

# Search engine consensus

|Aditi Narvekar, Andrea Chang, Felix Park, Nikhil Bhaip



Google Search

I'm Feeling Lucky

# Goal: an exploratory analysis

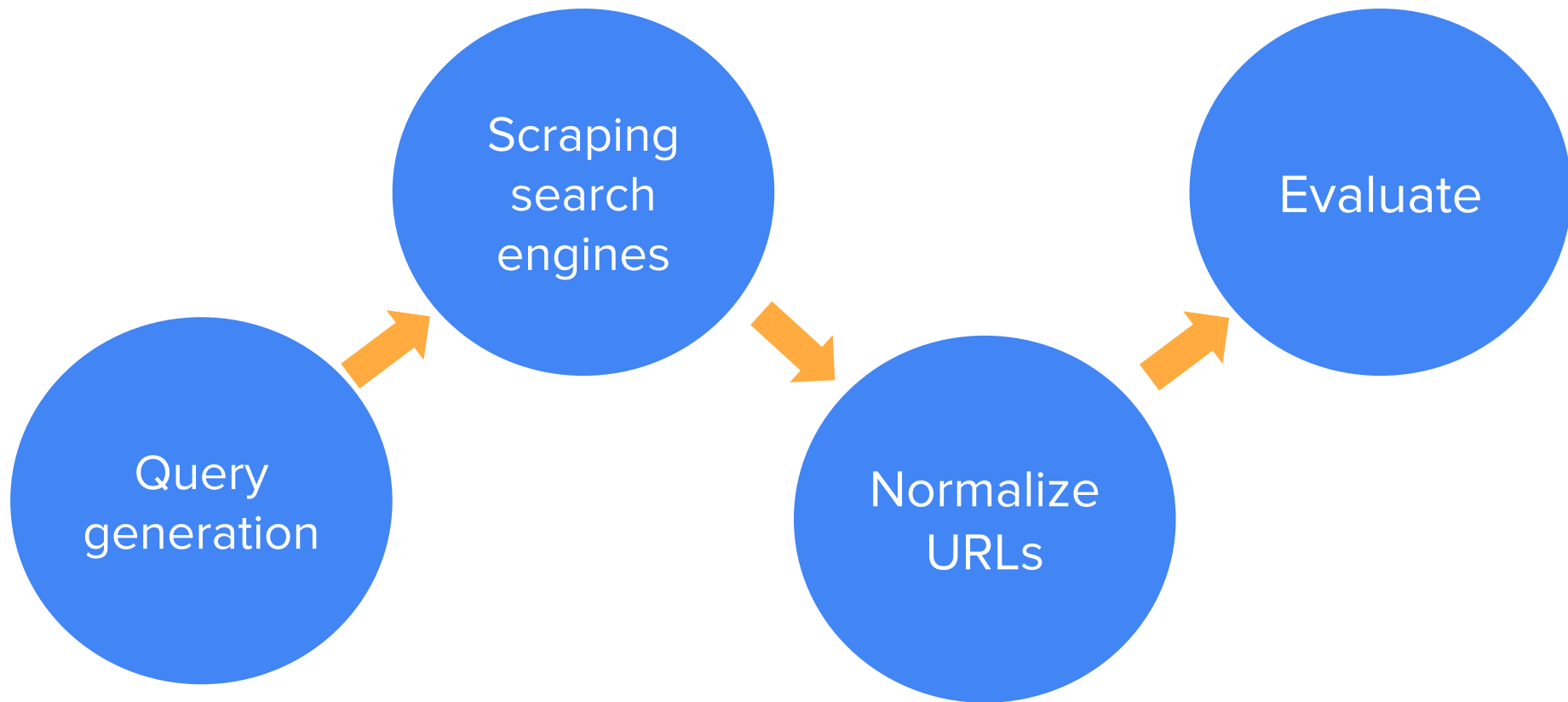
- Determine which/where search engines agree and disagree
- Do search engines have different ideas of relevance?



# Research questions

1. Which search engines have the most consensus?
2. Which search engines have the most disagreement?
3. On which topics do the above search engines agree/disagree?

