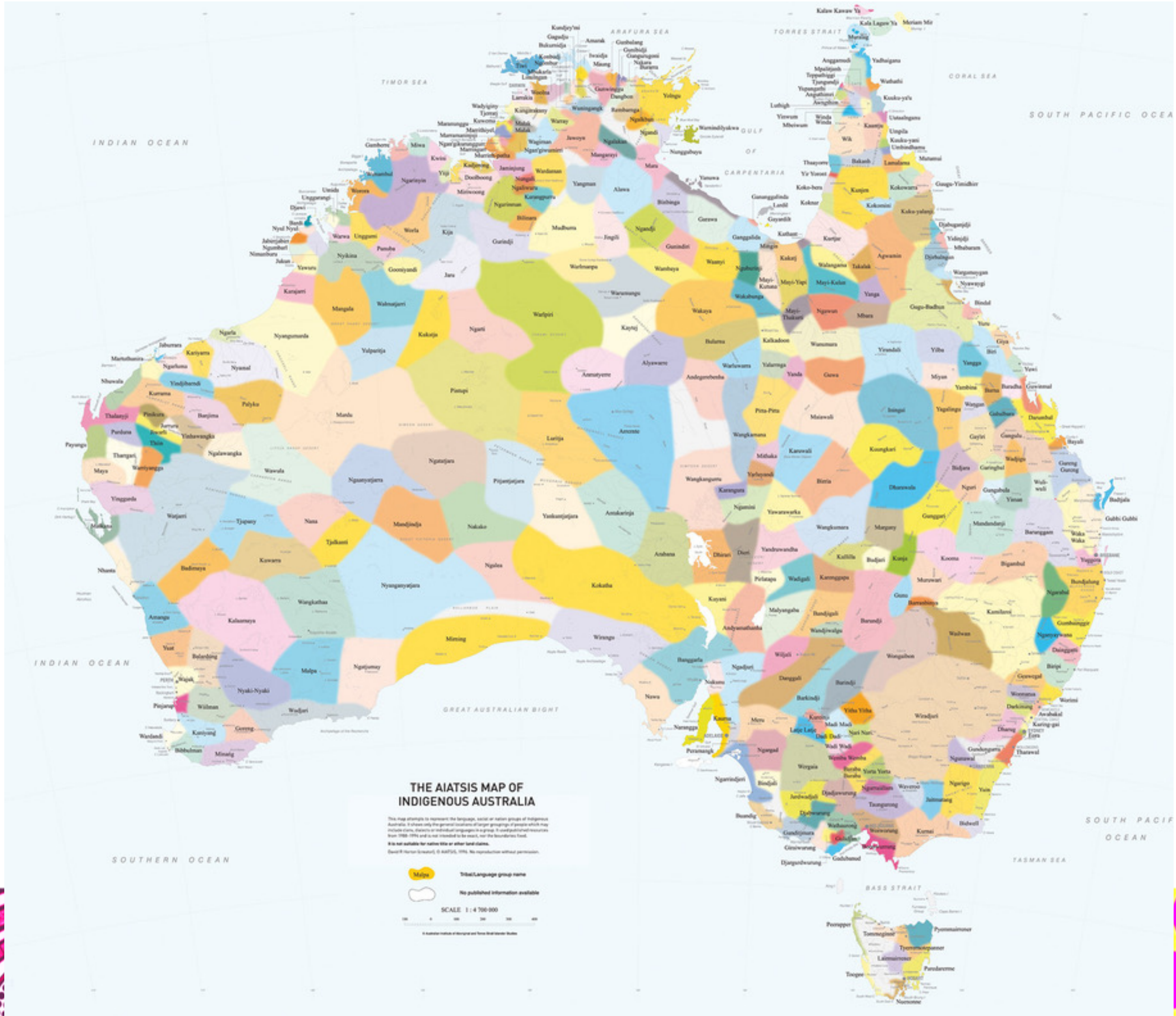
Diversity and Fairness From a Ranking Evaluation Perspective

Damiano Spina

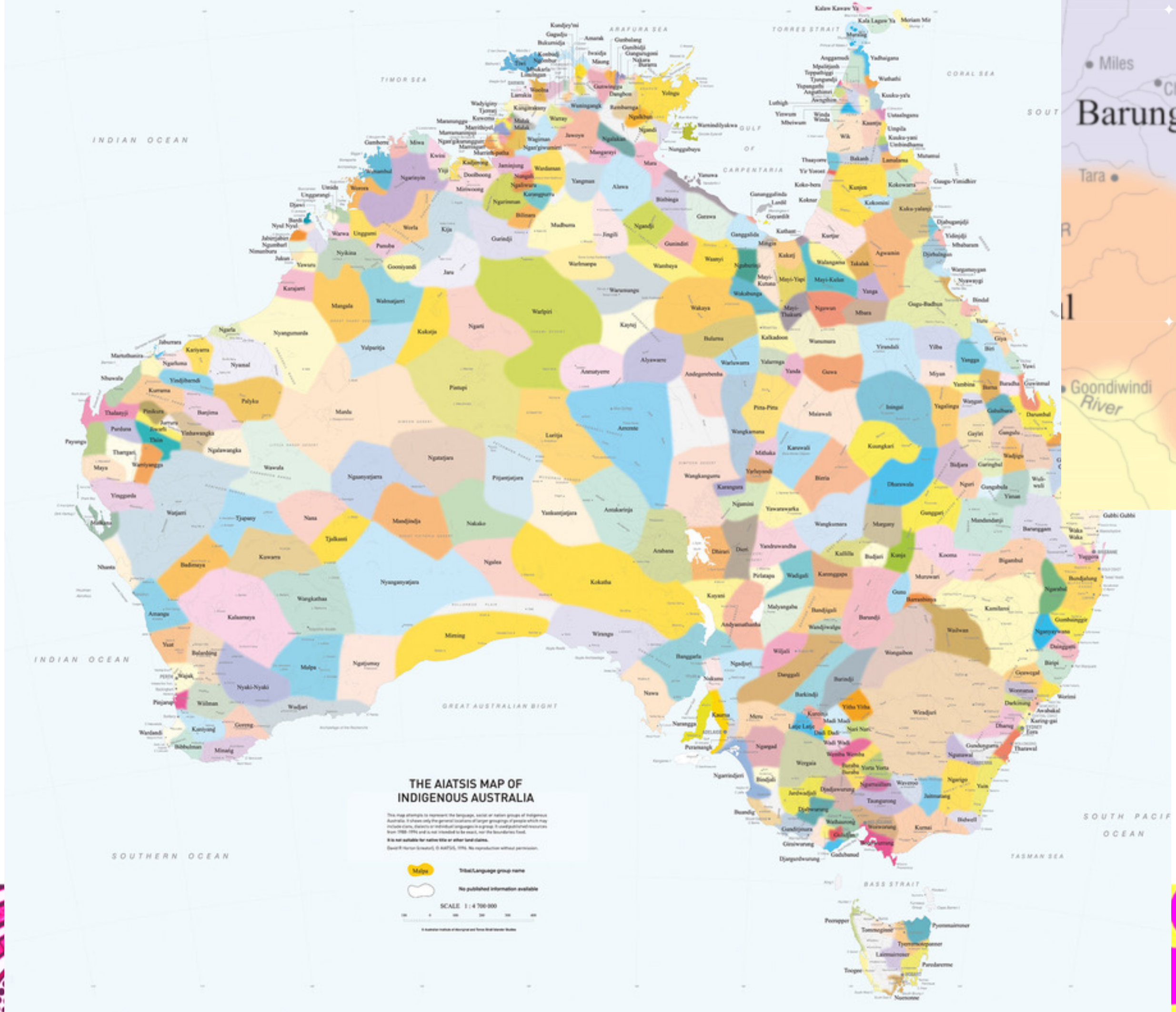
@damiano10

damiano.spina@rmit.edu.au

www.damianospina.com



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

Acknowledgement of Country

I would like to acknowledge the people of the Yuggera and Turrbal people whose unceded lands I visit today.

