Final Presentation on

CREATIVE QUERIES FOR EXPLORATORY SEARCH

Team Project By:
Shravya Rama
Ikram Ul Haq

Under the guidance of:
Dr.-Ing.Tatiana Gossen

Presentation Date: 7th March 2016

CONTENTS

- MOTIVATION & INTRODUCTION

- ALGORITHMS AND APPROACH

- APPLICATION ARCHITECTURE

- USER INTERFACE DEVELOPMENT

- EVALUATION

1.17 Billion People Use Google Search

Unique searchers worldwide in December 2012 (in millions)

Google — 1,168m
Baidu — 293m
Yahoo! — 292m
Microsoft — 269m
Yandex — 74m
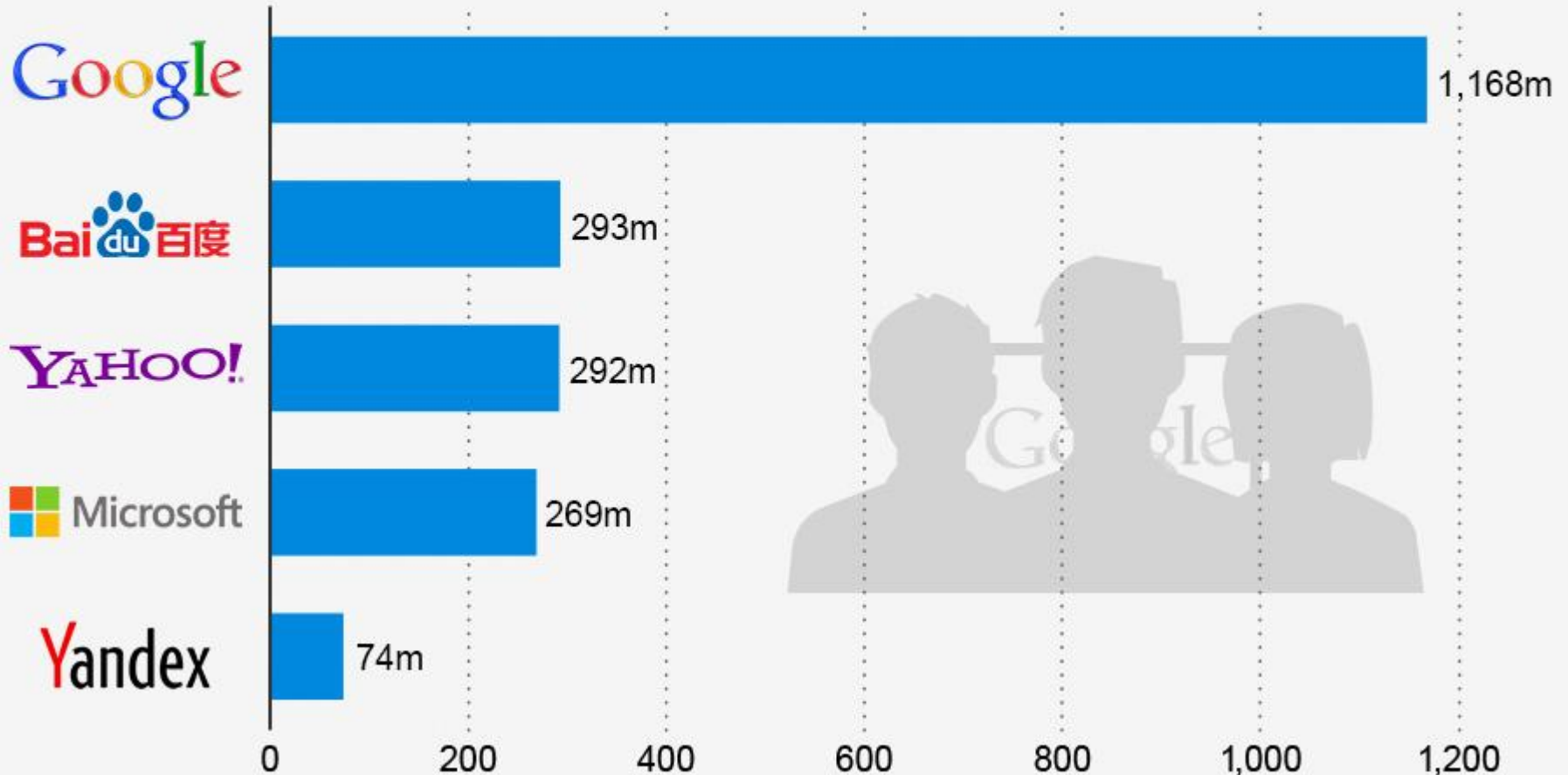
• GOOGLE STANDS AS THE TOP SEARCH ENGINE WITH 1.17 Billion USERS

MOTIVATION

TRADITIONAL SEARCH ENGINE ON SYSTEM 1

TRADITIONAL SEARCH ENGINE ON SYSTEM 2

- DOMINANCE OF COOKIES AND KEYWORD METHODOLOGY

SEARCH ON WEB

USER 1: Who knows specifically what he wants

For EXAMPLE:
- FIFA 2014 Results
- Latest Volkswagen Model
- JAVA Language Tutorials

http://10best4u.com/10-best-search-engines-on-web-to-make-your-searches-easy/

- Traditional Search Engine like GOOGLE implements searching on web

EXPLORATIVE SEARCH ON WEB

USER 2: Unfamiliar with the domain of their goal.
Unsure about the ways to achieve their goals.
Or even unsure about their goals in the first place.

For EXAMPLE :
ENTERTAINMENT

https://en.wikipedia.org/wiki/Exploratory_search

Reference to all images [1]

LITERATURE REVIEW

Creative Search Using Pataphysics (Digital Creativity 2013)
By: Fania Raczinski;
        Hongii Yang;
        Andrew Hugill;

The syzygy surfer: (Ab)using the semantic web to inspire
    Creativity(International Journal of Creative Computing, 2013)
By: James Hendler;
        Andrew Hugill;

- Consider the SYZYGY, CLINAMEN and the ANAMOLY of
the WORD rather than the word alone.

- WORDNET Dictionary is required to implement these algorithms.

ALGORITHMS AND APPROACH

SYZYGY:
We use WORDNET lexical database to find suitable results.
For a search term t
syno(t) = { s : s ∈ synonyms(t) } for $s \in$ syno(t)
hypo(t) = { h : h ∈ hyponyms(s) }
hyper(t) = { h : h ∈ hypernyms(s) }
holo(t) = { h : h ∈ holonyms(s) }
union(t) = hypo(t) ∪ hyper(t) ∪ holo(t)
syzygy(t) = { h : h ∈ union(t) ∧ ∃ h ∈ V }
V ∈ original vocabulary of the text

ALGORITHMS AND APPROACH

SYZYGY:

We use WORDNET lexical database to find suitable results.

For a search term t

$syno(\,t\,) = \{\,s : s \in synonyms(\,t\,)\,\}$ for $s \in$ syno$(\,t\,)$

$hypo(\,t\,) = \{\,h : h \in hyponyms(\,s\,)\,\}$

$hyper(\,t\,) = \{\,h : h \in hypernyms(\,s\,)\,\}$

$holo(\,t\,) = \{\,h : h \in holonyms(\,s\,)\,\}$

$union(\,t\,) = hypo(\,t\,) \cup hyper(\,t\,) \cup holo(\,t\,)$

$syzygy(\,t\,) = \{\,h : h \in union(\,t\,) \wedge \ \exists\,h \in V\,\}$

$V \in$ original vocabulary of the text

PROBLEM: Union of relations and intersection of union with original vocabulary.

SOLUTION: New Algorithm(E_SYZYGY)

ALGORITHMS AND APPROACH

E_SYZYGY:

We use WORDNET lexical database to find suitable results.