# Methodology



# Query generation

## US History Topics

Quizlet.com

Reveals differences in educational systems

## Abstract Ideas

Self-constructed

Reveals differences in philosophy and literature

## Religious Topics

<http://www.religioustolerance.org/>

Reveals differences in cultural views of religion

## Science Questions

Self-constructed

Avoid dictionary results for well-defined topics

## Current Events Topics

<https://trends.google.com/trends/>

Reveal differences in how users get their news

## Generic English Topics

<https://www.kaggle.com/rtatman/english-word-frequency>

Reveal differences in how neutral terms are treated

# Scraping search engines

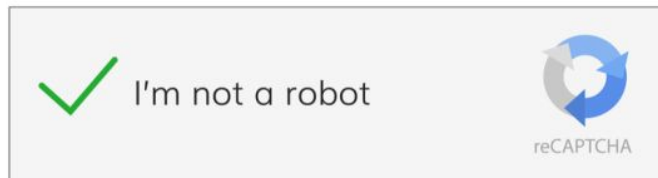
- Scraper API

- Handles proxies, browsers, and CAPTCHAs



- GoogleScraper

- Selenium browser emulator to avoid bot detection



# Normalizing URLs

```
"Taoism": [  
  {  
    "url": "https://en.wikipedia.org/wiki/Taoism",  
    "title": "Taoism - Wikipedia"  
  },  
  {  
    "url": "https://personaltao.com/taoism/taoism-101/",  
    "title": "Taoism 101: Introduction to the Tao and What is Taoism"  
  },  
  {  
    "url": "http://www.bbc.co.uk/religion/religions/taoism/",  
    "title": "BBC - Religion: Taoism"  
  },  
  {  
    "url": "https://www.ancient.eu/Taoism/",  
    "title": "Taoism - Ancient History Encyclopedia"  
  },  
  {  
    "url": "http://www.crystalinks.com/taoism.html",  
    "title": "Taoism - Crystalinks"  
  },  
]
```

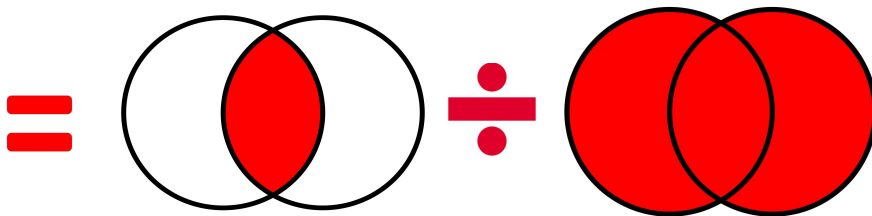


[  
 wikipedia.org,  
 personaltao.com,  
 bbc.co.uk,  
 ancient.eu,  
 Crystalinks.com  
]

# Evaluation Methods

- Jaccard Correlation

$$J(L_1, L_2) = \frac{|L_1 \cap L_2|}{|L_1 \cup L_2|}$$



- Top K symmetric difference

$$SD(L_1, L_2) = \sum_{n=1}^k \frac{1}{2k} (|L_1 \cup L_2| - |L_1 \cap L_2|)$$





# Modern evaluation methods

- Rank biased overlap
  - Incomplete sets
  - Weights higher rankings more

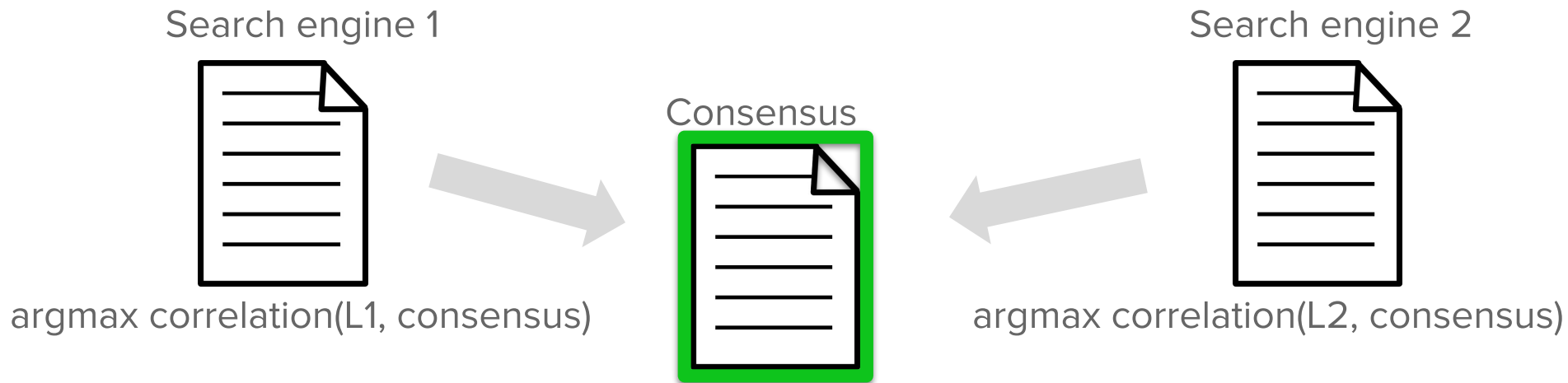
$$RBO(S, T, p) = (1 - p) \sum_{d=1}^{\infty} (p^{d-1}) \frac{I_{S, T, d}}{d}$$

