **Test Collection:** Ground truth of 2,964 relevance judgments across 49 topics.
- **Relevance:** Graded between 1 (least relevant) and 5 (most relevant).

Topic: Universal Basic Income	
Argument	**Stance**
...lifting society out of poverty….	PRO
UBI is individually destructive ...	CON

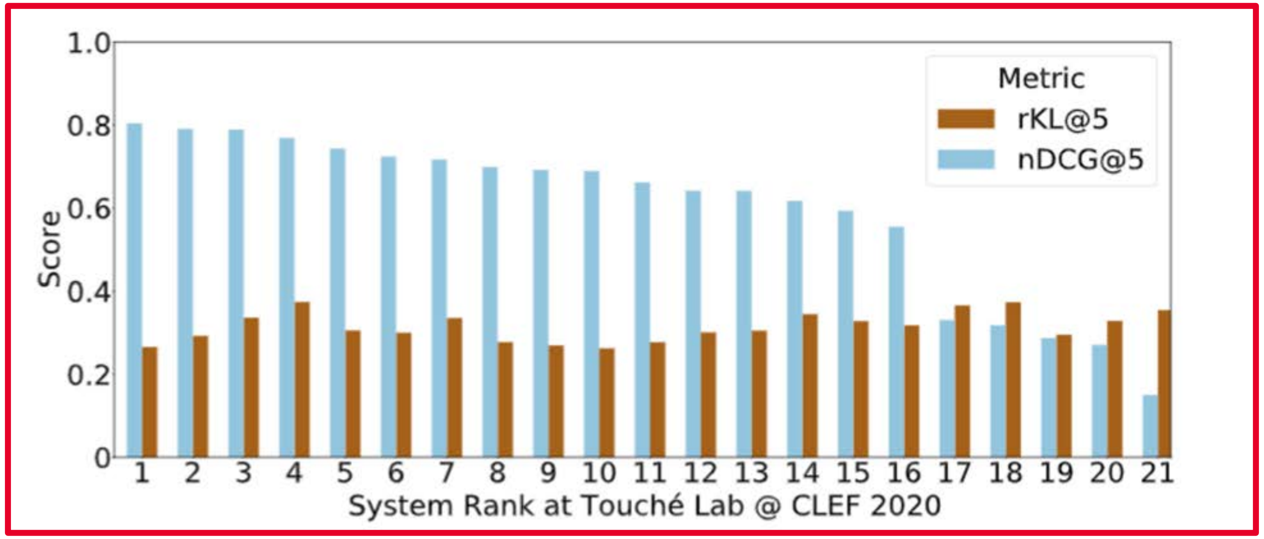
Code: https://github.com/rmit-ir/fair-arguments

Results



For nDCG@5 higher is better and for (un)fairness metric rKL@5, lower is better.

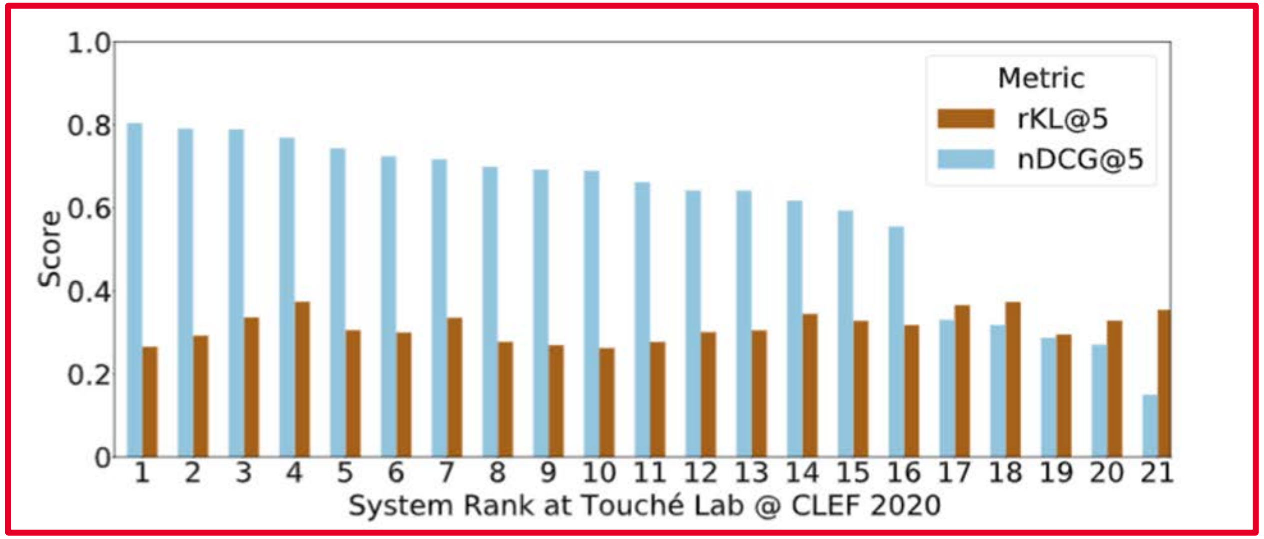
Results



For nDCG@5 higher is better and for (un)fairness metric rKL@5, lower is better.