For a search term t

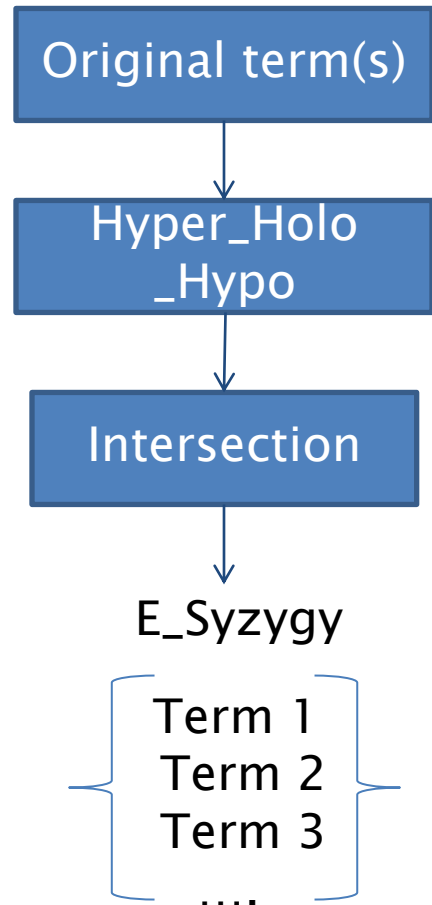$syno(t) = \{ s : s \in synonyms(t) \}$ for $s \in syno(t)$

$hypo(t) = \{ h : h \in hyponyms(s) \}$

$hyper(t) = \{ h : h \in hypernyms(s) \}$

$holo(t) = \{ h : h \in holonyms(s) \}$

$intersection(t) = hypo(t) \wedge hyper(t) \wedge holo(t)$

$E\_Syzygy(t) = \{ h : h \in intersection(t) \}$

Original term(s)

↓

Hyper_Holo _Hypo

↓

Intersection

↓

E_Syzygy

Term 1
Term 2
Term 3
….

ALGORITHMS AND APPROACH

ANOMOLY:
Anomaly function simply made use of WordNet's antonyms.

For a search term t
$syno(\ t\ ) = \{\ s\ :\ s \in synonyms(\ t\ )\ \}\ for\ s \in syno(\ t\ )$
$anto(\ t\ ) = \{\ h\ :\ h \in antonyms(\ s\ )\ \}$
$anomaly(\ t\ ) = \{\ h\ :\ h \in anto(\ t\ ) \wedge \exists\ h \in V\ \}$

ALGORITHMS AND APPROACH

ANOMOLY:
Anomaly function simply made use of WordNet's antonyms.

For a search term t
$syno(\,t\,) = \{\,s : s \in synonyms(\,t\,)\,\}\ for\ s \in syno(\,t\,)$
$anto(\,t\,) = \{\,h : h \in antonyms(\,s\,)\,\}$
$anomaly(\,t\,) = \{\,h : h \in anto(\,t\,) \wedge \exists\,h \in V\,\}$

PROBLEM: Intersection of antonyms of synonyms with the original
vocabulary.

SOLUTION: A New Algorithm(E_ANOMALY)
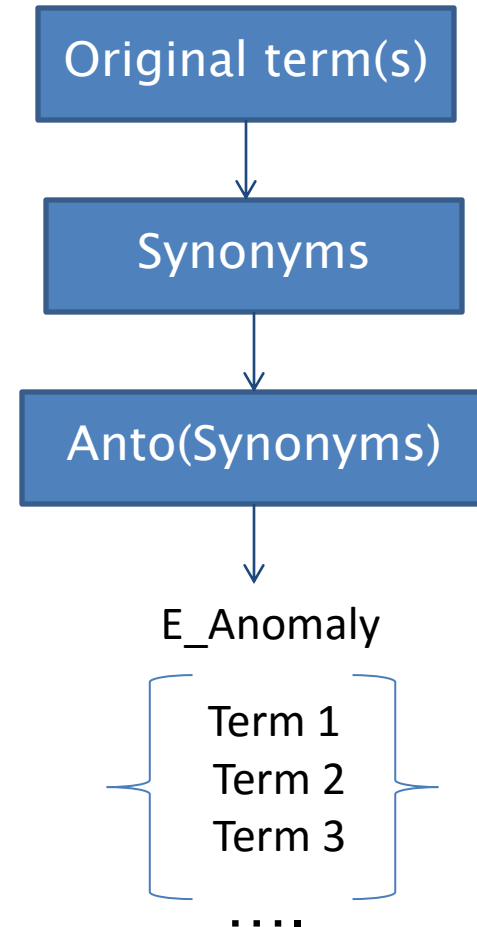
ALGORITHMS AND APPROACH

E_ANOMOLY:

Anomaly function simply made use of WordNet's antonyms.

For a search term t

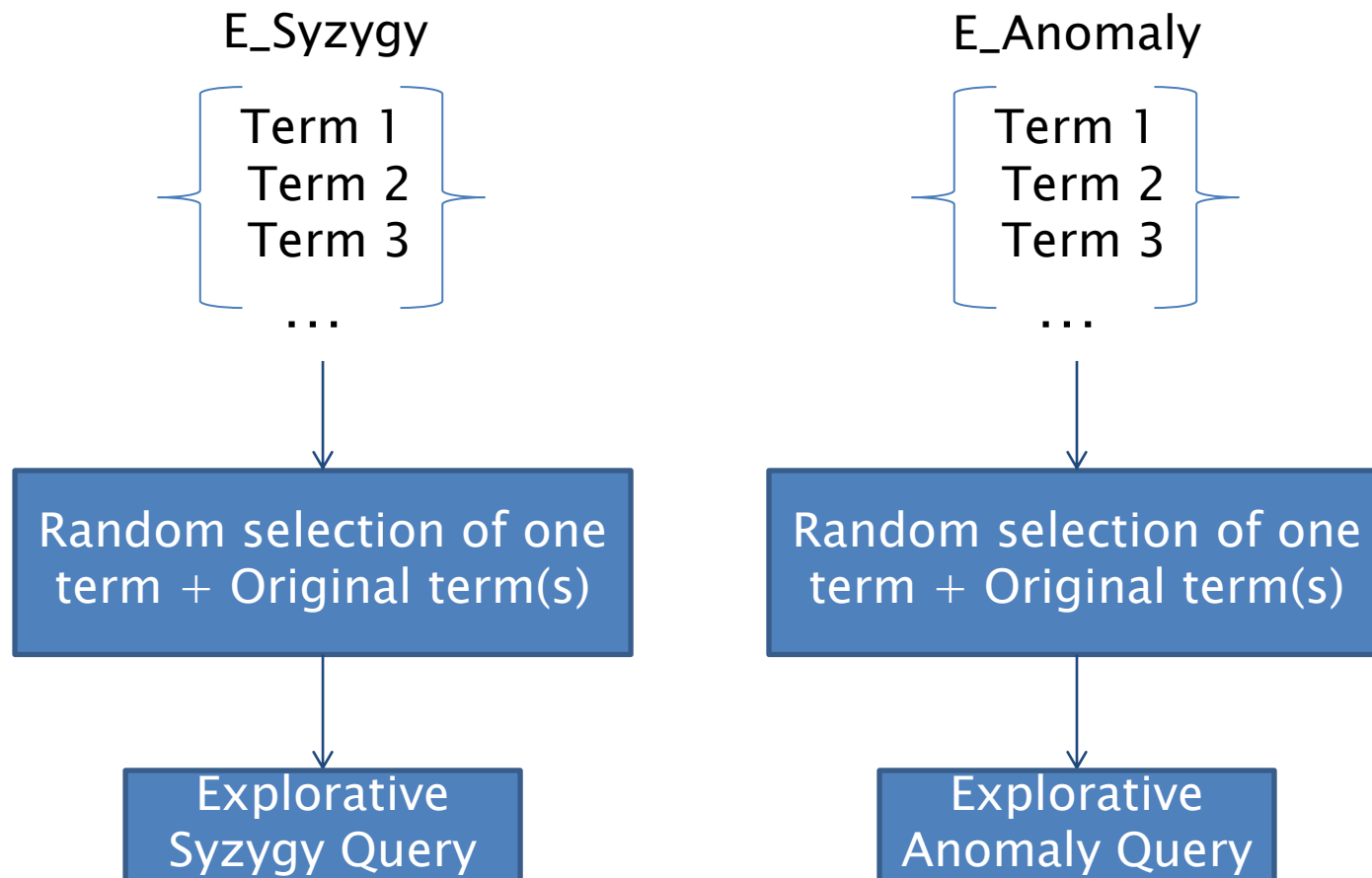$syno(\,t\,) = \{\,s : s \in synonyms(\,t\,)\,\}\,for\,s \in syno(\,t\,)$