- AnchorMAP
  - Assume 1 list as ground truth and compute MAP
  - Proposed a modified version

$$Anchor'(L_1, L_2) = \frac{Anchor(L_1, L_2) + Anchor(L_2, L_1)}{2}$$

# Proposed method for consensus ranking

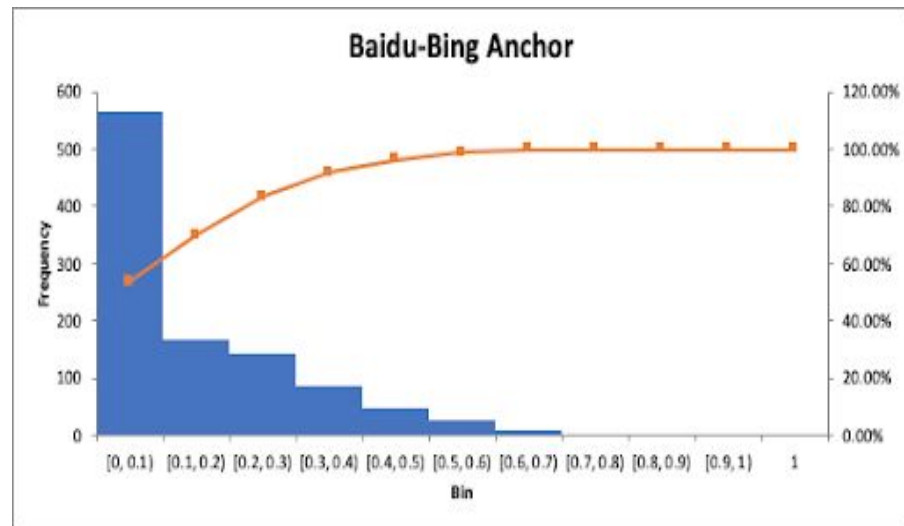
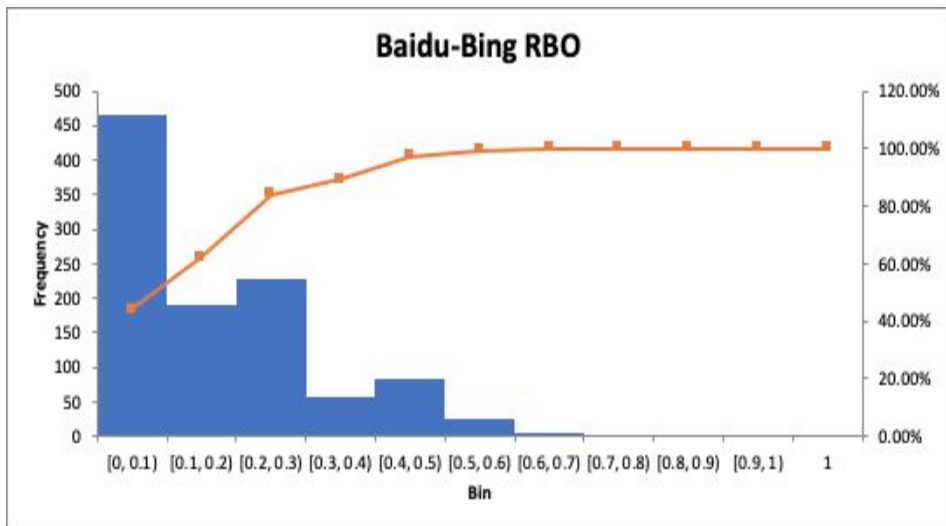
- Determine a consensus rank that would maximize the correlation between results