I respectfully acknowledge their Ancestors and Elders past and present.

I also acknowledge the Traditional Custodians and their Ancestors of the lands and waters across Australia.

About This Work

Collaborations with:

Enrique Amigó
Jorge Carrillo-de-Albornoz
Sachin P. Cherumanal
W. Bruce Croft
Johannes Kiesel
Stefano Mizzaro
Falk Scholer
Benno Stein
Henning Wachsmuth
…

About This Work

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 **3100 9957**

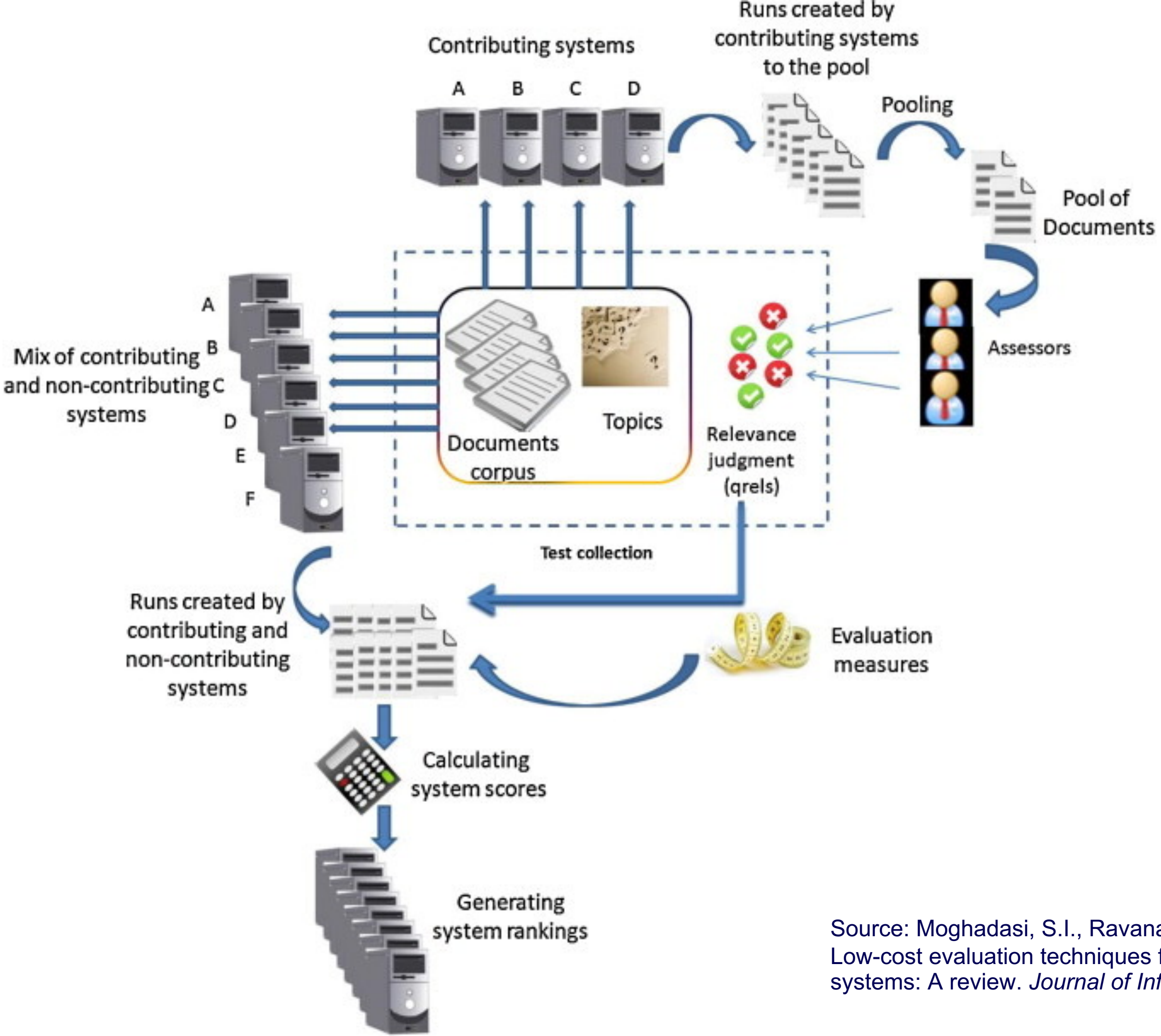
https://www.menti.com/m1173zzscx

Go to **www.menti.com** and use the code **3100 9957**

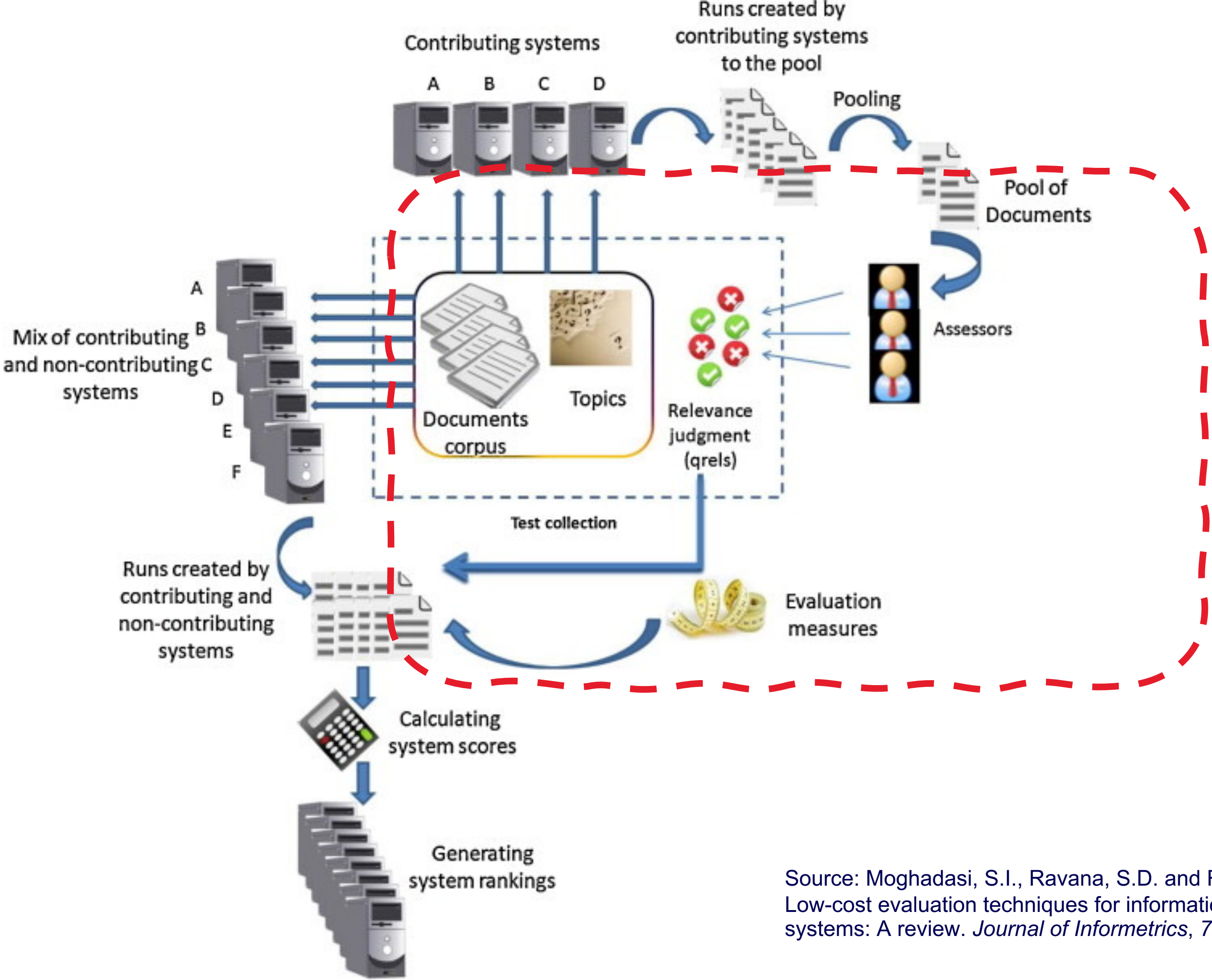
Mentimeter

List elements used to characterise IR evaluation

20



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 make a document relevant to a topic?

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

What does 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 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?

Go to **www.menti.com** and use the code **3100 9957**

Effectiveness evaluation metrics for ad-hoc retrieval

Mentimeter

user satisfaction

dcg

err

ap

rbu

ndcg

map

rbp

mrr

rr

6

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

User's information need → Topic

Intent

Topics, Intents, Queries, and Aspects

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

Topics, Intents, Queries, and Aspects

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

SIGIR'22 Perspectives Paper:

Where Do Queries Come From?

Marwah Alaofi
RMIT University
Melbourne, Australia

Luke Gallagher
RMIT University
Melbourne, Australia

Dana McKay
RMIT University
Melbourne, Australia

Lauren L. Saling
RMIT University
Melbourne, Australia

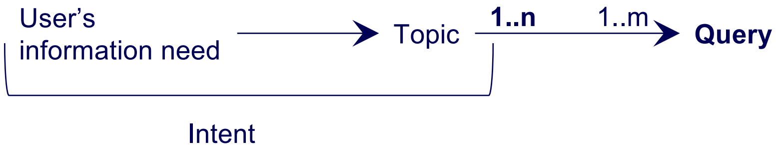
Mark Sanderson
RMIT University
Melbourne, Australia

Falk Scholer
RMIT University
Melbourne, Australia

Damiano Spina
RMIT University
Melbourne, Australia

Ryen W. White
Microsoft Research
Redmond, WA, USA

Topics, Intents, Queries, and Aspects



SIGIR'22 Perspectives Paper:

Where Do Queries Come From?

Marwah Alaofi
RMIT University
Melbourne, Australia

Luke Gallagher
RMIT University
Melbourne, Australia

Dana McKay
RMIT University
Melbourne, Australia

Lauren L. Saling
RMIT University
Melbourne, Australia

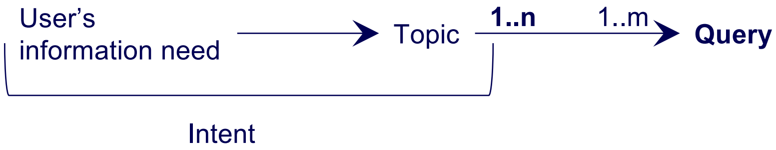
Mark Sanderson
RMIT University
Melbourne, Australia