Results

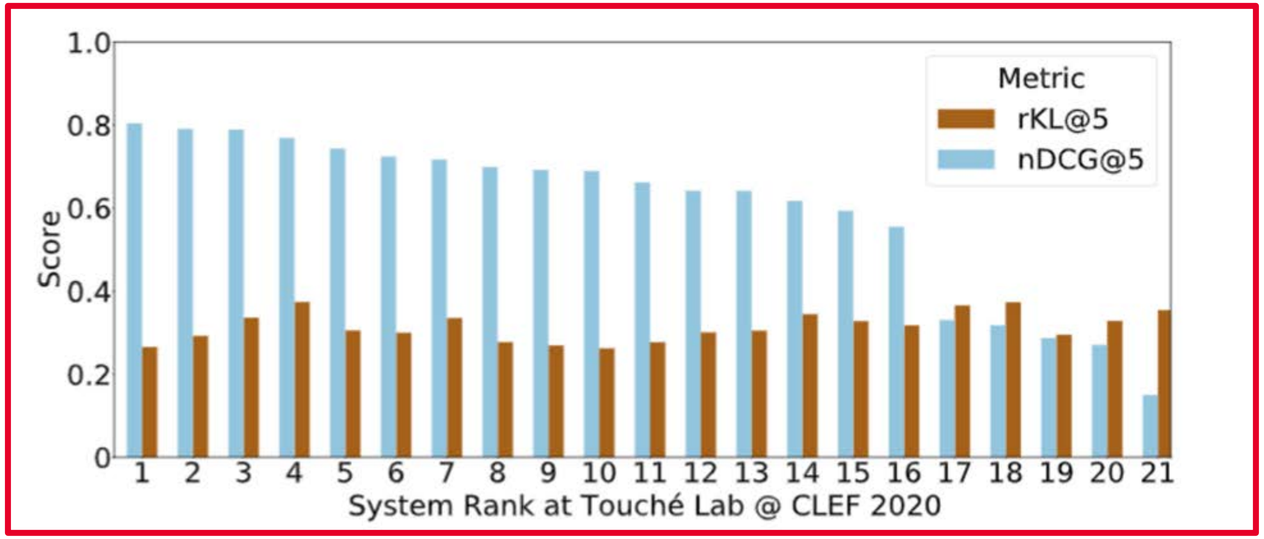
- (Un)fairness metrics do not increase monotonically w.r.t. NDCG@5



For nDCG@5 higher is better and for (un)fairness metric rKL@5, lower is better.

Results

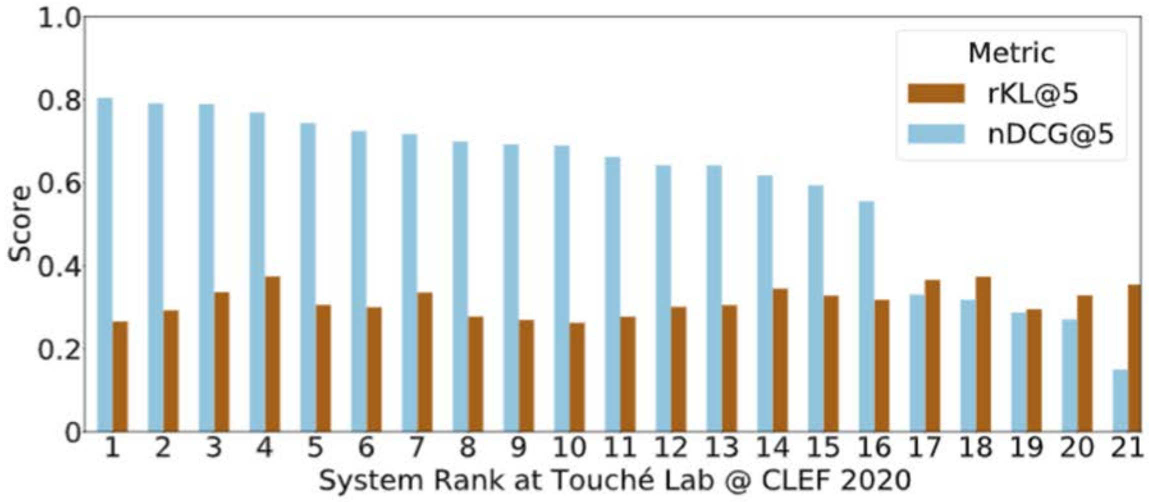
- (Un)fairness metrics do not increase monotonically w.r.t. NDCG@5
- System ranks would change when ranked using both (un)fairness and relevance.



For nDCG@5 higher is better and for (un)fairness metric rKL@5, lower is better.

Results

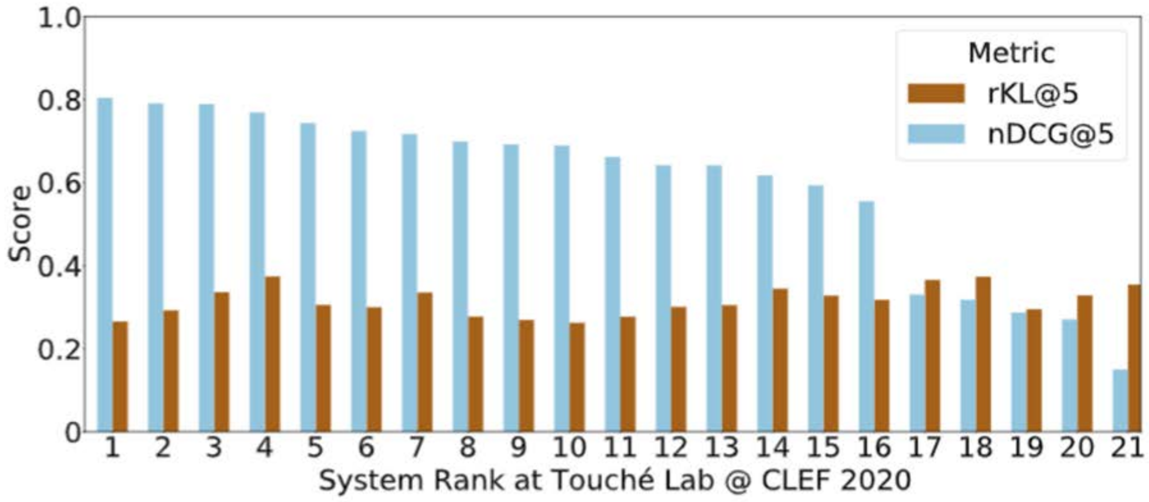
- (Un)fairness metrics do not increase monotonically w.r.t. NDCG@5
- System ranks would change when ranked using both (un)fairness and relevance.