$anto(\,t\,) = \{\,h : h \in antonyms(\,s\,)\,\}$

$anomaly(\,t\,) = \{\,h : h \in anto(\,t\,)\,\}$

Original term(s)

↓

Synonyms

↓

Anto(Synonyms)

↓

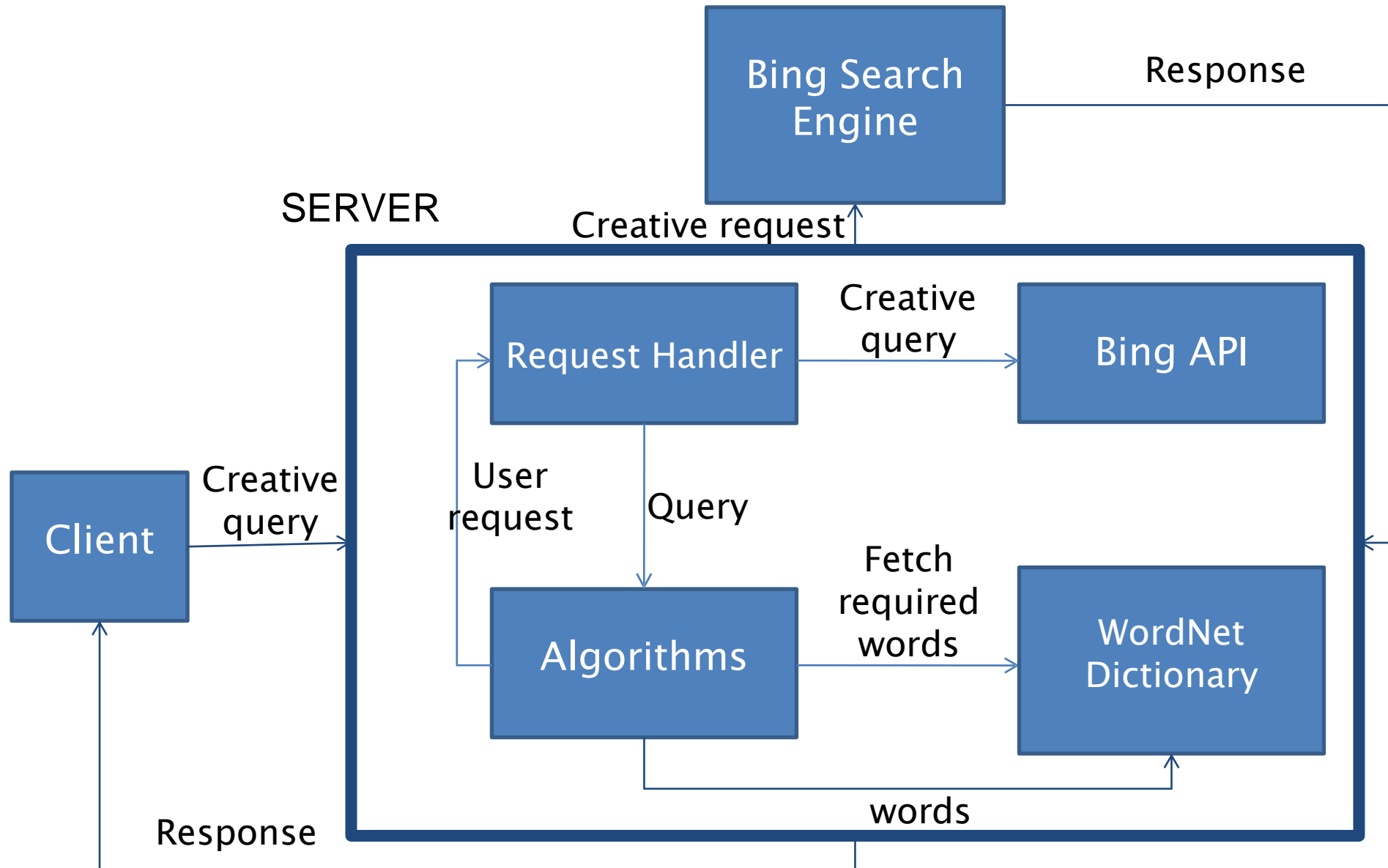E_Anomaly

Term 1
Term 2
Term 3
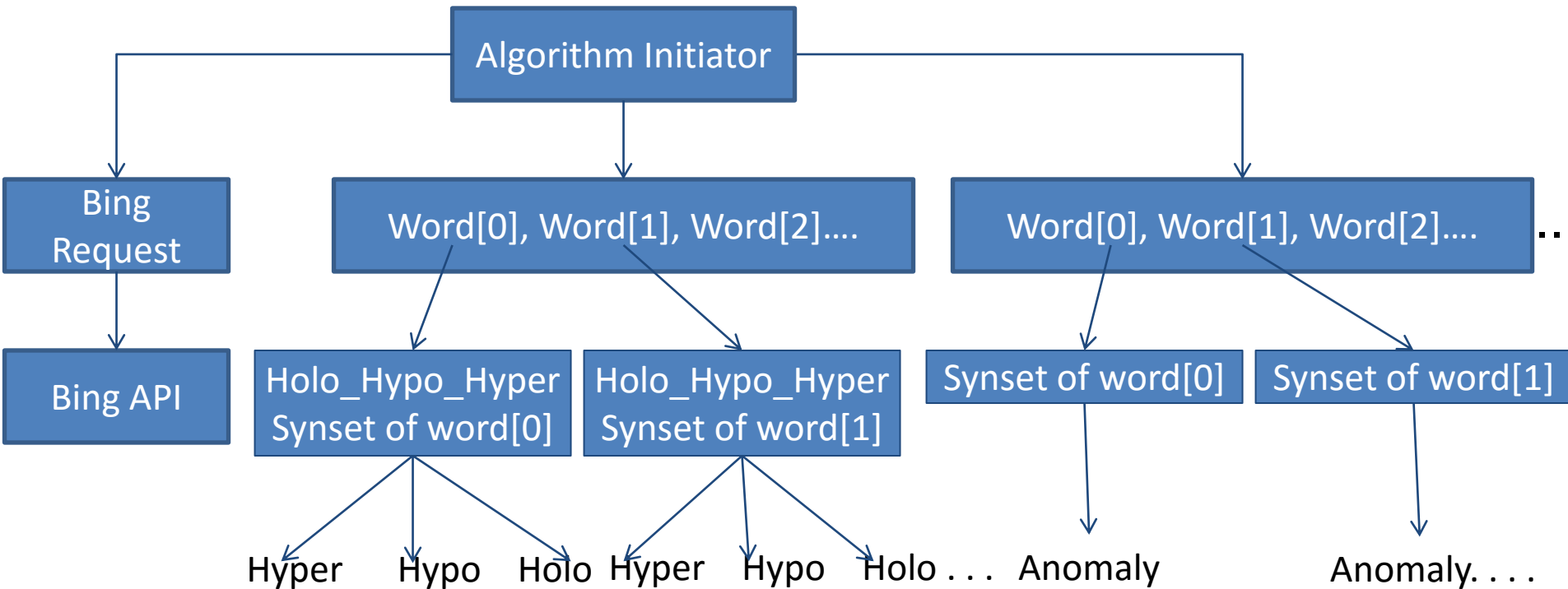
....

