**SOCIAL INFORMATION NETWORKS**

CSE3021

PROJECT REPORT

**NETWORK ANALYSIS OF Q&A WEBSITES**

Submitted by:

*Aditya Firoda – 16BCE2184*

*Anshul Shrivastava – 16BCE2017*

Slot – A1

Project Guide – **Prof. Anuradha G.**



OCTOBER 2018

ACKNOWLEDGEMENTS

I take immense pleasure in thanking Dr. G. Viswanathan, my beloved Chancellor, VIT University and respected Dean, Dr. R. Saravanan, for having permitted me to carry out the project.

I express gratitude to my guide, Prof. Anuradha G, for guidance and suggestions that helped me to complete the project on time. Words are inadequate to express my gratitude to the faculty and staff members who encouraged and supported me during the project.

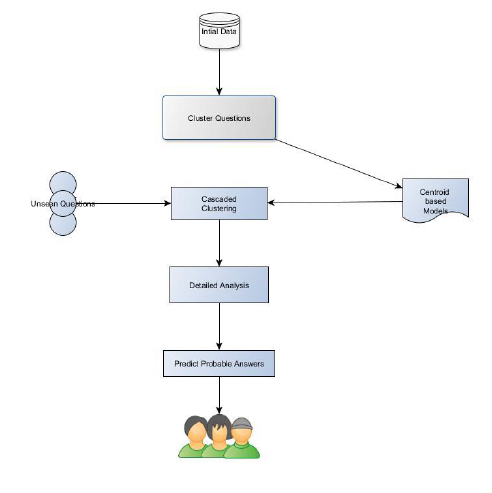
Finally, I would like to thank my ever-loving parents for their blessings and my friends for their timely help and support.

Signature of Student

**Abstract**

Historically, mailing lists have been the preferred means for coordinating development and user support activities. With the emergence and popularity growth of social Q&A sites such as the StackExchange network (e.g., StackOverflow), this is beginning to change. Such sites offer different sociotechnical incentives to their participants than mailing lists do, e.g., rich web environments to store and manage content collaboratively, or a place to showcase their knowledge and expertise more visibly to peers or potential recruiters. A key difference between StackExchange and mailing lists is gamification, i.e., StackExchange participants compete to obtain reputation points and badges. Using a case study of R, a popular data analysis software, in this paper we investigate how mailing list participation has evolved since the launch of StackExchange. Our main contribution is assembling a joint data set from the two sources, in which participants in both the r-help mailing list and StackExchange are identifiable. This allows for linking their activities across the two resources and also over time. With this data set we found that user support activities are showing a strong shift away from r-help. In particular, mailing list experts are migrating to StackExchange, where their behaviour is different. First, participants active both on r-help and on StackExchange are more active than those who focus exclusively on only one of the two. Second, they provide faster answers on StackExchange than on r-help, suggesting they are motivated by the gamified environment. To our knowledge, our study is the first to directly chart the changes in behaviour of specific contributors as they migrate into gamified environments, and has important implications for knowledge management in software engineering.

1. **Introduction**
   1. **Theoretical Background**

Historically, mailing lists have been the preferred medium for coordinating development and user support activities. In particular, mailing lists have been viewed as the de facto communication medium between knowledge seekers (e.g., users of the software asking for support) and knowledge providers (e.g., other users, more knowledgeable about the topic, or the developers themselves) in models of knowledge sharing in open source. The two categories of knowledge actors have been reported to co-exist in a symbiotic relationship, wherein “the community learns from its participants, and each individual learns from the community”. However, their motivations for participation may differ. For instance, knowledge seekers may directly benefit from having their problems solved, while knowledge providers may be motivated intrinsically (e.g., by altruism), or by learning about the problems other users are experiencing. Recent years have witnessed the emergence and growing popularity of software-development-related social media sites, such as GitHub1 (coding), Jira2 (issue tracking), or the StackExchange network (question and answer websites, e.g., StackOverflow for “professional and enthusiast programmers,”3 or CrossValidated for “statisticians, data analysts, data miners and data visualization experts”4 ). Such sites are rapidly changing the ways in which developers collaborate, learn, and communicate among themselves and with their users. Moreover, they are offering different socio-technical incentives to their participants, e.g., rich Web 2.0 platforms to store and manage content collaboratively, or a place to showcase their knowledge and expertise more visibly to peers and potential recruiters. In addition, StackExchange sites employ gamification to engage users more: questions and answers are voted upon by the community; the number of votes is reflected in the poster’s reputation and badges; exceeding various reputation thresholds grants access to additional features (e.g., moderation rights on topics and posts); reputation and badges can also be seen as a measure of one’s expertise by potential recruiters, and are known to motivate users to contribute more. Activity on StackExchange sites can also elevate one to celebrity status within the developer community (see, e.g., the discussion around Jon Skeet5 , the most prolific contributor to StackOverflow).