For nDCG@5 higher is better and for (un)fairness metric rKL@5, lower is better.

Results

- (Un)fairness metrics do not increase monotonically w.r.t. NDCG@5
- System ranks would change when ranked using both (un)fairness and relevance.

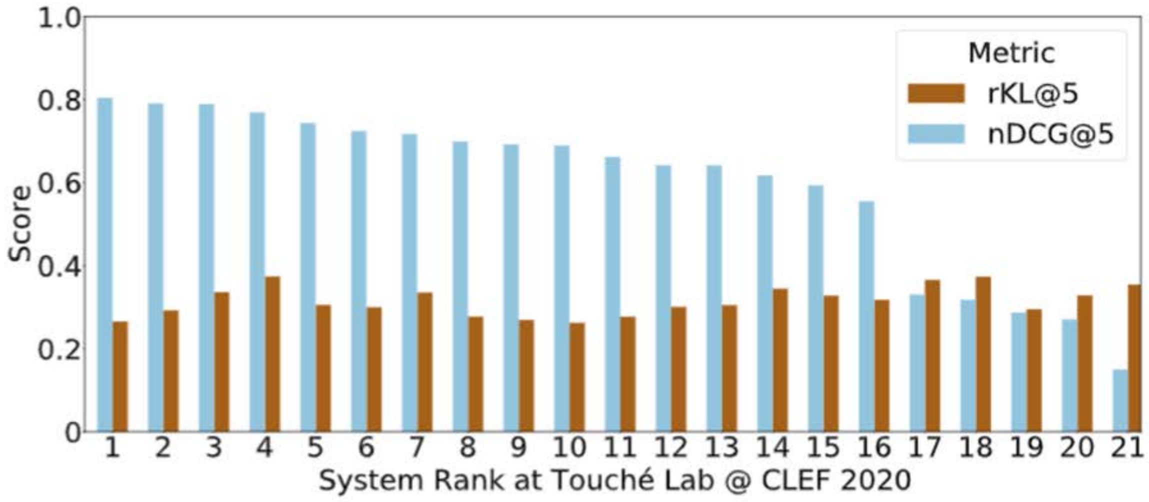


For nDCG@5 higher is better and for (un)fairness metric rKL@5, lower is better.

(un)fairness	nDCG@5	α-nDCG@5	rND@5	rKL@5
rND@5	−0.0762	−0.0667		
rKL@5	−0.2667	−0.3333*	0.3524*	
rRD@5	−0.2571	−0.3238*	−0.1143	0.2857

Results

- (Un)fairness metrics do not increase monotonically w.r.t. NDCG@5
- System ranks would change when ranked using both (un)fairness and relevance.

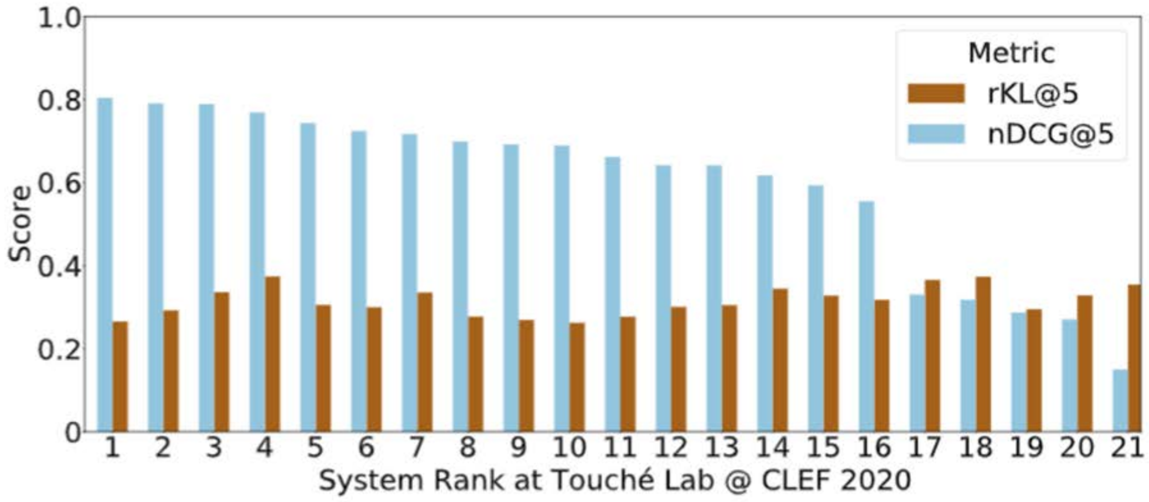


For nDCG@5 higher is better and for (un)fairness metric rKL@5, lower is better.

