Google

Desktop Search Comparative Research

go/desktop-search-comparative-research-july2020

Redacted @, **Redacted** @

7.15.20

Colab analysis:
https://colab.corp.google.com/drive/17ubghA6OAFETdUG6sjulmd_rbIT8SbIT?usp=sharing

Ex. No.

UPX0220

1:20-cv-03010-APM

Redacted

Executive Summary

- **Bing and Google were compared on 500 popular US queries** using two different methods of SxSs:
  - **Fullpage scrape** (where Bing scrapes did not include user location)
  - **Live links** SxS (where raters opened a dynamic link for the query in their own browser and rated their experiences)

- The **methodology of the SxS made a difference** (**-0.192** ± 0.060, p-value = 8.398e-10)
  - Full page scrapes were strongly positive (**0.185** ± 0.055, Google > Bing)
  - **Live link SxS was neutral** (**-0.007** ± 0.035).

- Both methodologies indicated that **Google is stronger in ranking** (authoritativeness, local context), while **Bing is stronger in feature diversity and interactive results**.

- Next steps: deep dive into features and rerunning experiment monthly

Google

Confidential + Proprietary

Motivation & Previous Work

Motivation

Quantify quality and feature gap between Bing vs. Google.

Previous Work

In light of covid, Redacted@ and Redacted@ looked at Bing

- Bing covid OSRP (Redacted @)
- Bing vs. Google Full page scrape (Redacted@, Redacted@)
  - In-depth analysis of experiment (Redacted @, Redacted, Redacted@)

https://docs.google.com/presentation/d/1ri7TlhDrQqbnncoROhuYGwOjELT_kd5pntn0R6i
S-Jc/edit#slide=id.g82eb967944_0_416

Redacted

Experiment Design

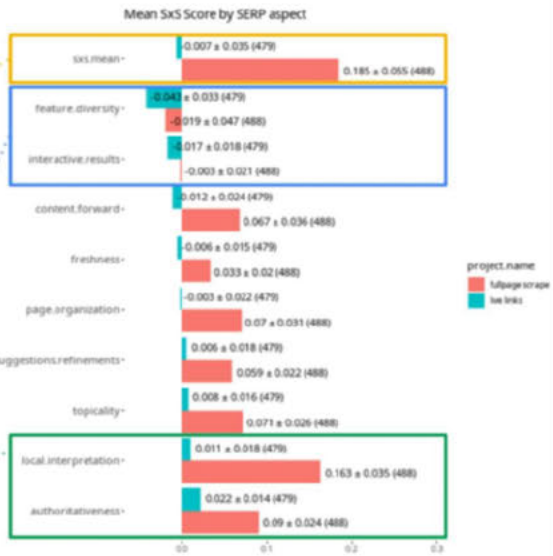**Queries**: ~500 popular queries sampled from April-June 2020

Methodology

- **Full page scrape (template)**
  - No Ads on Google
  - Scraped content captures what raters saw
  - Bing does not utilize user location
  - Results not interactive

- **Dynamic live links (template)**
  - Bing shows results for the rater's location
  - Results will be interactive
  - No scraped content (i.e. black box)

Mean SxS Scores

Mean SxS Score by SERP aspect

- **Overall mean SxS score**
  - Positive (**0.185** ± 0.055, Google > Bing) on the fullpage SxS
  - Neutral (**-0.007** ± 0.035) on the dynamic SxS.
  - Mean difference is **-0.192** ± 0.060 (p-value = 8.398e-10)

- **Bing is relatively stronger on features** (feature diversity and interactive results)

- **Google is relatively stronger in ranking** (authoritativeness, local interpretation)

SERP aspect	Value
sxs.mean	-0.007 ± 0.035 (479)
	0.185 ± 0.055 (488)
feature.diversity	-0.043 ± 0.033 (479)
	-0.019 ± 0.047 (488)
interactive.results	-0.017 ± 0.018 (479)
	-0.003 ± 0.021 (488)
content.forward	0.012 ± 0.024 (479)
	0.067 ± 0.036 (488)
freshness	-0.006 ± 0.015 (479)
	0.033 ± 0.02 (488)
page.organization	-0.003 ± 0.022 (479)
	0.07 ± 0.031 (488)
suggestions.refinements	0.006 ± 0.018 (479)
	0.059 ± 0.022 (488)
topicality	0.008 ± 0.016 (479)
	0.071 ± 0.026 (488)
local.interpretation	0.011 ± 0.018 (479)
	0.163 ± 0.035 (488)
authoritativeness	0.022 ± 0.014 (479)
	0.09 ± 0.024 (488)

project.name
- fullpage scrape
- live links

Live Links vs. Full Page: No user location in Full Page

In the full-page scrape, **Bing has no user location** so Google is much better. In live links SxS, both search engines utilize user location, but Bing shows rich features (map, listings) while Google only shows blue links. 15/20 top differences between SxSs need local context.

Live Links vs. Full Page: Richer features trigger in Live Links

In the fullpage scrape, raters prefer seeing the stock graph and price at top (not seen in the same format in Bing). In the live links SxS, "[Bing] has special stock price news block and one with quotes from analyst. It offers much more value in terms of SCRBs showing historical prices too." 4/20 top differences between SxSs showed different ranking and/or features on Bing when viewed live.

Live Links vs. Full Page: Richer features trigger in Live Links

Broad/ambiguous query (movie/greek god/monitoring system). Bing top result is for the Greek god interpretation while in Google, top interpretation is the monitoring system (minor interpretation). In the Live Links SxS, raters converge on the movie being the dominant interpretation, in which case they prefer Google RHS to Bing RHS. 1/20 top differences between SxSs were due to inconsistencies in rating (note: bing did show different content ranking between SxSs)

Mean SxS Scores

- **Overall mean SxS score**
  - Positive (**0.185** ± 0.055, Google > Bing) on the fullpage SxS
  - Neutral (**-0.007** ± 0.035) on the dynamic SxS.
  - Mean difference is **-0.192** ± 0.060 (p-value = 8.398e-10)

- **Bing is relatively stronger on features** (feature diversity and interactive results)

- **Google is relatively stronger in ranking** (authoritativeness, local interpretation, topicality)