Falk Scholer
RMIT University
Melbourne, Australia

Damiano Spina
RMIT University
Melbourne, Australia

Ryen W. White
Microsoft Research
Redmond, WA, USA

Topics, Intents, Queries, and Aspects



SIGIR'22 Perspectives Paper:

Where Do Queries Come From?

Marwah Alaofi
RMIT University
Melbourne, Australia

Luke Gallagher
RMIT University
Melbourne, Australia

Dana McKay
RMIT University
Melbourne, Australia

Lauren L. Saling
RMIT University
Melbourne, Australia

Mark Sanderson
RMIT University
Melbourne, Australia

Falk Scholer
RMIT University
Melbourne, Australia

Damiano Spina
RMIT University
Melbourne, Australia

Ryen W. White
Microsoft Research
Redmond, WA, USA

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
```

CiDDA

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

Document Id

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
```

CiDDA

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
```

```
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
```

```
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 **3100 9957**

https://www.menti.com/m1173zzscx

Go to **www.menti.com** and use the code **3100 9957**

Effectiveness evaluation metrics for diversity

Mentimeter

allha ndch

s-recall

alpha-ndcg

irbu

err-ia

err

6

How Do We Pick The Right Metric?

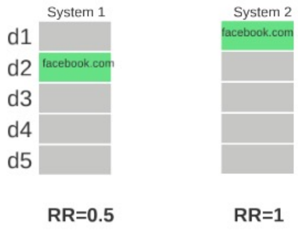
Example

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

query: facebook

Reciprocal Rank (RR)

P@3?



How Do We Pick The Right Metric?

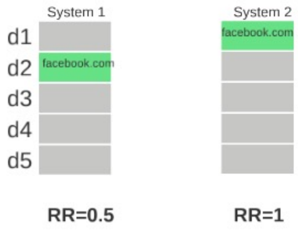
Example

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

query: facebook

Reciprocal Rank (RR)

P@3?



and for **Search Result Diversification**?

query: things to do in Meanjin

query aspects:{ , , , }

How Do We Pick The Right Metric?

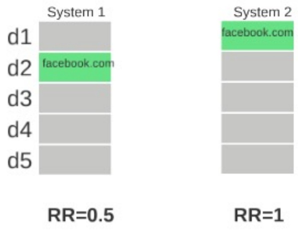
Example

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

query: facebook

Reciprocal Rank (RR)

P@3?



and for **Search Result Diversification**?

***query: things to do in* Meanjin**

How Do We Pick The Right Metric?

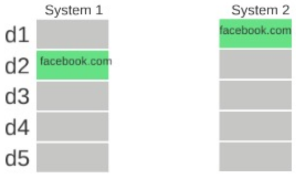
Example

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

query: facebook

Reciprocal Rank (RR)

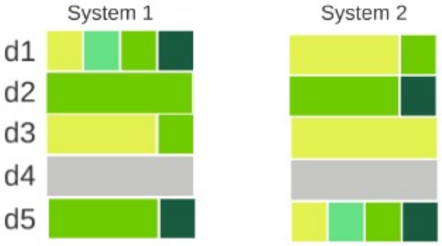
P@3?



RR=0.5 **RR=1**

and for **Search Result Diversification**?

query: things to do in Meanjin



Which Metric Would You Pick?