(un)fairness	nDCG@5	α-nDCG@5	rND@5	rKL@5
rND@5	−0.0762	−0.0667		
rKL@5	−0.2667	−0.3333*	0.3524*	
rRD@5	−0.2571	−0.3238*	−0.1143	0.2857

Results

- (Un)fairness metrics do not increase monotonically w.r.t. NDCG@5
- System ranks would change when ranked using both (un)fairness and relevance.
- Diversity is related but not equivalent to (un)fairness.



For nDCG@5 higher is better and for (un)fairness metric rKL@5, lower is better.

(un)fairness	nDCG@5	α-nDCG@5	rND@5	rKL@5
rND@5	−0.0762	−0.0667		
rKL@5	−0.2667	−0.3333*	0.3524*	
rRD@5	−0.2571	−0.3238*	−0.1143	0.2857

Summary

Summary

- Search results for controversial topics often had a biased exposure of perspectives/stances.

Summary

- Search results for controversial topics often had a biased exposure of perspectives/stances.
- Quantify this bias in exposure of different perspectives/stances.

Summary

- Search results for controversial topics often had a biased exposure of perspectives/stances.
- Quantify this bias in exposure of different perspectives/stances.
- Real and controlled experiments conducted to study the relationship between diversity and (un)fairness metrics.

Summary

- Search results for controversial topics often had a biased exposure of perspectives/stances.
- Quantify this bias in exposure of different perspectives/stances.
- Real and controlled experiments conducted to study the relationship between diversity and (un)fairness metrics.

Future Work:

Summary

- Search results for controversial topics often had a biased exposure of perspectives/stances.
- Quantify this bias in exposure of different perspectives/stances.
- Real and controlled experiments conducted to study the relationship between diversity and (un)fairness metrics.

Future Work:

- Fairness of topical categories along with stance i.e., Multi-Attribute fairness.

Summary

- Search results for controversial topics often had a biased exposure of perspectives/stances.
- Quantify this bias in exposure of different perspectives/stances.
- Real and controlled experiments conducted to study the relationship between diversity and (un)fairness metrics.

Future Work:

- Fairness of topical categories along with stance i.e., Multi-Attribute fairness.
- Analyze other fairness and diversity metrics

Formal Properties (Diversity)

Aspect Relevance

Aspects with higher weights have more effect in score of the ranking quality.

CONSTRAINT 10 (ASPECT RELEVANCE, ASPREL). *Aspects with higher weights have more effect in score of the ranking quality. Formally, assuming no aspect overlap, and being d_i and d'_i documents that are relevant to different aspects that have not been observed before, $\forall j < i.\ r(d_j, t) = r(d_j, t') = 0$, and $r(d_i, t) = r(d'_i, t') > 0$ then:*

$$w(t) < w(t') \Longrightarrow Q\left(\vec{d}_{d_i \leftrightarrow d'_i}\right) > Q\left(\vec{d}\right) \quad (10)$$

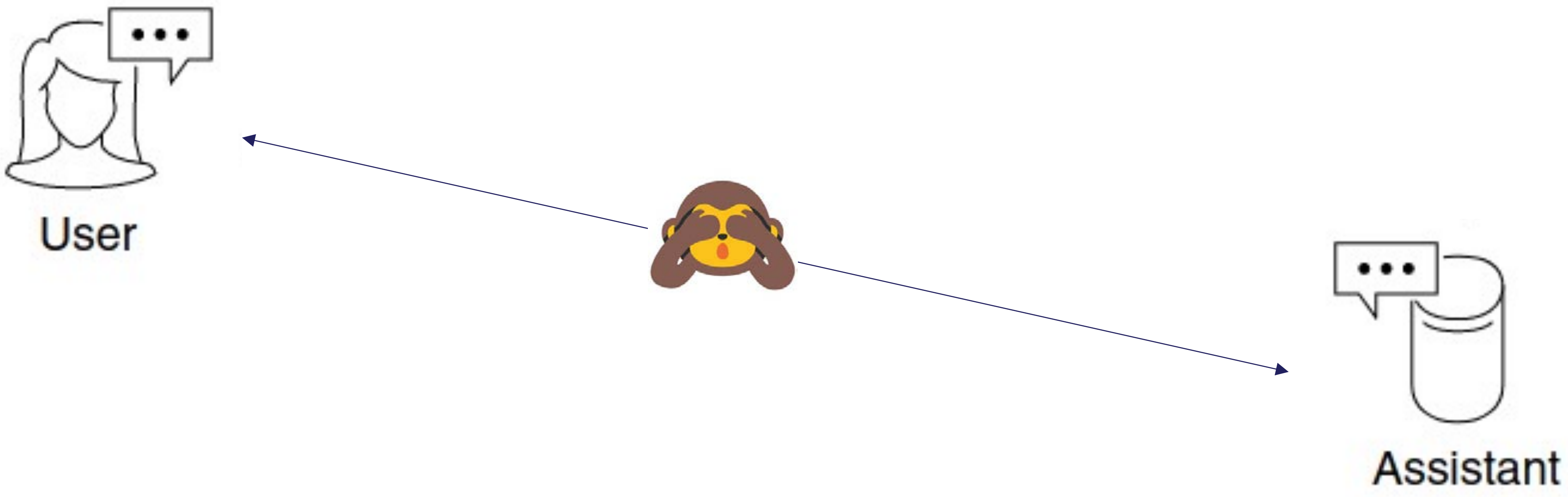
Assessing Viewpoint Diversity in Search Results Using Ranking Fairness Metrics

Tim Draws, Nava Tintarev, and Ujwal Gadiraju. 2021. Assessing Viewpoint Diversity in Search Results Using Ranking Fairness Metrics. *SIGKDD Explor. Newsl. 23, 1 (June 2021)*, 50–58. DOI:https://doi.org/10.1145/3468507.3468515

The Meant, the Said, and the Understood: Conversational Argument Search and Cognitive Biases

Johannes Kiesel, Damiano Spina, Henning Wachsmuth, and Benno Stein. 2021. The Meant, the Said, and the Understood: Conversational Argument Search and Cognitive Biases. *In Proceedings of the 3rd Conference on Conversational User Interfaces (CUI '21)* DOI:https://doi.org/10.1145/3469595.3469615

“Ok Google/Alexa/Siri, should students wear school uniforms?



