

# Search Engine for Mobile Devices

Mansi Iyengar
Raymond Christopher
Shilpita Roy
Ziyi Yang



#### Content

- Domain & Targeted User Base
- Logging
- Search Engine Components
- Short Demonstration
- User Study Results
- Future Work & Conclusion

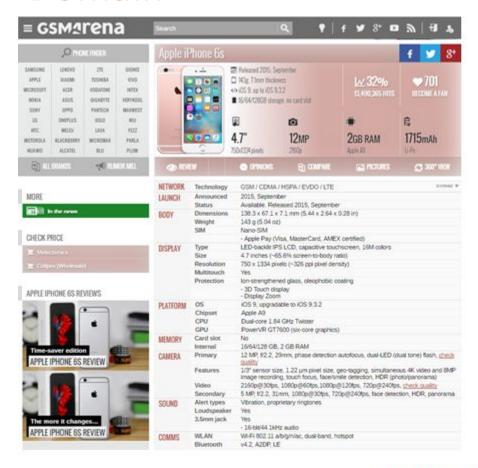


- Mobile Devices
  - cellphones, tablets
- Specifications
  - brand, screen size, camera quality, memory, price



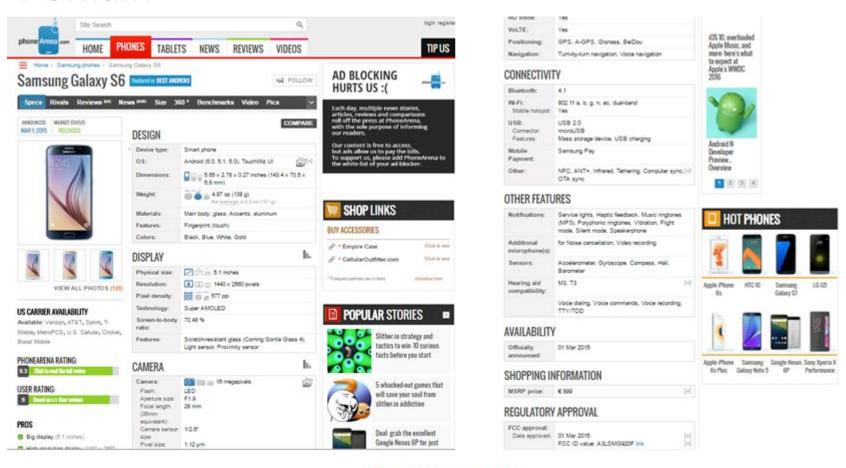
- Crawled websites
  - gsmarena.com
  - phonearena.com
  - phonescoop.com











phonearena.com





Features	Ungroup Features • Show all feature
Fingerprint Reader	in lock button, on side
Flashlight	Yes
GPS / Location	Type: A-GPS
Water- Resistant	Rating: IPX8 (SCUBA)
Accessibility	
Adjustable Text Size	Yes
Hearing Aid Compatible	Rating: M3, T4 (very tele-coil compatible)
Multiple Languages	Yes
TTY/TDD (Digital)	Yes
Alerts	
Vibrate	Yes
Connectivity	
Bluetooth	Supported Profiles: <u>HSP 1.2</u> , <u>HFP 1.6</u> , <u>OPP 1.2</u> , <u>PBA 1.1</u> , <u>A2OP 1.2</u> , <u>AVRC 1.3</u> , <u>HID. PAN. MAP 1.1</u> version 4.1 / supports <u>aptX</u>
MirrorLink	Yes
Mobile Hotspot	Yes
NEC	Yes
USB	Connector: Micro-USB
Wi-Fi	Version(s): 802.11 a,ac,b,g,n supports Wi-Fi Direct

phonescoop.com



## Targeted User Base

- People who are looking for new devices or want to upgrade their current devices
- Users who want to understand more about the specifications of their devices



# **Logging Session**

- User logging session is recorded in JSON
  - Queries
  - Number of clicks
  - Number of facets clicked
  - URL clicked
  - List of satisfying URLs
  - List of relevant URLs
  - Session start and end time



## Search Engine Components

- Apache Solr
- Stopping and Stemming at query time
  - default english stopword list
    - a, and, at, or, be, ...
  - EnglishMinimal stemmer
    - remove plurals only
  - Query expansion
    - GB, MB, MP
- Default Solr ranking algorithm
  - Boolean Model and Vector Space
     Model
    - weighted by tf-idf



# Search Engine Components

- Ajax-Solr
  - github.com/evolvingweb/ajax-solr
- 'Manager' sends requests to Solr
- 'Widget' handles reponses
  - ResultWidget, MultiSelectWidget, TextWidget,...
- Snippets are taken from description in document metadata



# Search Engine Components

- Facet classification documents during crawling
  - brands, camera quality, display size, system memory, price
  - store in metadata
- During indexing, Solr extracts information to fields
  - faceted search on fields
- Removed noise for improved rankings
  - advertisements, navigation, footers



## **Short Demonstration**

- Basic Search Engine
- Advanced Faceted Search Engine



# **User Study Methodology**

- User given 2 tasks one on basic and one in advanced
  - Find a 12 MP, 3.5 Inches Phone
  - Find a \$300 Samsung Phone
- User asked to mark the URL satisfying the task or relevant to the task.
- User asked to fill SUS and TLX questionnaires
- Age group of 15-35 considered in collection of 5 users.



# Efficiency

- The total number of clicks is averaged for all users for the basic search and then the advanced search.
  - Avg Basic = 5 Avg Adv. = 3
- The total number of Queries is averaged for all users for the basic search and then the advanced search.
  - Avg Basic = 1.8 Avg Adv. = 1.4
- The paired t-Tests is performed for time per task for basic and advanced search
  - P-score = 0.0932



## Effectiveness

- User clicks the checkboxes given for the links satisfying the task
- For each of the tasks users found relevant pages amongst the top ten ranked pages.
- Task 6 and 9 had no answers
  - User did not mark any page as relevant for them



## **User Satisfaction**

- Users were given SUS and TLX questionnaires
- SUS scoring system was used
- TLX was scored through online TLX scoring application
- The scores were averaged for the 5 users
- SUS (Score should be above 68)
  - Basic = 68.5 Advanced = 73
- TLX (Smaller score is considered)
  - Basic = 25.1 Advanced = 19.5



### **Future Work & Conclusion**

- Visualization
  - Number of hits per link
  - Histogram of fetched documents
- Logging session to be stored to database
- Concluding that advanced faceted search has a little edge over basic search with respect to efficiency and user satisfaction.