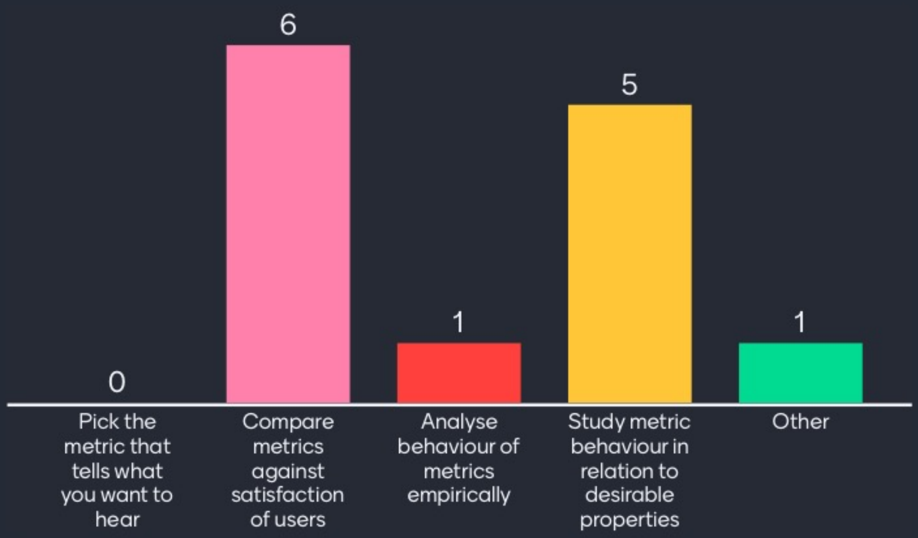
Go to **www.menti.com** and use the code **3100 9957**

https://www.menti.com/m1173zzscx

Go to **www.menti.com** and use the code **3100 9957**

Which approach would you follow to pick the right evaluation measure?

Mentimeter



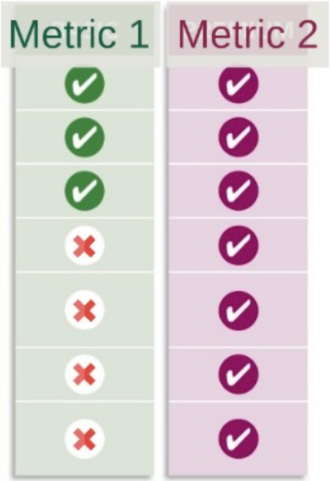
13

Axiomatic Analysis

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

Axiomatic Analysis

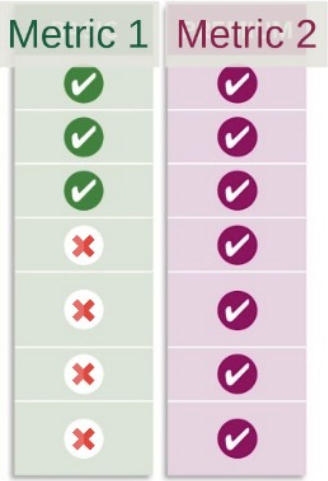
We may not be able to prescribe how to design the evaluation metric for our task...



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)

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

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

CiDDA

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	●	●	●	●	●	●	●	●	●	●

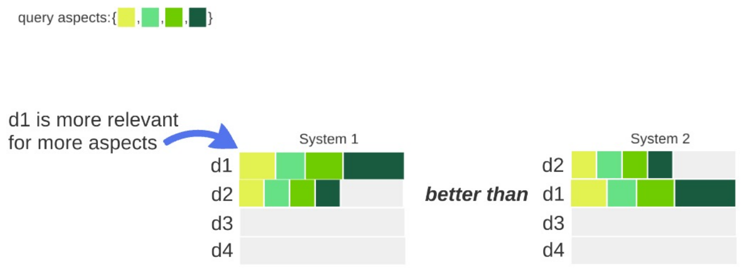
NEW!

	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	●	●	●	○	○	●	●	○	●	●
NEW! 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

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

patience parameter (RBP)

effort (EU, Expected Utility)

ranking $\vec{d}$

k = <u>$\mathit{min}(\mathit{cutoff}, |\vec{d}|)$</u>

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/

Ranking Interruptus: When Truncated Rankings Are Better and How to Measure That

Enrique Amigó
UNED NLP & IR Group
Madrid, Spain
enrique@lsi.uned.es

Stefano Mizzaro
University of Udine
Udine, Italy
mizzaro@uniud.it

Damiano Spina
RMIT University
Melbourne, Australia
damiano.spina@rmit.edu.au

Ranking Interruptus: When Truncated Rankings Are Better and How to Measure That

		Non-Truncated				*Truncated*		
	Metric	Priority	Top-W.	Deep. Th.	Shall. Th.	Confidence	Recall	Redundancy
Full ranking metrics	Spearman	✓			✓	✓		
	Kendall	✓			✓	✓		
	AUC-ROC	✓			✓	✓		
Binary relevance	P@*N*			✓	✓			
	R@*N*			✓	✓		✓	
	R-p						✓	
	RR							✓
	F-measure					✓	✓	✓
	AP	✓	✓				✓	
Graded relevance	NDCG	✓	✓		✓		✓	
	Q-measure	✓	✓		✓		✓	
	BPref	✓	✓		✓		✓	
Probabilistic user model based	ERR	✓	✓	✓				✓
	RBP	✓	✓	✓	✓			
	iRBU	✓	✓	✓	✓			✓
Terminal document based	NDCGT	✓	✓		✓	✓	✓	
	ERRT	✓	✓	✓		✓		✓
	RBPT	✓	✓	✓	✓	✓		
Utility based	Flat Utility	✓			✓	✓		
	RBU	✓	✓	✓	✓	✓		✓
	DCGU	✓	✓		✓	✓		
	ERRU	✓	✓	✓		✓		✓
	RBPU	✓	✓	✓	✓	✓		
Information based	OIE	✓	✓	✓	✓	✓	✓	✓

Stefano Mizzaro
University of Udine
Udine, Italy
mizzaro@uniud.it

Damiano Spina
RMIT University
Melbourne, Australia
damiano.spina@rmit.edu.au

Ranking Interruptus: When Truncated Rankings Are Better and How to Measure That

	Metric	Priority	Top-W.	Deep. Th.	Shall. Th.	Confidence	Recall	Redundancy
		Non-Truncated				*Truncated*		
FULL	Spearman	✓			✓	✓		
RANKING	Kendall	✓			✓	✓		
METRICS	AUC-ROC	✓			✓	✓		
BINARY	P@N			✓	✓			
RELEVANCE	R@N			✓	✓		✓	
	R-p						✓	
	RR							✓
	F-measure					✓	✓	✓
	AP	✓	✓				✓	
GRADED	NDCG	✓	✓		✓		✓	
RELEVANCE	Q-measure	✓	✓		✓		✓	
	BPref	✓	✓		✓		✓	
PROBABILISTIC	ERR	✓	✓	✓				✓
USER MODEL	RBP	✓	✓	✓	✓			
BASED	iRBU	✓	✓	✓	✓			✓
TERMINAL	NDCGT	✓	✓		✓	✓	✓	
DOCUMENT	ERRT	✓	✓	✓		✓		✓
BASED	RBPT	✓	✓	✓	✓	✓		
UTILITY	Flat Utility	✓			✓	✓		
BASED	RBU	✓	✓	✓	✓	✓		✓
	DCGU	✓	✓		✓	✓		
	ERRU	✓	✓	✓		✓		✓
	RBPU	✓	✓	✓	✓	✓		
INFORMATION BASED	OIE	✓	✓	✓	✓	✓	✓	✓