Report on the Future Conversations Workshop at CHIIR 2021

Damiano Spina*
RMIT University
damiano.spina@rmit.edu.au

Johanne R. Trippas*
The University of Melbourne
johanne.trippas@unimelb.edu.au

Paul Thomas*
Microsoft
pathom@microsoft.com

Hideo Joho*
University of Tsukuba
hideo@slis.tsukuba.ac.jp

Katriina Byström[†] Leigh Clark[†] Nick Craswell[†] Mary Czerwinski[†]
David Elsweiler[†] Alexander Frummet[†] Souvick Ghosh[†]
Johannes Kiesel[†] Irene Lopatovska[†] Daniel McDuff[†] Selina Meyer[†]
Ahmed Mourad[†] Paul Owoicho[†] Sachin Pathiyan Cherumanal[†]
Daniel Russell[†] Laurianne Sitbon[†]

Abstract

The Future Conversations workshop at CHIIR'21 looked to the future of search, recommendation, and information interaction to ask: where are the opportunities for conversational interactions? What do we need to do to get there? Furthermore, who stands to benefit?

The workshop was hands-on and interactive. Rather than a series of technical talks, we solicited position statements on opportunities, problems, and solutions in conversational search in all modalities (written, spoken, or multimodal). This paper –co-authored by the organisers and participants of the workshop– summarises the submitted statements and the discussions we had during the two sessions of the workshop. Statements discussed during the workshop are available at https://bit.ly/FutureConversations2021Statements.

Report on the Future Conversations Workshop at CHIIR 2021

Damiano Spina*
RMIT University
damiano.spina@rmit.edu.au

Johanne R. Trippas*
The University of Melbourne
johanne.trippas@unimelb.edu.au

Paul Thomas*
Microsoft
pathom@microsoft.com

Hideo Joho*
University of Tsukuba
hideo@slis.tsukuba.ac.jp

Katriina Byström† Leigh Clark† Nick Craswell† Mary Czerwinski†
David Elsweiler† Alexander Frummet† Souvick Ghosh†
Johannes Kiesel† Irene Lopatovska† Daniel McDuff† Selina Meyer†
Ahmed Mourad† Paul Owoicho† Sachin Pathiyan Cherumanal†
Daniel Russell† Laurianne Sitbon†

Abstract

The Future Conversations workshop at CHIIR'21 looked to the future of search, recommendation, and information interaction to ask: where are the opportunities for conversational interactions? What do we need to do to get there? Furthermore, who stands to benefit?

The workshop was hands-on and interactive. Rather than a series of technical talks, we solicited position statements on opportunities, problems, and solutions in conversational search in all modalities (written, spoken, or multimodal). This paper –co-authored by the organisers and participants of the workshop– summarises the submitted statements and the discussions we had during the two sessions of the workshop. Statements discussed during the workshop are available at https://bit.ly/FutureConversations2021Statements.

2.3 Arguing with Search Engines

Johannes Kiesel, Bauhaus-Universität Weimar

Report on the Future Conversations Workshop at CHIIR 2021

Damiano Spina*
RMIT University
damiano.spina@rmit.edu.au

Johanne R. Trippas*
The University of Melbourne
johanne.trippas@unimelb.edu.au

Paul Thomas*
Microsoft
pathom@microsoft.com

Hideo Joho*
University of Tsukuba
hideo@slis.tsukuba.ac.jp

Katriina Byström† Leigh Clark† Nick Craswell† Mary Czerwinski†
David Elsweiler† Alexander Frummet† Souvick Ghosh†
Johannes Kiesel† Irene Lopatovska† Daniel McDuff† Selina Meyer†
Ahmed Mourad† Paul Owoicho† Sachin Pathiyan Cherumanal†
Daniel Russell† Laurianne Sitbon†

Abstract

The Future Conversations workshop at CHIIR'21 looked to the future of search, recommendation, and information interaction to ask: where are the opportunities for conversational interactions? What do we need to do to get there? Furthermore, who stands to benefit?

The workshop was hands-on and interactive. Rather than a series of technical talks, we solicited position statements on opportunities, problems, and solutions in conversational search in all modalities (written, spoken, or multimodal). This paper –co-authored by the organisers and participants of the workshop– summarises the submitted statements and the discussions we had during the two sessions of the workshop. Statements discussed during the workshop are available at https://bit.ly/FutureConversations2021Statements.