# Results

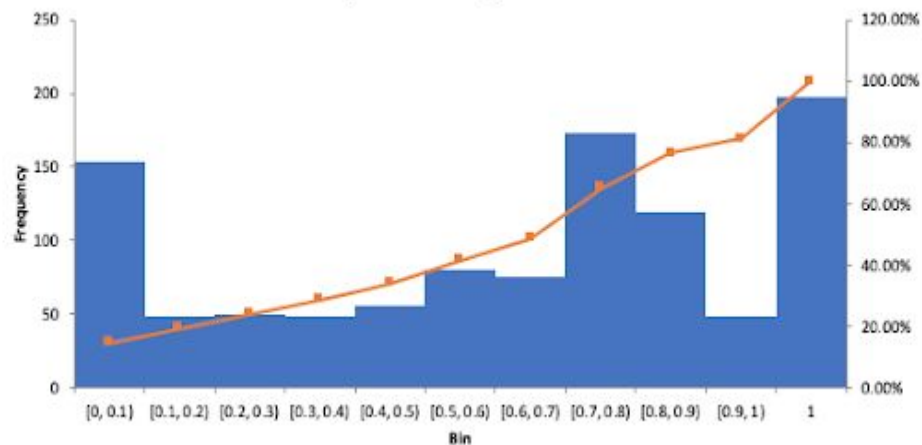
- Bing and DuckDuckGo had the most consensus
- Baidu and Bing have the most disagreement
- Topics were independent from correlation results