Stefano M
University
Udine,
mizzaro@

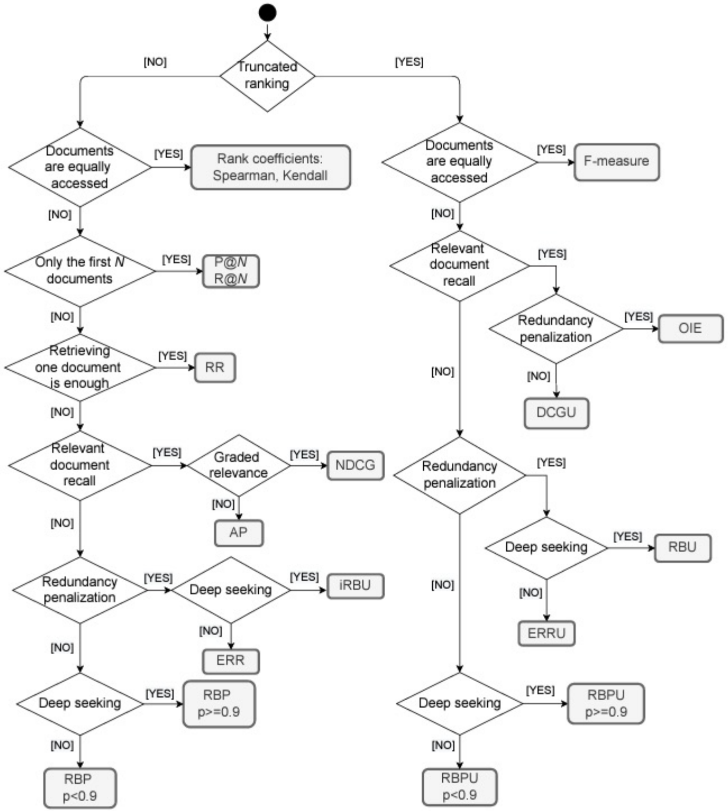


Figure 4: Workflow for metric selection in different ranking problems.

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 **3100 9957**

https://www.menti.com/m1173zzscx

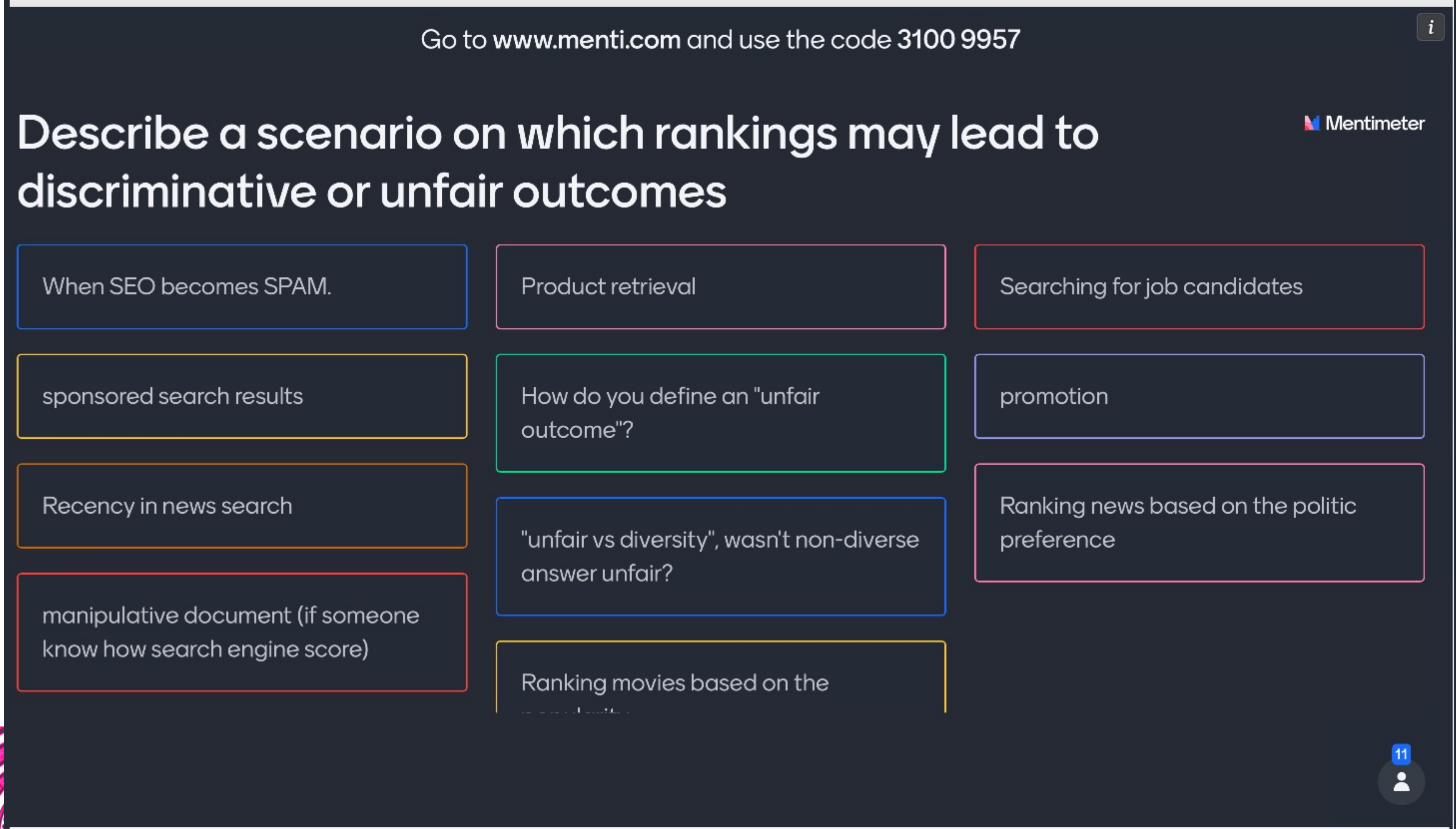

Djoerd Hiemstra
Research, Teaching and More
Home Contact Publications

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/


Go to www.menti.com and use the code 3100 9957
Describe a scenario on which rankings may lead to discriminative or unfair outcomes
Mentimeter
When SEO becomes SPAM.
Product retrieval
Searching for job candidates
sponsored search results
How do you define an "unfair outcome"?
promotion
Recency in news search
"unfair vs diversity", wasn't non-diverse answer unfair?
Ranking news based on the politic preference
manipulative document (if someone know how search engine score)
Ranking movies based on the
11

Fairness

Fairness

Individual Fairness

Similar individuals should be treated similarly.

Talent search: 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.

Talent search: 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.

Talent search: 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

CIDDA

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 ▾

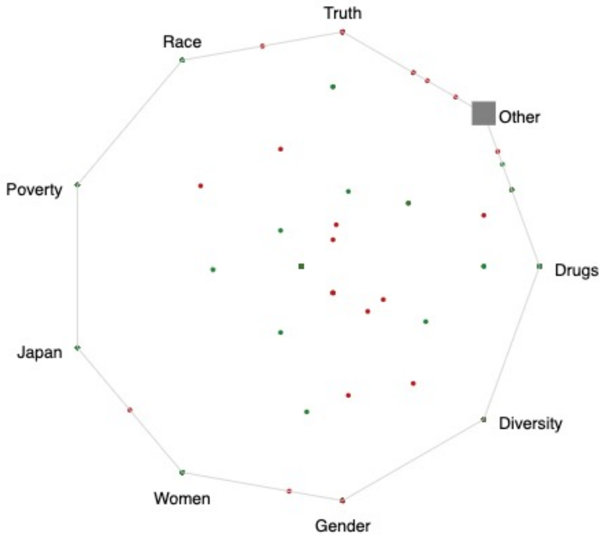
PRO I understand that there is always a good chance that...