CREATIVE QUERY GENERATION

CREATIVE QUERIES FOR EXPLORATIVE SEARCH:

APPLICATION ARCHITECTURE

MULTILEVELS OF PARALLELISM

USER INTERFACE DEVELOPMENT

- QUERY DISPLAY INTERFACE

USER INTERFACE DEVELOPMENT

- GRID DISPLAY INTERFACE
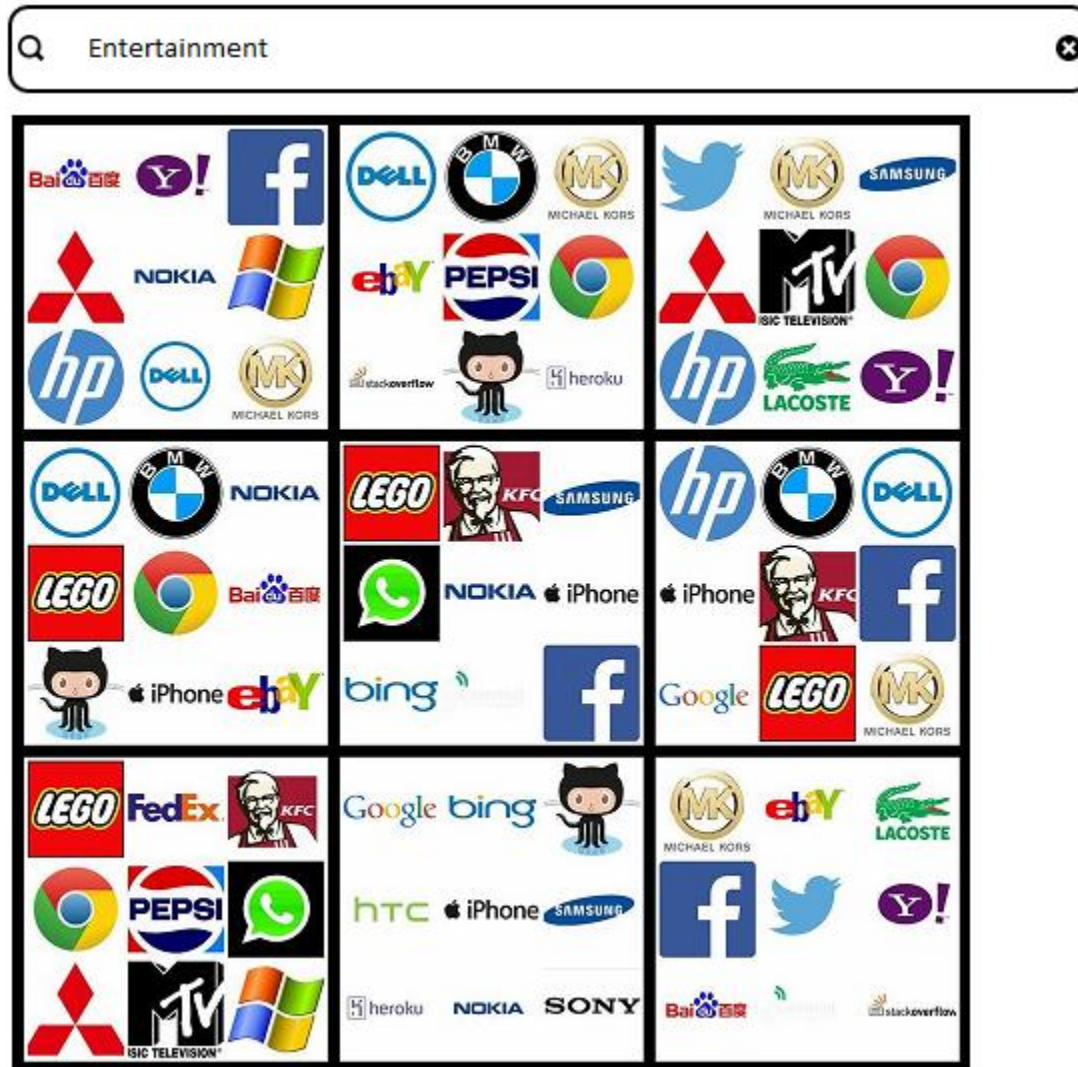
USER INTERFACE DEVELOPMENT

- INTERACTIVE INTERFACE

PRELIMINARY EVALUATION

Statistics of survey:

- Total no of Participants = 45
- No of Female Participants = 78.3.7%
- No of Female Participants = 21.7%
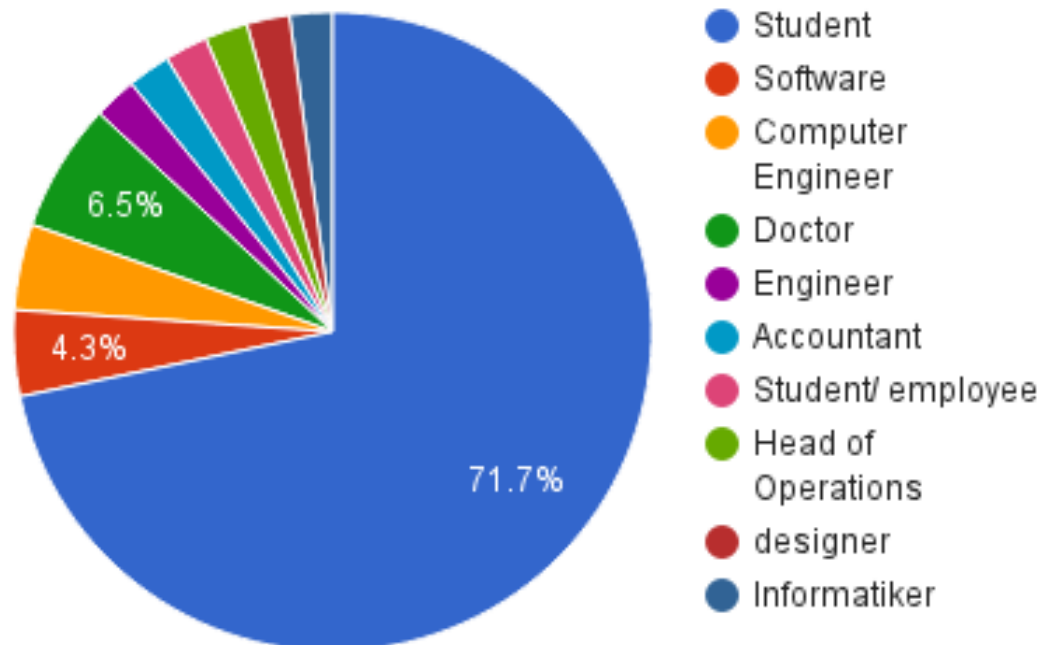
Percentage of participants according to professions