# Baidu and Bing had the most disagreement

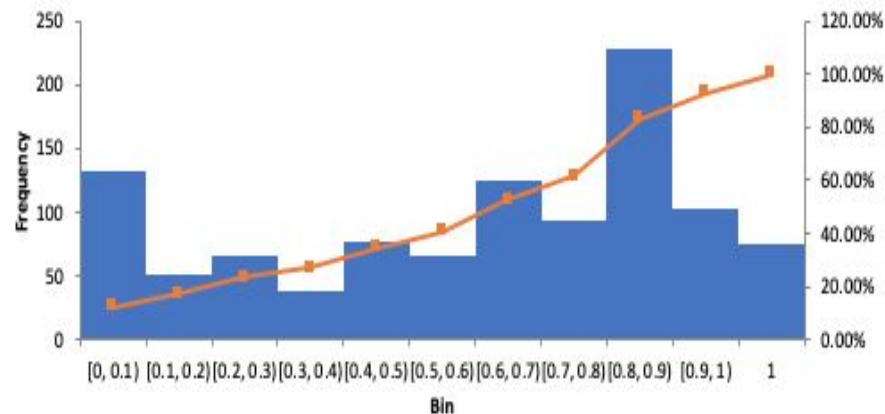


# Bing and DuckDuckGo had the most consensus

Bing-Duckduckgo Anchor



Bing-Duckduckgo RBO



# Statistics

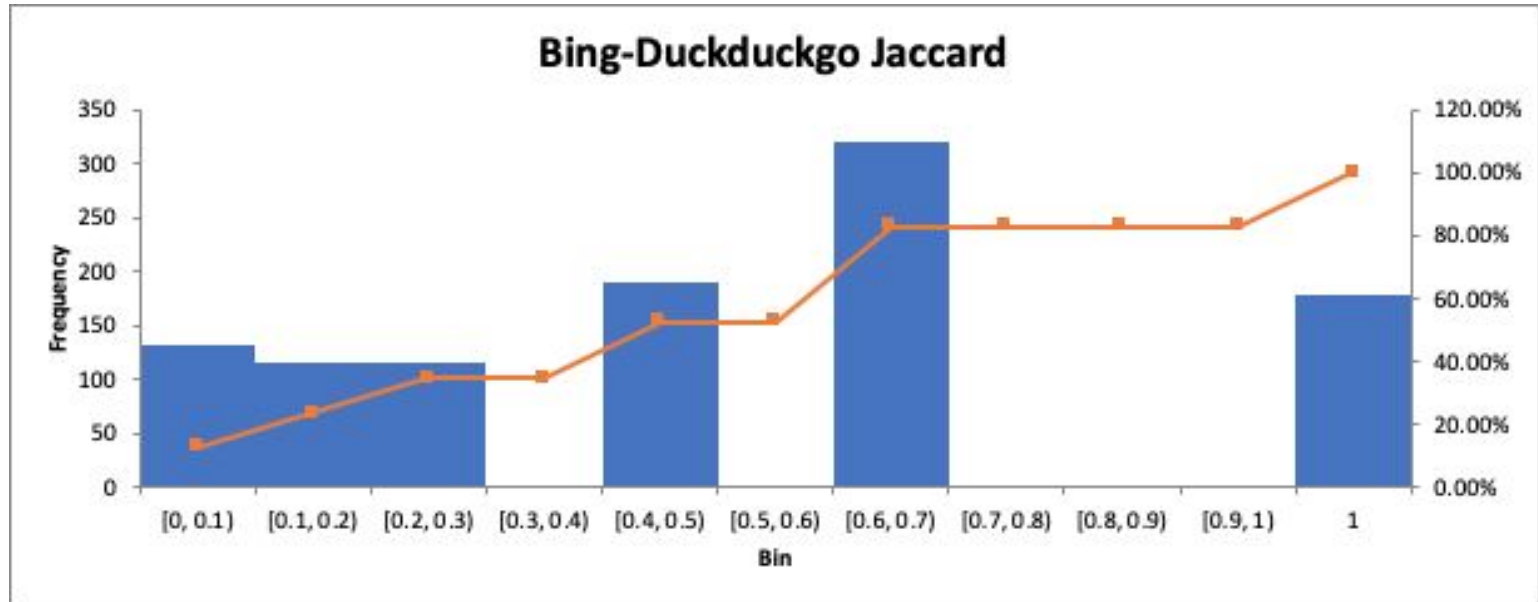
AnchorMAP Statistics

	Baidu-Bing	Baidu-DDG	Baidu-Google	Bing-DDG	Bing-Google	DDG-Google
<b>Median</b>	0.075	0.083	0.120	0.710	0.367	0.487
<b>Mean</b>	0.129	0.133	0.145	0.591	0.369	0.469

Baidu-Bing Wilcoxon Signed-Rank Test (Anchor)

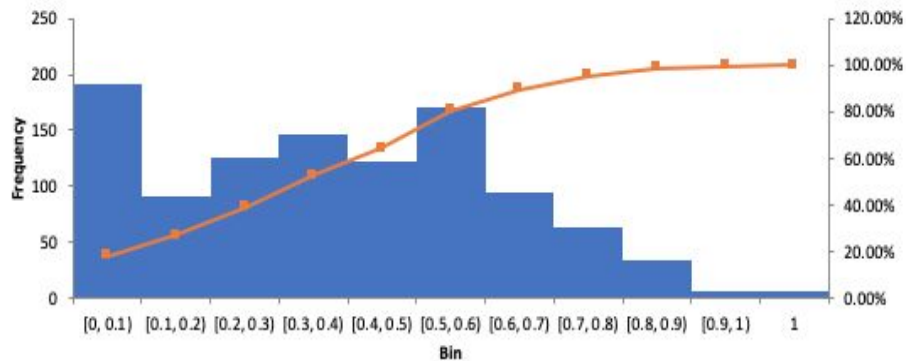
	Baidu-DDG	Baidu-Google	Bing-DDG	Bing-Google	DDG-Google
<b>p-value</b>	7.056e-2	7.354e-6	4.089e-142	2.311e-120	1.011e-157