▶ 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 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

(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 Ratio (**rRD**)
- 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

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

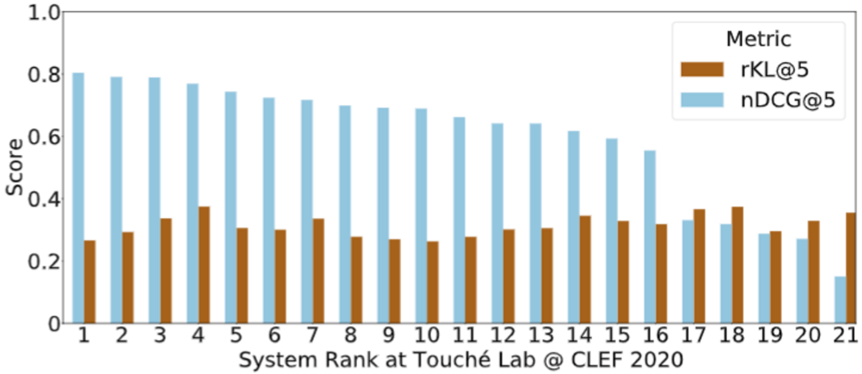
- System runs from Touché Lab at CLEF 2020.
- Relevance (NDCG), (Un)fairness metrics, and diversity (α-NDCG)

- **Test Collection:** Ground truth of 2,964 relevance judgments across 49 topics.
- **Relevance:** Graded between 1 (least relevant) and 5 (most relevant).

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

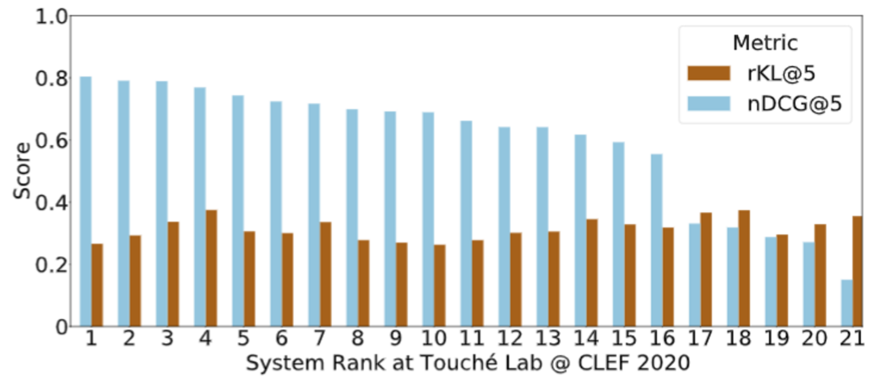
Topic: Universal Basic Income	
Argument	**Stance**
...lifting society out of poverty….	PRO
UBI is individually destructive ...	CON

Results



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

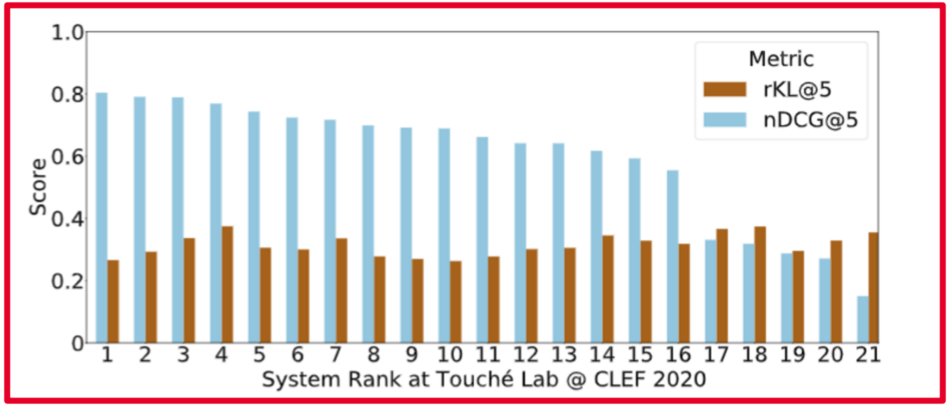
Results



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

Results

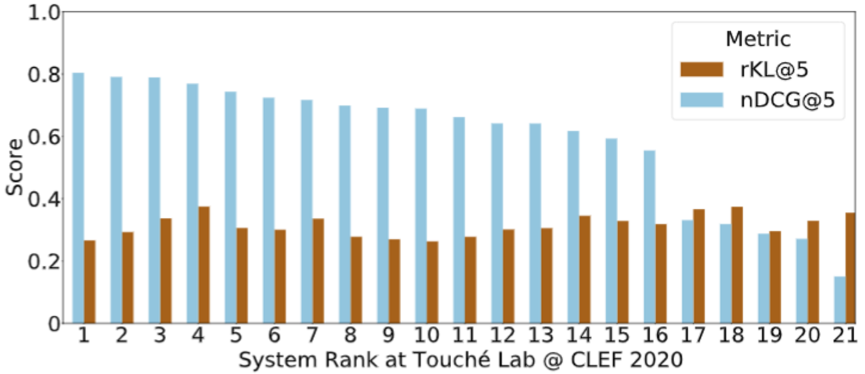
- (Un)fairness metrics do not increase monotonically w.r.t. NDCG@5



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

Results

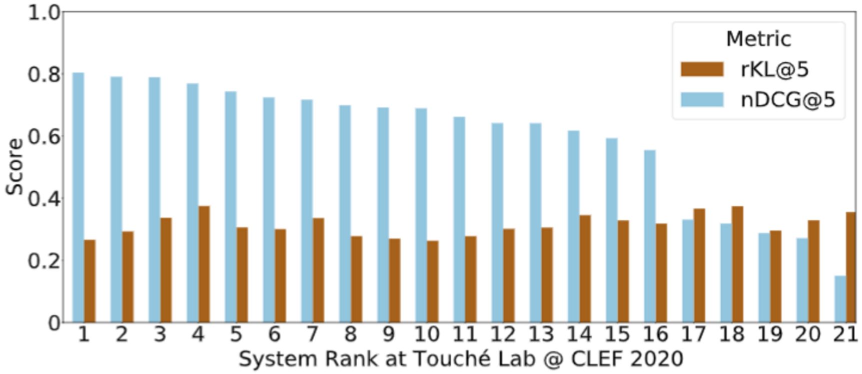
- (Un)fairness metrics do not increase monotonically w.r.t. NDCG@5
- System ranks would change when ranked using both (un)fairness and relevance.



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.

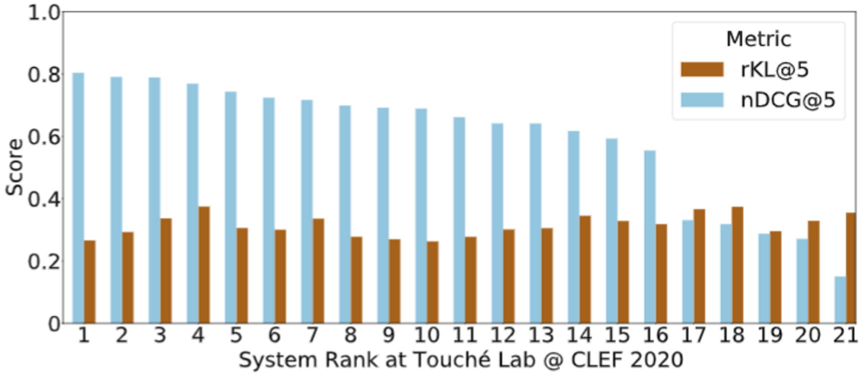


For nDCG@5 higher is better and for (un)fairness metric rKL@5, lower is better.

CiDDA

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.