PRELIMINARY EVALUATION

- AGREEMENT MODEL
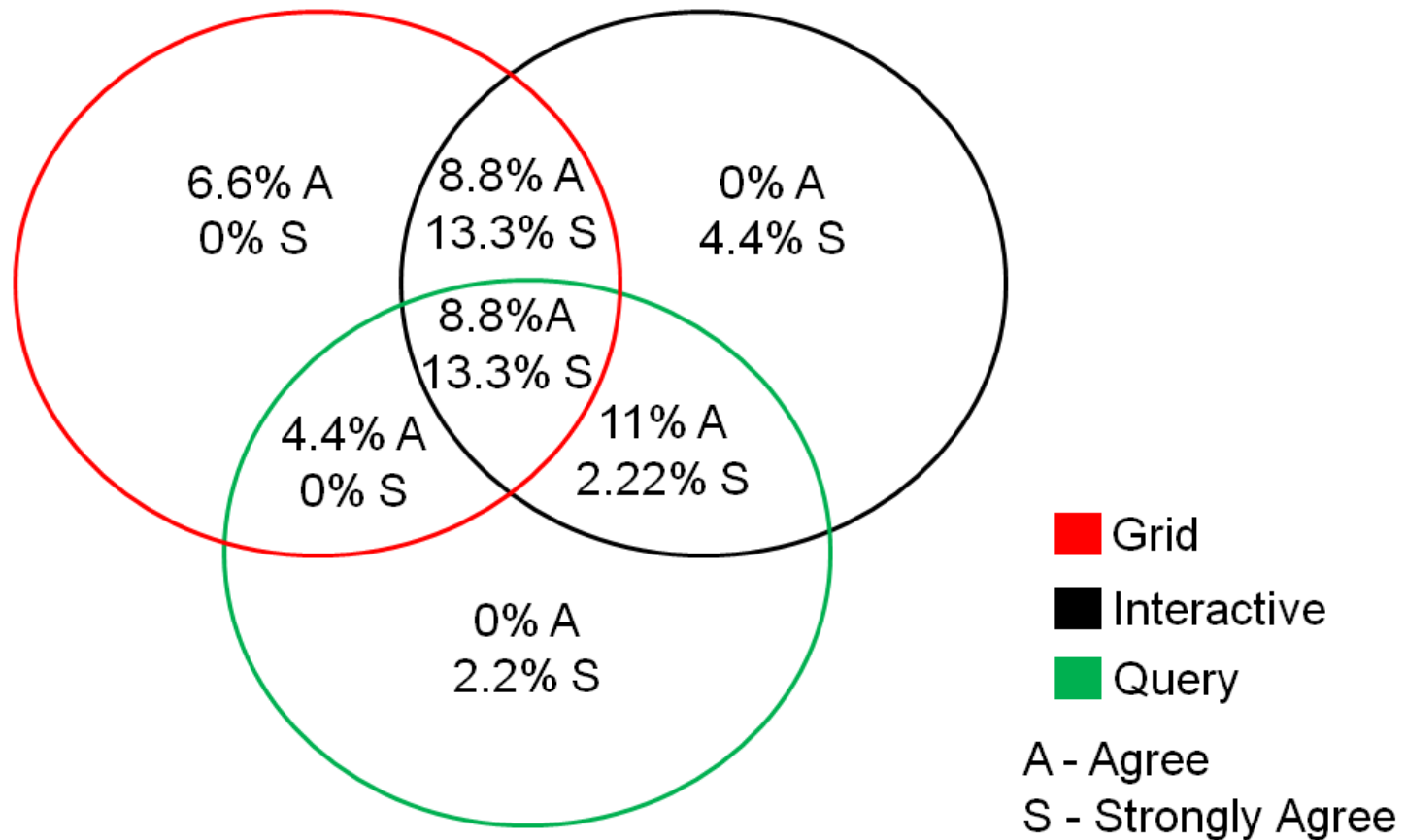
PRELIMINARY EVALUATION

- AGREE–DISAGREE MODEL

13.3%

4.4%

2.2%
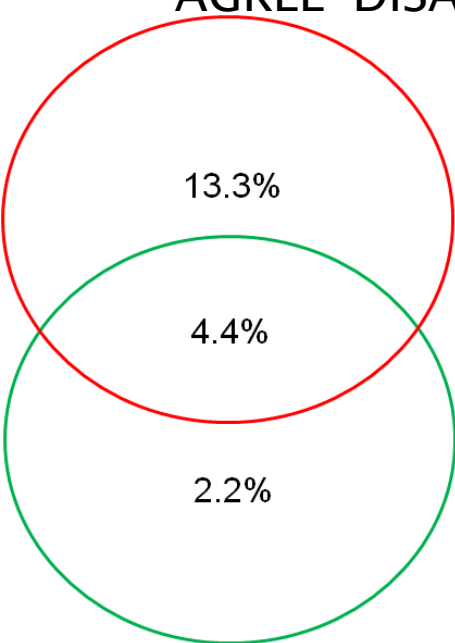
Agree -
■ Interactive

Disagree -
■ Query
■ Grid

*

*

11.11%
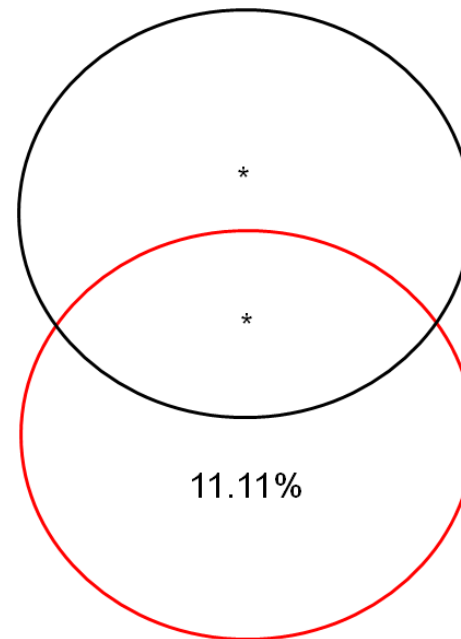
Agree -
■ Query

Disagree -
■ Interactive
■ Grid

*

*

2.2%
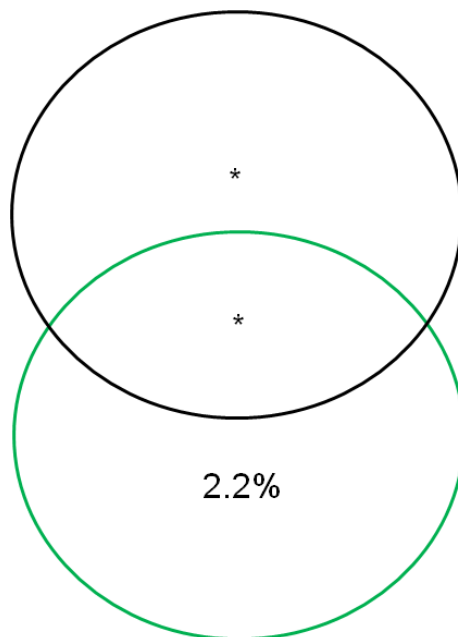
Agree -
■ Grid

Disagree -
■ Query
■ Interactive

PRELIMINARY EVALUATION

- LINEAR COMPARISION
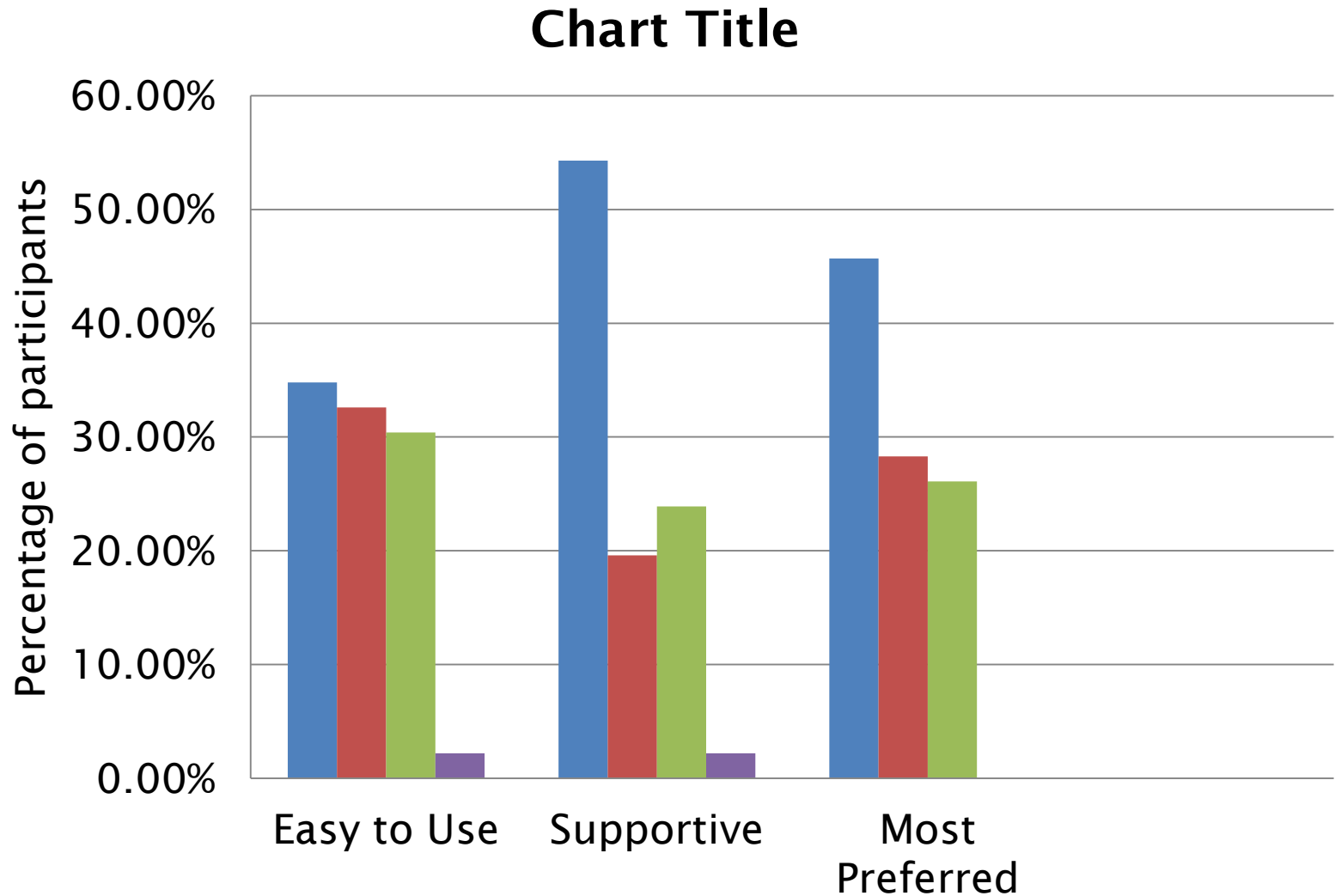
INTERACTIVE INTERFACE