# Limitations of Jaccard Correlation

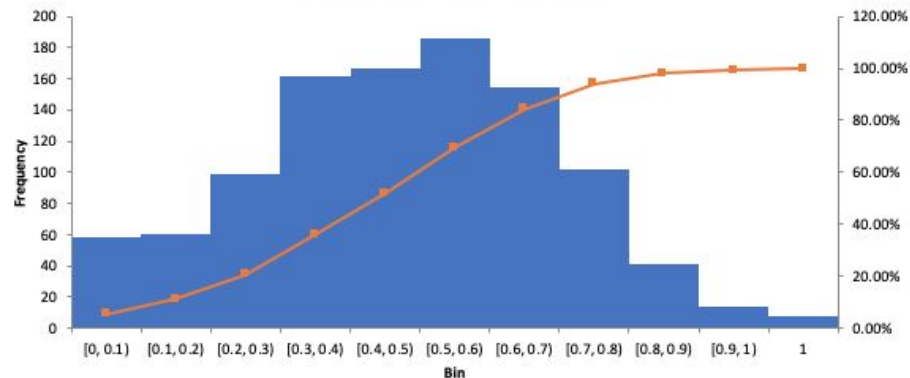


# American search engine comparison

**Bing-Google Anchor**



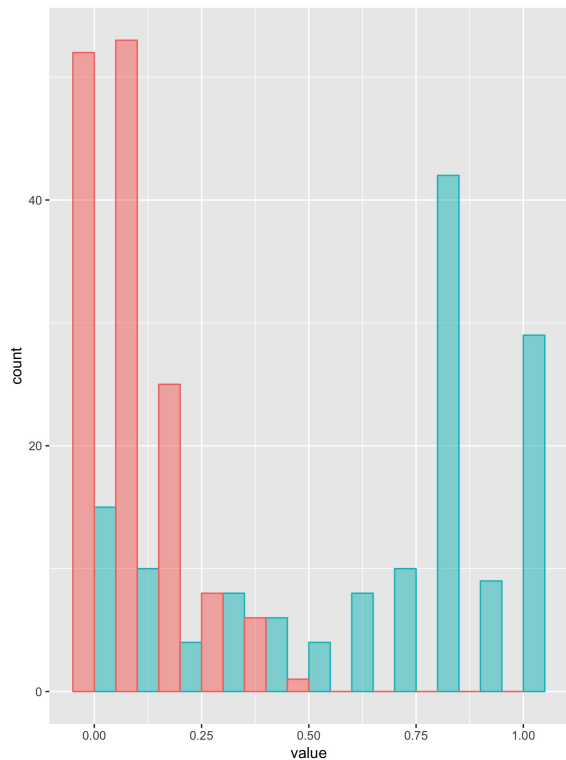
**Duckduckgo-Google Anchor**



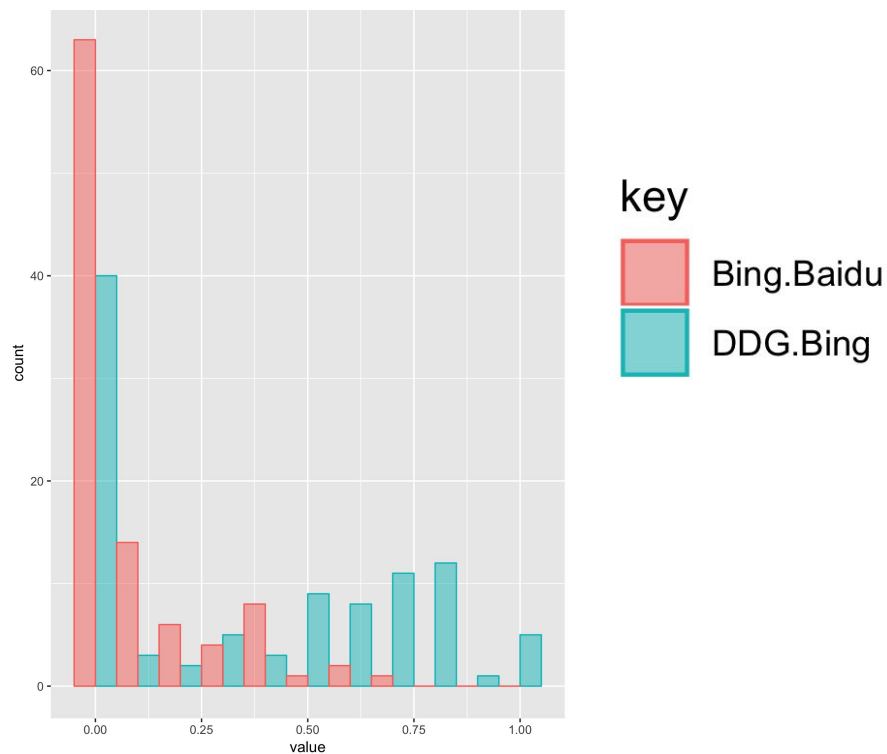


# Results: Cross-Topic Comparisons

“Ideas” with AnchorMAP



“Current” with AnchorMAP



# Limitations

- Consensus ranking approach was too naive for interesting results
- Jaccard and Top K Symmetric Difference were naive, and only re-confirmed what RBO and AnchorMAP revealed
- More search engines

# Conclusions

- Bing and DuckDuckGo perhaps define relevance similarly
- Baidu and Bing have the most disagreement
- These pairs have differing rankings across all topics

**Questions?**