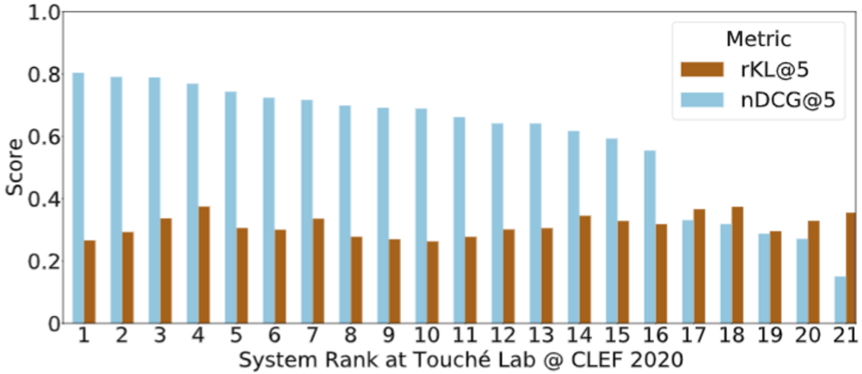


For nDCG@5 higher is better and for (un)fairness metric rKL@5, lower is better.

(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.

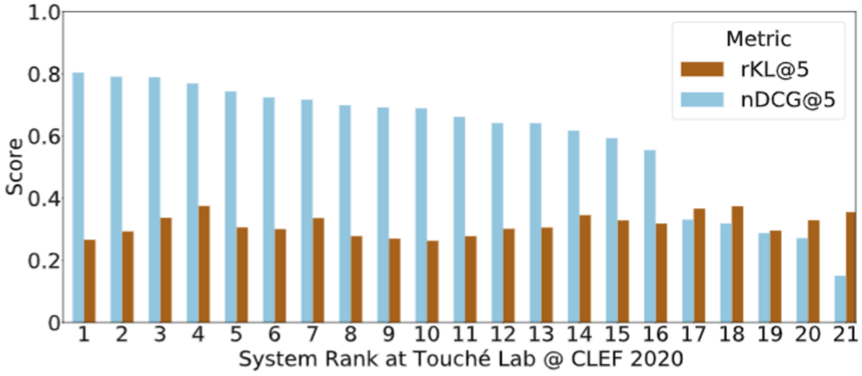


For nDCG@5 higher is better and for (un)fairness metric rKL@5, lower is better.

(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.



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

- Diversity is related but not equivalent to (un)fairness.

(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

Diversity

and

Fairness

from a

Ranking Evaluation

perspective

Diversity and Fairness From a Ranking Evaluation Perspective

Damiano Spina

@damiano10

damiano.spina@rmit.edu.au

www.damianospina.com

RMIT UNIVERSITY

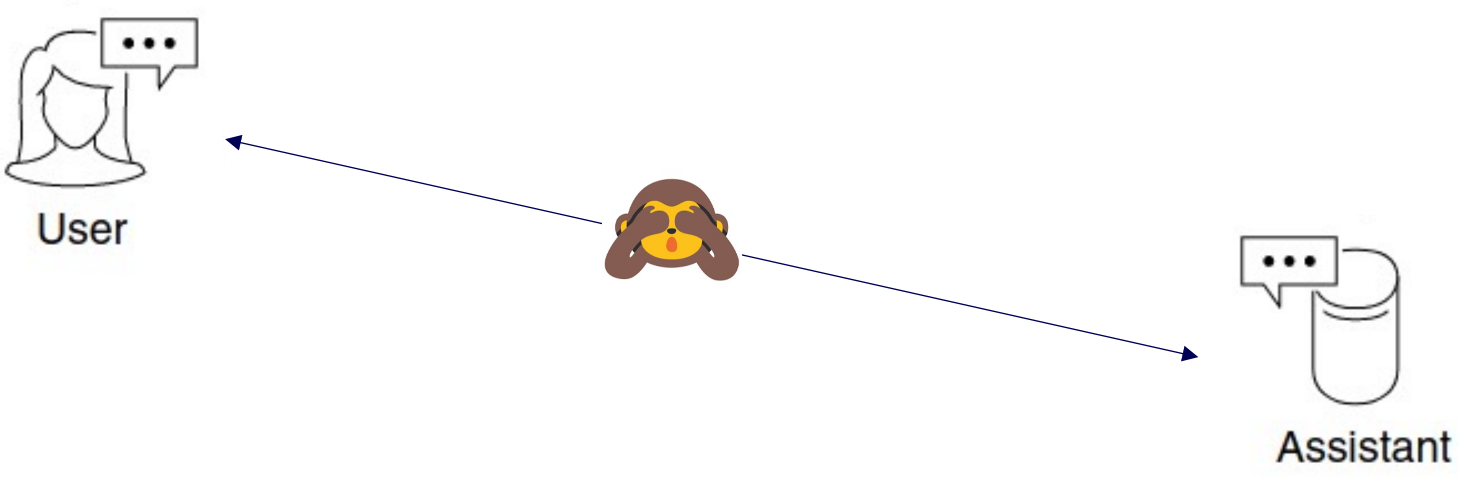
CiDDA

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

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.