| Live | | Exploratory Search |

Available Terms

Be Experience Go through Recorded

Explorative Syzygy results for *live*

. . . .

Explorative Syzygy results for *live*

. . . .

Conventional results for *live*

. . . .

Prev   Page#1   Next

FINAL EVALUATION

Statistics of survey:

- Total no of Participants = 28
- No of Female Participants = 50%
- No of Female Participants = 50%
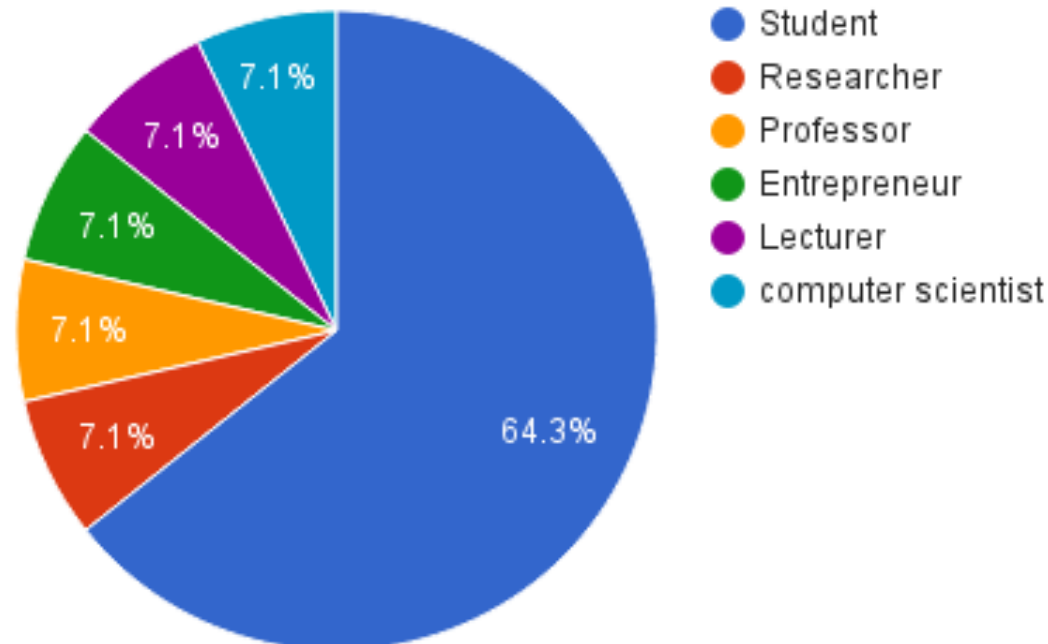
Percentage of participants according to professions

FINAL EVALUATION

- Which best describes your mindset when you are searching using a Search Engine?

- Could you adapt yourself to a new Search Engine interfaces if useful?

```
                          Search
   Ad-hoc Search                    Explorative Search
       50%                                50%

        Adapt                              Adapt

  No     Maybe    Yes            No      Maybe    Yes
14.28%  28.57%  57.14%           0%     28.57%  71.42%
```
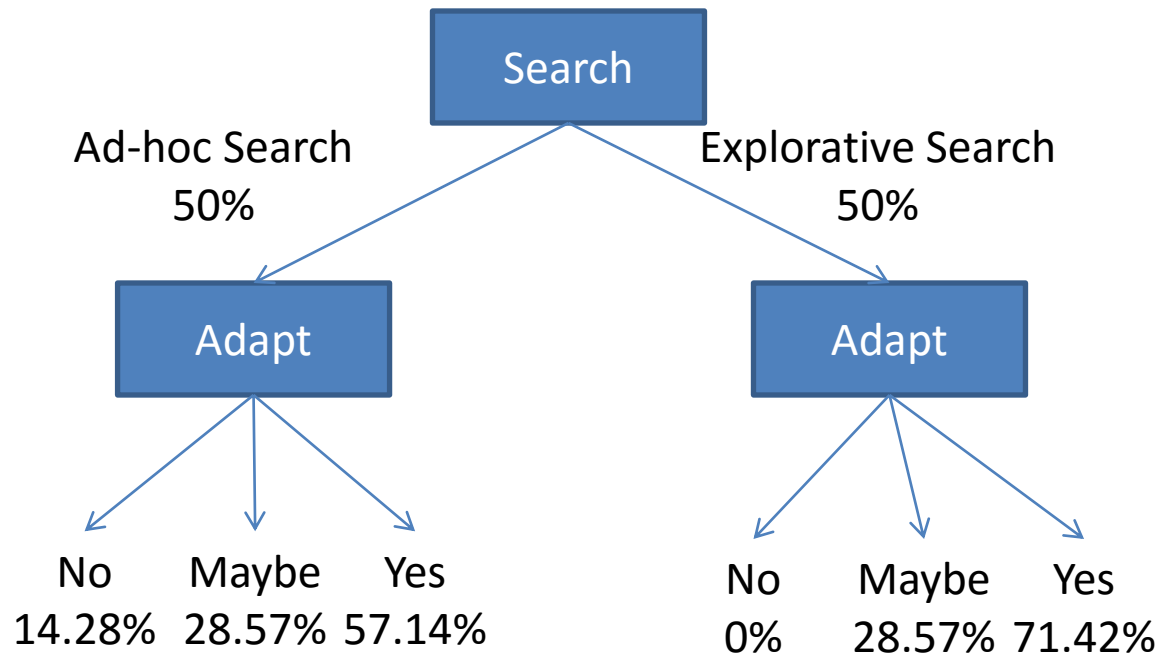
User Behavior before getting familiar with the interface

FINAL EVALUATION

How far do you agree that the Available Terms suggestions are related to the search

How far do you agree that the Available Terms suggestions are helpful for the researcher?