2.3 Arguing with Search Engines

Johannes Kiesel, Bauhaus-Universität Weimar

+

2.7 Future *Fair* Conversations

Damiano Spina, RMIT University

=

EVENT REPORT

Report on the Future Conversations Workshop at CHIIR 2021

Damiano Spina*
RMIT University
damiano.spina@rmit.edu.au

Johanne R. Trippas*
The University of Melbourne
johanne.trippas@unimelb.edu.au

Paul Thomas*
Microsoft
pathom@microsoft.com

Hideo Joho*
University of Tsukuba
hideo@slis.tsukuba.ac.jp

Katriina Byström† Leigh Clark† Nick Craswell† Mary Czerwinski†
David Elsweiler† Alexander Frummet† Souvick Ghosh†
Johannes Kiesel† Irene Lopatovska† Daniel McDuff† Selina Meyer†
Ahmed Mourad† Paul Owoicho† Sachin Pathiyan Cherumanal†
Daniel Russell† Laurianne Sitbon†

Abstract

The Future Conversations workshop at CHIIR'21 looked to the future of search, recommendation, and information interaction to ask: where are the opportunities for conversational interactions? What do we need to do to get there? Furthermore, who stands to benefit?

The workshop was hands-on and interactive. Rather than a series of technical talks, we solicited position statements on opportunities, problems, and solutions in conversational search in all modalities (written, spoken, or multimodal). This paper –co-authored by the organisers and participants of the workshop– summarises the submitted statements and the discussions we had during the two sessions of the workshop. Statements discussed during the workshop are available at https://bit.ly/FutureConversations2021Statements.

2.3 Arguing with Search Engines

Johannes Kiesel, Bauhaus-Universität Weimar

+

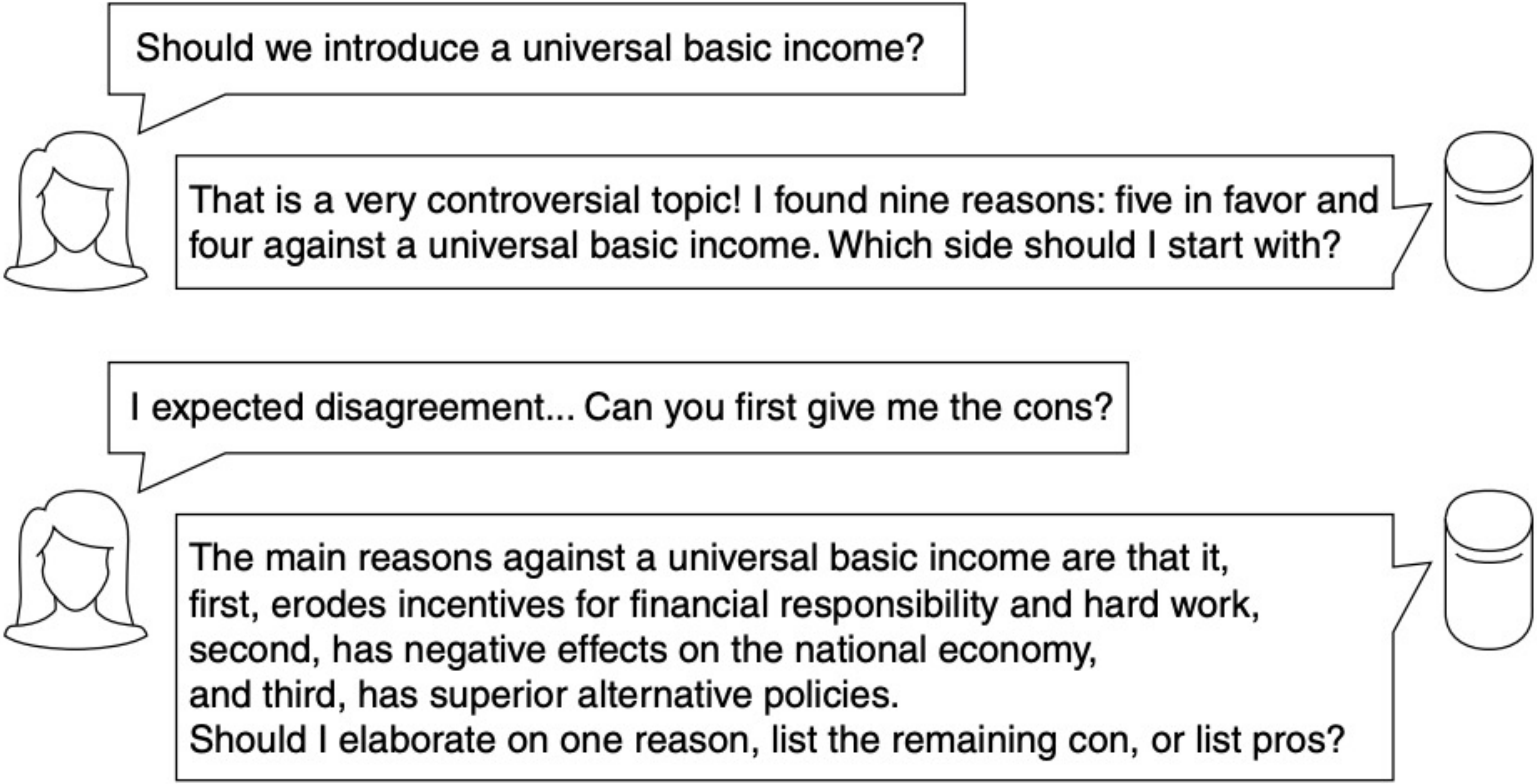
2.7 Future *Fair* Conversations

Damiano Spina, RMIT University

=

The Meant, the Said, and the Understood: Conversational Argument Search and Cognitive Biases

Provocation Paper at


Should we introduce a universal basic income?
That is a very controversial topic! I found nine reasons: five in favor and four against a universal basic income. Which side should I start with?
I expected disagreement... Can you first give me the cons?
The main reasons against a universal basic income are that it, first, erodes incentives for financial responsibility and hard work, second, has negative effects on the national economy, and third, has superior alternative policies. Should I elaborate on one reason, list the remaining con, or list pros?