CIDDA

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

CiDDA

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

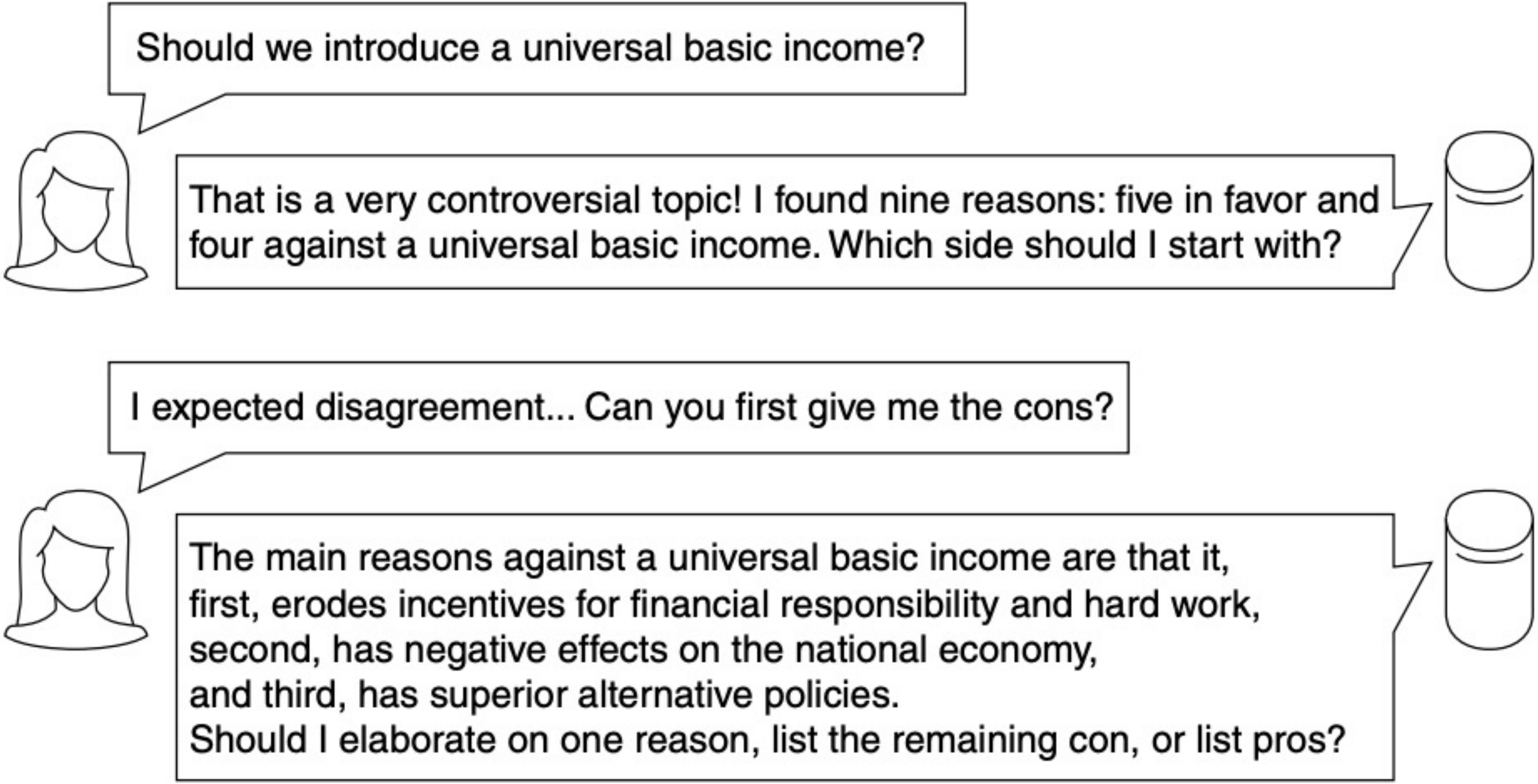
=

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

Provocation Paper at

CUI 2021
Conversational User Interfaces
27th–29th July 2021
Online virtual conference

CiDDA


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

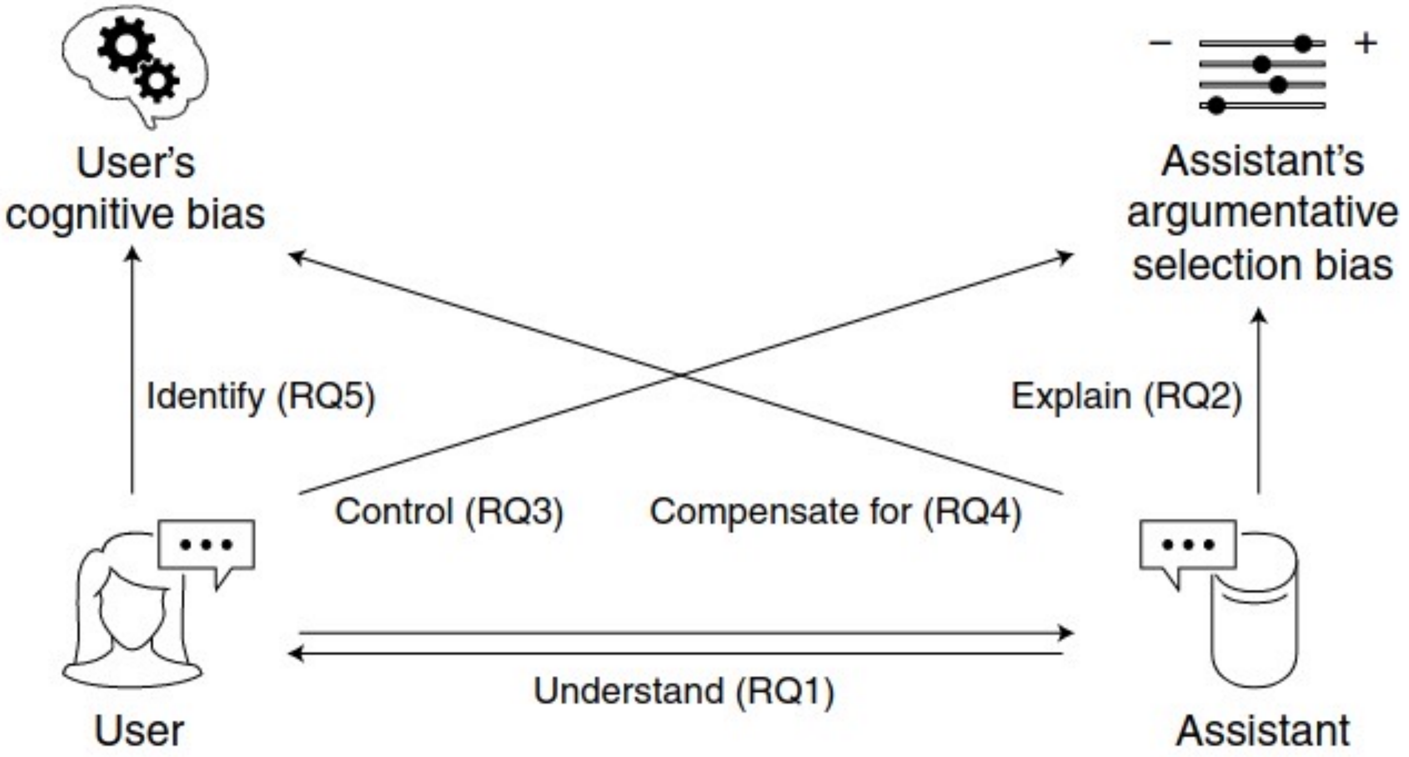


Figure 3: Overview of the proposed research questions. Read arrows as "How can the <entity 1> <verb> the <entity 2>?"

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

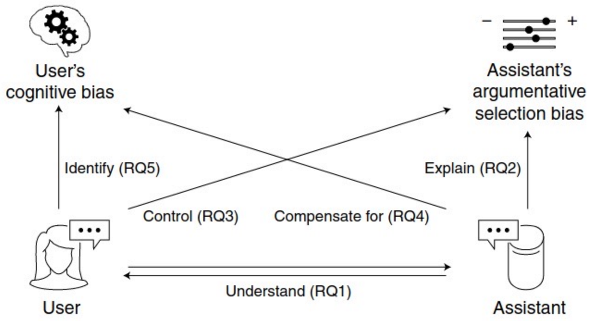
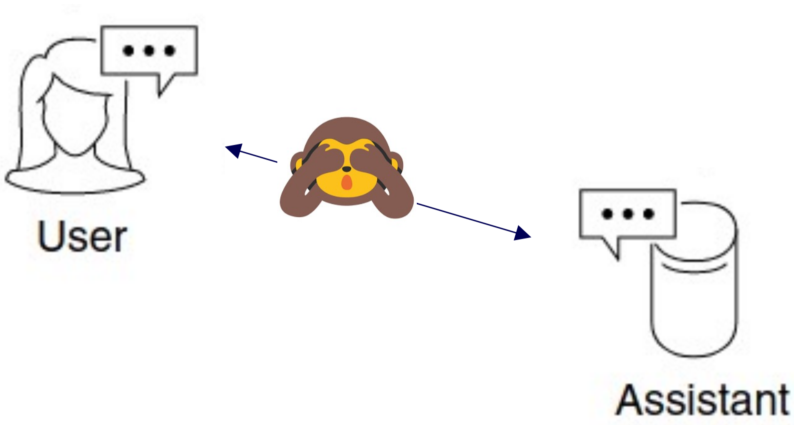


Figure 3: Overview of the proposed research questions. Read arrows as "How can the <entity 1> <verb> the <entity 2>?"

RMIT UNIVERSITY CIDDA

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



FACT CHECK

Scott Morrison claimed Australia was outperforming other countries with its coronavirus vaccination rollout. Was he correct?

RMIT ABC Fact Check

Posted Fri 16 Apr 2021 at 6:20am, updated Sun 2 May 2021 at 9:05pm

Misleading