User

	Ready to Adopt 64.29%	Not Sure 28.57%	Not Ready to Adopt 7.14%
Disagree	11.1%, 11.1%	0%, 0%	0%, 0%
Agree	55.5%, 33.3%	50.0%, 50.0%	100%, 100%
Strongly Agree	33.3%, 55.5%	25.0%, 50.0%	0%, 0%
Neither Agree Nor Disagree	0%, 0%	25.0%, 0%	0%, 0%
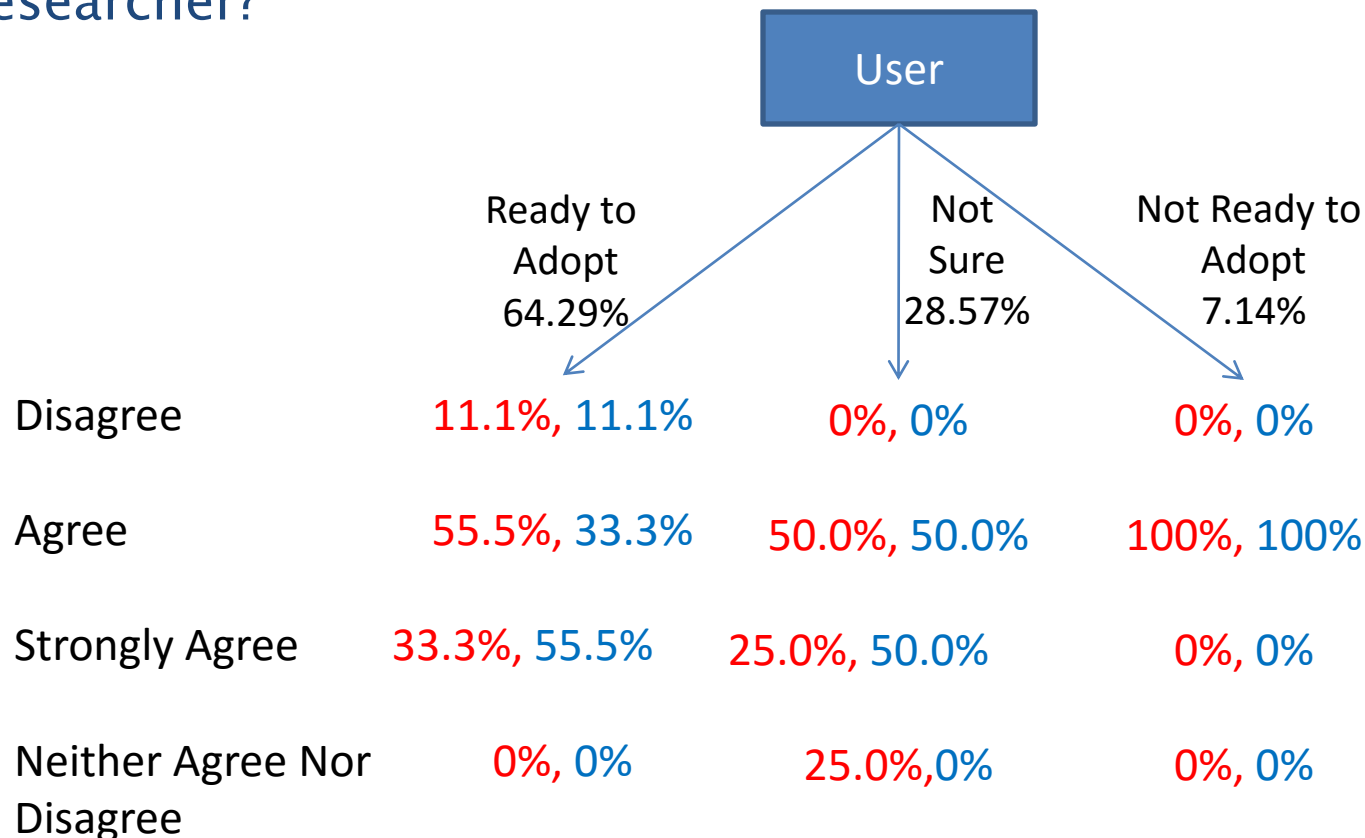
User Behavior after getting familiar with the interface

FINAL EVALUATION

- Comparison of Conventional and Explorative Results

User

Explorative Results

Conventional Results

Learnt New
Yes-92.86%
No-7.1%

Learnt New
Yes-76.78%
No-23.21%

Diversity	53.57%	39.28%
Usefulness	64.28%	32.14%
Interesting	50.00%	46.42%
Research	67.8%	32.14%
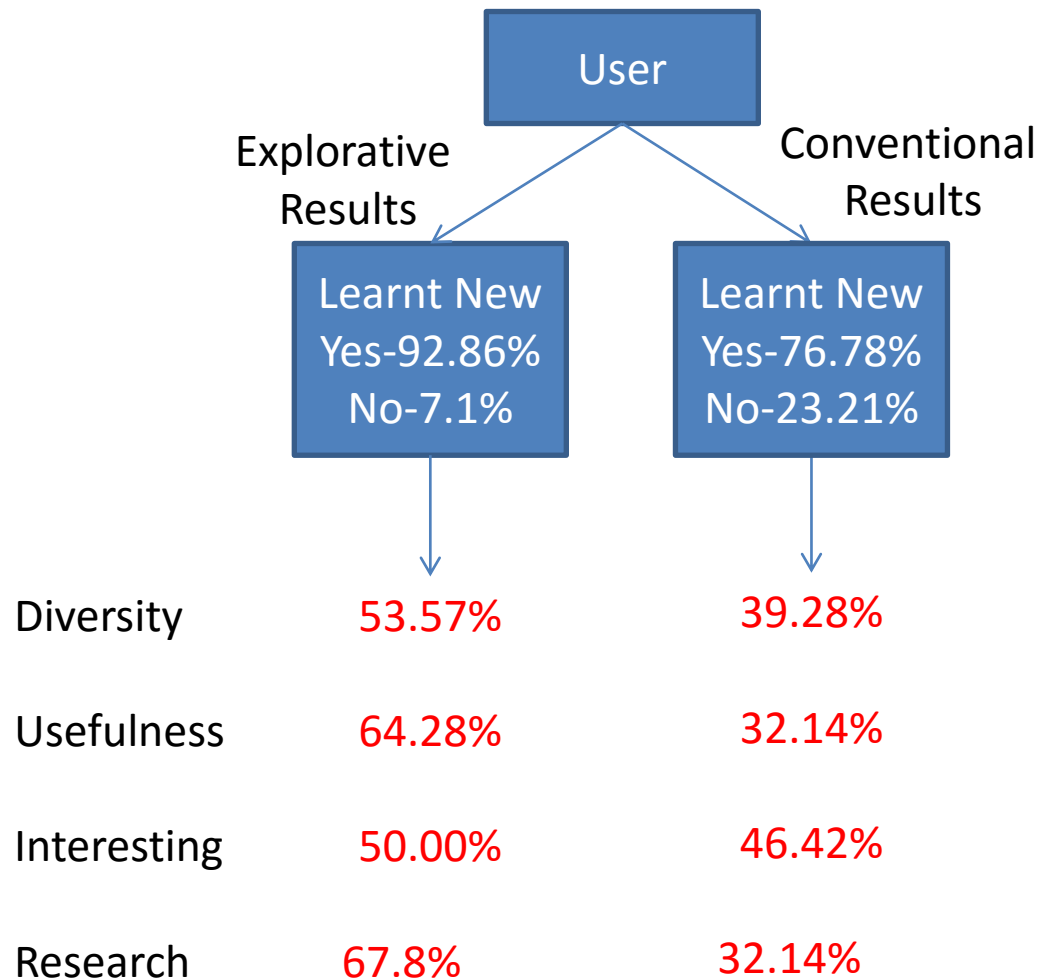
CONCLUSION AND FUTURE WORK

- An Interactive Search Engine Interface which generates effective, diverse, fruitful and compelling Explorative results is developed.

- To decrease the processing time(currently 30 seconds), in future, following mechanisms can be deployed.
    1. Inverted Indexing process.
    2. Kill with iron process(Adding Hardware).

- Limited set of terms in WordNet Dictionary also has to be enhanced.

Thank you for listening. Any questions?