Conversational Argument Search

Large discrepancy between:

- Amount of information an intelligent assistant can convey
- Exploration a complex/controversial topic demands

- Decision made by the system to *expose* information to the user via a speech-only communication channel may create or reinforce unintended cognitive bias

Bias in Conversational Information Access

- Bias in Search (e.g., Azzopardi CHIIR'21)
- Bias in Conversational Systems and Personalized Knowledge Graphs (e.g., Gerritse et al. ICTIR'20)

  Strategies: ignore, tell the user, or provide options to counterweigh/diversify

- Priming, anchoring, framing, availability bias…

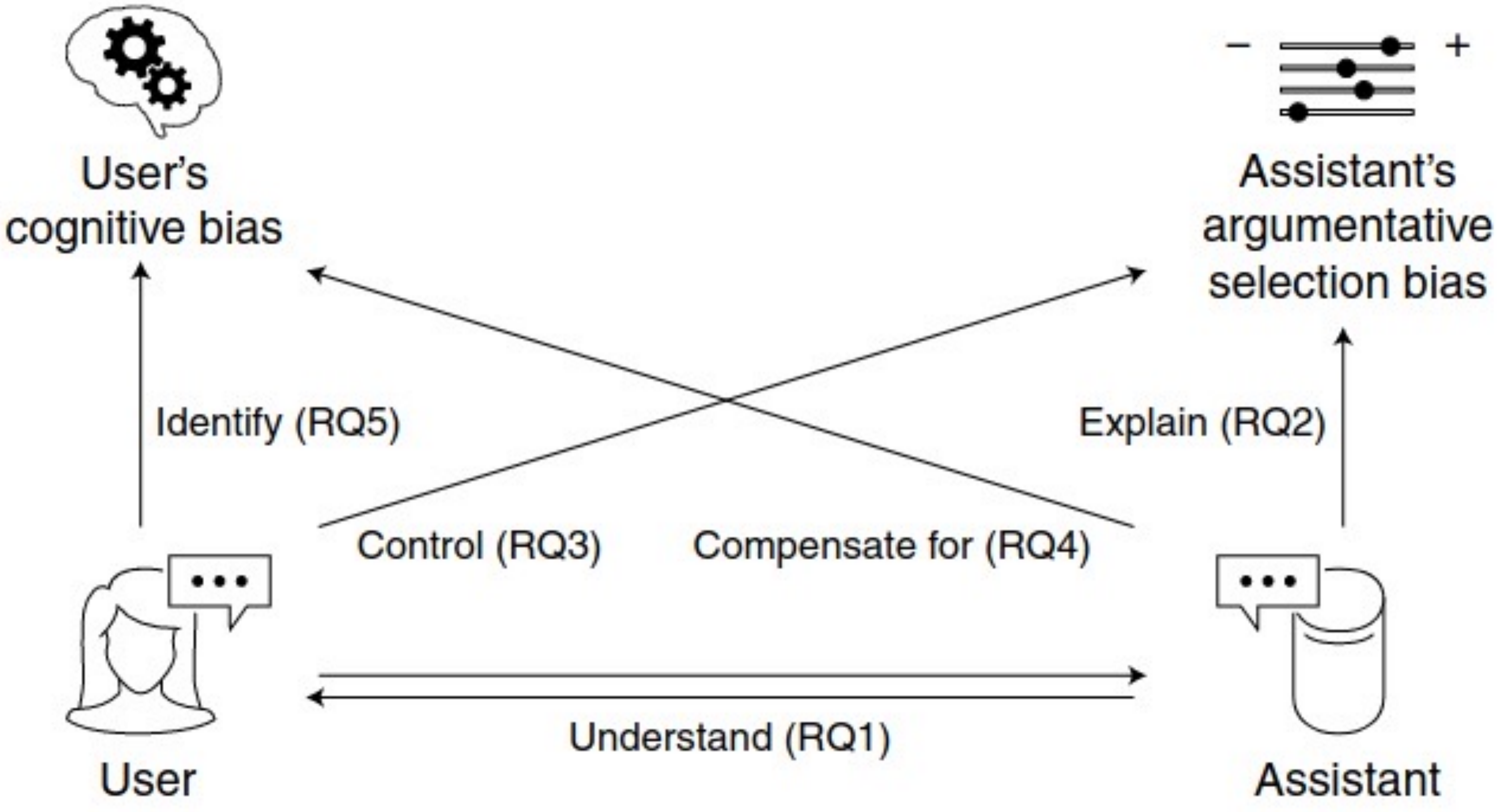


Figure 3: Overview of the proposed research questions. Read arrows as "How can the <entity 1> <verb> the <entity 2>?"

Research Agenda

RQ	Action Step
1. How can the user and the assistant understand each other?	
	- Investigate on short/long-term effects and mental models
	- Develop privacy-aware interaction guidelines
2. How can the assistant explain its argumentative selection bias?	
	- Identify intuitively understandable bias categories
	- Investigate how to make bias explicit
3. How can the user control the assistant's argumentative selection bias?	
	- Identify cue phrases that specify argumentative selection biases
	- Investigate on personas for different argumentative selection biases
4. How can the assistant compensate for the user's cognitive biases?	
	- Investigate strategies to encourage users to explore
	- Identify conversation styles that least provoke cognitive biases
5. How can the assistant help the user identify their cognitive biases?	
	- Identify hints at the application of cognitive biases
	- Identify strategies to explain the users their cognitive biases