**Fig. 1.1.1 Flow diagram for our proposed system**

A practical scenario of an internal wiki generally involves a large database. The application of linear search in such a scenario proves to be infeasible - the time taken for lookup is high and the results are inaccurate due to mismatching content.

Naturally, the richer user interfaces, wider audiences, or different incentives and motivations for participation inherent in social Q&A sites are challenging the supremacy of mailing lists as the de facto communication medium between knowledge seekers and knowledge providers. For example, StackOverflow is known to provide good technical solutions and to provide them fast. At the same time, mailing list participants are signalling the need for more modern support,6 and are even promoting a transition to StackExchange.7 Our goal here is to study in detail the effects of such a transition on contributors and their work. Are mailing list participants transitioning to StackExchange? If so, do they behave differently on StackExchange than on the mailing list?

**1.2 Motivation**

The importance of user support for the adoption, growth, and success of open source projects is well-recognized. Traditionally, user support was organised through mailing lists, forums, user groups, etc. However, as new venues of information and tools for information access are emerging (e.g., weblogs, wikis, social Q&A sites), people’s online information seeking behaviour is also evolving. When it comes to user support activities around open source software, though, what is common among both traditional and new venues of information is that this “mundane but necessary task” is typically carried out by unpaid volunteers. Often, the developers themselves take part in these activities, but thriving open source projects also succeed in enlisting some of their users to offer assistance to peers. But what motivates knowledge providers to answer other people’s questions? Together with online information seeking behaviours, these motives are also evolving. In traditional information-sharing venues, knowledgeproviders participate for reasons related to, albeit different, than those of developers contributing to open source. Developers contribute for reasons such as: a direct need for the software; enjoyment of the work itself; or the enhanced reputation arising from high-quality contributions. Knowledge-providers, on the other hand, are reportedly motivated by: learning about problems other users are experiencing; an enhanced feeling of being part of a community; personal benefits of learning through teaching; an enhanced likelihood of receiving help in the future; or a sense of obligation from having received help from others in the past.

* 1. **Aim of the Proposed Work**

The aim of the project was that as social Q&A sites, e.g., the StackExchange network, are becoming more popular, instances of what the economic literature calls “signalling incentives” start to better explain what motivates knowledge providers to help others, in addition to the previous reasons. Career concerns (e.g., online activities in social Q&A sites are more visible to employers and recruiters, who may use them to find qualified people or a desire for peer recognition (e.g., reputation building) are among the listed motives. However, such signalling incentives may not be equally applicable to all knowledge providers. For example, we can expect that more knowledgeable providers would draw greater benefit from signalling, and thus signal more. Therefore, to obtain a more fine-grained understanding of the transition from traditional information venues to social Q&A, it is important to distinguish between different groups of participants (e.g., developers, with likely more knowledge about the software, versus non-developers).

**1.5 Report Organization**

The report has been built in-accordance to the set guidelines and covers all the stated topics. In the upcoming sections we discuss about the existing literature survey along with the proposed system architecture for our project. We then proceed by stating in detail the various functional and nonfunctional requirements identified for developing our search engine platform. The report concludes with the results and discussion section, in which we debate about the performance of our system with the existing ranking systems.

**2. Literature Survey**

**2.1 Summary of the Existing Models/Work**

1. Facts or Friends? Distinguishing Informational and Conversational Questions in Social Q&A Sites

Authors - F. Maxwell Harper, Daniel Moy, Joseph A. Konstan GroupLens Research University of Minnesota

Summary – This paper discusses the fact that social websites generate mainly two types of questions conversational or informational. While the latter needs to archived we face the problem of classification which can be solved by any humans and by using the proposed algorithms of the paper they could achieve the results with up to 82.1 accuracy.