REFERENCES

Digital Creativity:
https://scholar.google.co.in/citations?view_op=view_citation&hl=en&user=sRce9ogAAAAJ&citation_for_view=sRce9ogAAAAJ:TFP_iSt0sucC

International Journal of Creative Computing,2013:
https://scholar.google.co.in/citations?view_op=view_citation&hl=en&user=sRce9ogAAAAJ&citation_for_view=sRce9ogAAAAJ:e5wmG9Sq2KIC

[1]Sports:
https://www.google.co.in/search?newwindow=1&biw=1242&bih=585&tbm=isch&sa=1&q=sports&oq=sports&gs_l=img.3..0l10.103378.104573.0.105407.6.6.0.0.0.0.131.671.1j5.6.0....0...1c.1.64.img..0.6.663.siDaaxFZ2QI#imgrc=-1DK2TJ5lVuXqM%3A

[1]TV shows:
https://www.google.co.in/search?newwindow=1&biw=1242&bih=585&tbm=isch&sa=1&q=tv+shows&oq=tv+shows&gs_l=img.3..0l10.79917.81712.0.82233.8.7.0.1.1.0.141.754.1j6.7.0....0...1c.1.64.img..0.8.765._PJaAWoTtnw#imgrc=_v6UYojsFyD0UM%3A

[1]facebook:
https://www.google.co.in/search?newwindow=1&biw=1242&bih=585&tbm=isch&sa=1&q=facebook&oq=facebok&gs_l=img.3.0.0i10j0j0i10l2j0j0i10l3j0j0i10.29428.30764.0.32350.7.5.0.2.2.0.134.558.0j5.5.0....0...1c.1.64.img..0.7.574._9NfFoykG_8#imgrc=LUjnnkvhX4rfvM%3A

[1]games:
https://www.google.co.in/search?newwindow=1&biw=1242&bih=585&tbm=isch&sa=1&q=games&oq=games&gs_l=img.3..0l10.48184.49089.0.49919.5.5.0.0.0.0.135.534.2j3.5.0....0...1c.1.64.img..0.5.530.uSx1Ezd9qX0#imgrc=6K8fHew_LyI1rM%3A

[1]technology:
https://www.google.co.in/search?newwindow=1&biw=1242&bih=585&tbm=isch&sa=1&q=technology&oq=technology&gs_l=img.3..0l10.48782.50732.0.51413.10.8.0.2.2.0.113.843.1j7.8.0....0...1c.1.64.img..0.10.855.Kt2GCGGFKuM#imgrc=lSRH5n6g3a9wzM%3A

[1]Movies:
https://www.google.co.in/search?newwindow=1&biw=1242&bih=585&tbm=isch&sa=1&q=movies&oq=movies&gs_l=img.3..0l10.45832.46892.0.47092.6.6.0.0.0.0.162.696.1j5.6.0....0...1c.1.64.img..0.6.692.1k0ZdGwFaBE#imgrc=FvhmHpwP0i4_xM%3A

DESCRIPTION OF WORDNET DICTIONARY

Lexical Categories in WORDNET
- NOUN(117798+82115)
- ADVERB(11529+13767)
- ADJECTIVE(21479+18156)
- VERBS(4481+3621)

but ignores prepositions, determiners and other function words.

WordNet Dictionary determines
- Similarity between the words
- Distance amongst the words
- Produces Synsets

Word Net dictionary comprises a total of 206941 Word sense pairs (Unique strings+synsets)

DESCRIPTION OF WORDNET DICTIONARY

RELATIONS	Also Called	Example
Hypernym	Superordinate	Breakfast - meal
Hyponym	Subtype	Meal - lunch
Member Meronym	Has member	Faculty - professor
Member Holonym	Member of	
Has Instance		Composer - MJ
Instance		MJ– author
Part Meronym	Has part	Table - leg
Part Holonym	Part of	Wheel - bike
Antonym	Opposite	Leader- follower

SEARCH REQUEST HANDLER
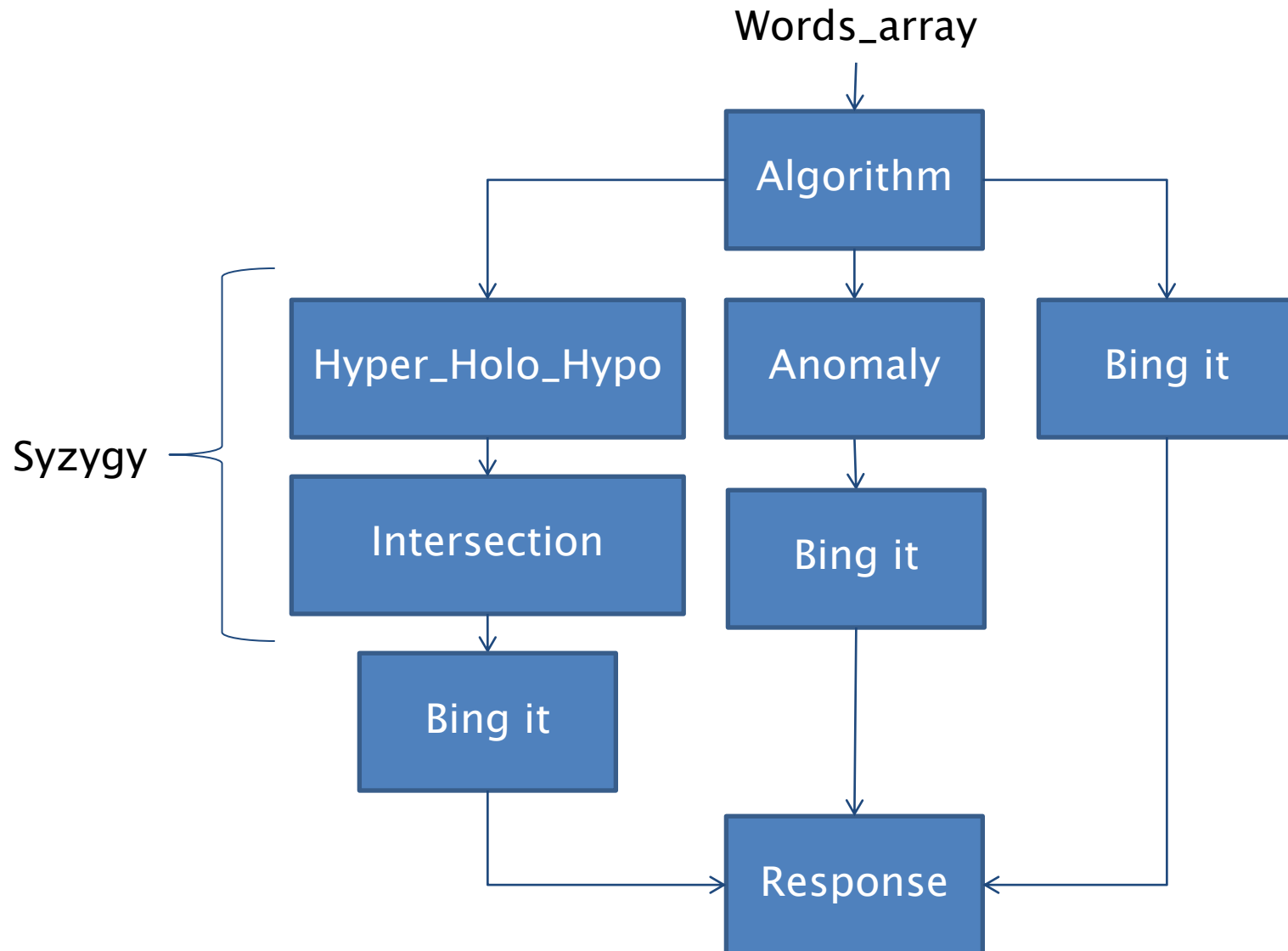
ALGORITHMS AND APPROACH

CLINAMEN:

The Clinamen function uses the Damerau–Levenshtein algorithm which measures the distance between two strings

For a specific search term t
Clinamen (t) =$\{ v : 0 <$ dameraulevenshtein $($t$, v) \leq 2 \}$, for $v \in V$

For e.g. Clinamen of <u>LIVE</u>= <u>LO</u>VE, <u>LI</u>ES, S<u>IZE</u>, R<u>IVE</u>R

- Clinamen is completely ignored due to it's enormous distinctiveness.

WORK LOAD DISTRIBUTION

TASK	IKRAM	SHRAVYA
STATE OF THE ART	Ca 25 hrs	Ca 25hrs
MOCK UP FOR FRONT END DESIGN		Ca 10 hrs
DESIGNING FRONT END (Html & CSS)	Ca 10 hrs	Ca 40 hrs
INTEGRATION OF BING API AND LEXICAL DICTIONARY(NODE.JS)	Ca 25 hrs	Ca 15 hrs
INSTALLING NODE.JS SERVER , NPM PACKAGE STRUCTURE	Ca 40 hrs	
IMPLEMENTATION OF ALGORITHMS(Javascript)	Ca 70 hrs	Ca 70 hrs
PRELIMINARY + FINAL USER STUDY	Ca 40 hrs	Ca 10 hrs