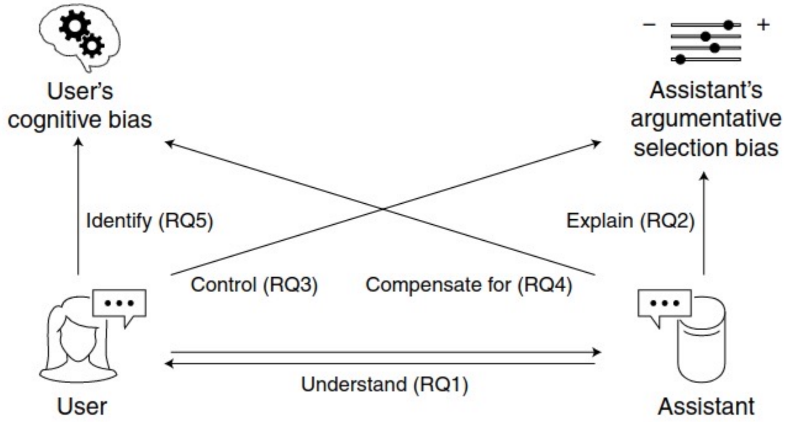
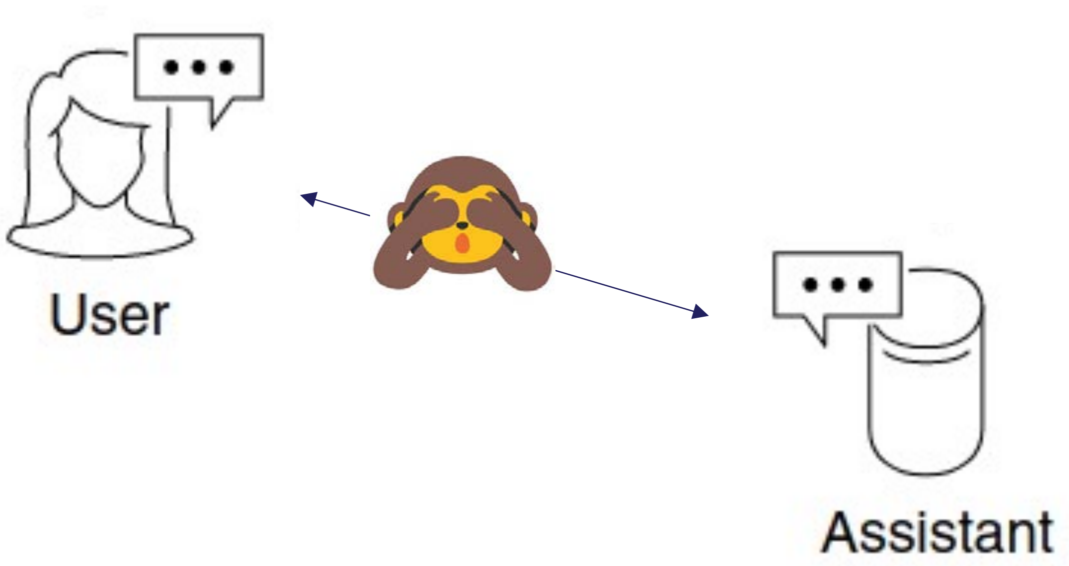


Figure 3: Overview of the proposed research questions. Read arrows as "How can the <entity 1> <verb> the <entity 2>?"

“Ok Google/Alexa/Siri, Is Australia outperforming other countries with its coronavirus vaccination rollout?”



Diversity

and

Fairness

from a

Ranking Evaluation

perspective

Dank Je Wel!

Diversity and Fairness From a Ranking Evaluation Perspective

Damiano Spina

@damiano10

damiano.spina@rmit.edu.au

www.damianospina.com

What's next...

ADCS 2021

The 25th Australasian Document Computing Symposium

Virtual Seminar Series

http://adcs-conference.org/2021/ @AusDocCompSymp Register here: https://bit.ly/ADCS2021Registration

December 9, 2021:
Conference Opening
Keynote 1: *What users tell us: User Understanding and Modeling for Personalized Recommendation*
Assoc. Prof. Min Zhang, Tsinghua University

January 13, 2022:
Hierarchical Clustering of Corals using Image Clustering - R. Sawant, G. Demartini, and T. Bridge
Cost-Effective Updating of Distributed Reordered Indexes - J. Mackenzie and A. Moffat

January 18, 2022:
Document Clustering vs Topic Models: A Case Study - M. Yuan, P. Lin, and J. Zobel
MeSH Term Suggestion for Systematic Review Literature Search - S. Wang, H. Li, H. Scells, D. Locke, and G. Zuccon

January 20, 2022:
Annotation of Struck-out Text in Handwritten Documents - H. Nisa, V. Ciesielski, J. Thom, and R. Tennakoon
Cohort-based Clinical Trial Retrieval - B. Koopman and G. Zuccon

February 3, 2022:
Passage Based Answer-Set Graph Approach for Query Performance Prediction - G. Sarwar and C. O'Riordan
Crowdsourcing Backstories for Complex Task-Based Search - M. Steiner, D. Spina, F. Scholer, and L. Cavedon

February 11, 2022:
An Analysis of the Australian Political Discourse in Sponsored Social Media Content - L. Han, R. Sawant, S. Fan, G. Kefford, and G. Demartini
Business Meeting and Social Event

February 24, 2022:
Keynote 2: *Maximising the potential of documents, NLP, and AI in an online employment marketplace: Technical and non-technical challenges* - Dr. Terrence Szymanski, SEEK
Conference Closing

Register here

RMIT UNIVERSITY seek THE UNIVERSITY OF MELBOURNE