1. Self-Presentation and the Value of Information in Q&A Websites

Author - Daphne Ruth Raban Graduate School of Management and the Center for the Study of the Information Society, University of Haifa, Mount Carmel, Haifa

Summary – Centre idea of this paper social interaction will lead to more interaction and which will attract these users. So author tries to understand all the interactions between the users whether of implicit or explicit form in nature. These behaviours are given weight and implicit behaviour is studied more carefully as most people don’t give always feedback and comment to others.

1. Exploring characteristics and effects of users participation in online Q&A sites

Authors – Chirag Shah, Jung Sun Oh and Sanghee Oh

Summary – This paper mainly assesses the performance of 2 websites i.e. Google Answers and Yahoo Answers which are based mainly on user participation and sees how their user profiles and interaction leads to certain design choices and quality of participation in which it proves that Yahoo Answers encourages users to have an active participation.

1. **Topic**: Knowledge Sharing in the Online Social Network of Yahoo!

**Authors:** P. Vijayalakshmi, V. Hemalatha, Associate Professor, P G Student Department of M.C.A Panimalar Engineering College, India.

**Summary:** In this paper, the authors have described the need of online Q&A platforms and the issues related with Yahoo answers’ spamming. There is no effective ways to control duplication of answers. The advantages and disadvantages of the existing algorithms are described. The authors have implemented an algorithm where the answers with most likes are shown early than other answers.

1. **Topic:** Tie Strength in Question & Answer on Social Network Sites

**Author:** Katrina Panovich, Robert C. Miller, David R. Karger, MIT CSAIL

**Summary:** In this paper, the authors have evaluated the difference in the quality of answers between strong ties (close friends) and weak ties (acquaintances). The authors have referred to the Granovetter’s work “The Strength of Weak Ties”. The authors have given different view on the earlier research, they inferred that tie strength does nothave the same effect in question answering as would be predicted. In fact, there seems to be a subtle increase in answer quality from strong ties, rather than weak ties.

1. **Topic:** Analysis of Question and Answering Behaviour in Question Routing Services

**Author:** Zhe Liu, Bernard J. Jansen

**Summary:** In this paper, the authors have used a Chinese Q&A platform named Wenwo to analyse the Q&A behaviour in Question Routing Services. On the platform, they analysed over 340658 questions posted in ten months and they also analysed over 1.7 million answers. They noticed that the users asked less questions than the answers i.e. there were more knowledge contributors.

1. **Topic**: Facts or Friends? Distinguishing Informational and Conversational Questions in Social Q&A Sites

**Authors** - F. Maxwell Harper, Daniel Moy, Joseph A. Konstan GroupLens Research University of Minnesota

**Summary** – This paper discusses the fact that social websites generate mainly two types of questions conversational or informational. While the latter needs to archived we face the problem of classification which can be solved by any humans and by using the proposed algorithms of the paper they could achieve the results with up to 82.1 accuracy.

1. **Topic:** Self-Presentation and the Value of Information in Q&A Websites

**Author** - Daphne Ruth Raban Graduate School of Management and the Center for the Study of the Information Society, University of Haifa, Mount Carmel, Haifa

**Summary** – Centre idea of this paper social interaction will lead to more interaction and which will attract these users. So author tries to understand all the interactions between the users whether of implicit or explicit form in nature. These behaviours are given weight and implicit behaviour is studied more carefully as most people don’t give always feedback and comment to others.

1. **Topic:** Exploring characteristics and effects of users participation in online Q&A sites

**Authors** – Chirag Shah, Jung Sun Oh and Sanghee Oh

**Summary** – This paper mainly assesses the performance of 2 websites i.e. Google Answers and Yahoo Answers which are based mainly on user participation and sees how their user profiles and interaction leads to certain design choices and quality of participation in which it proves that Yahoo Answers encourages users to have an active participation.

**2.2 Threats to Validity**

Despite our detailed efforts in experimental design, data gathering and data analysis, we do note several threats to the validity of our methodology and conclusions. The core step in the data gathering, extracting information from the StackExchange data dump and the mail archives, is subject to threats to validity akin to those identified for digital trace data. Following the classification of Howison , system and practice issues in our work may be related to communication through other mailing lists than r-help and other StackExchange sites than StackOverflow and CrossValidated. Indeed, advanced R packages such as lme4 have separate mailing lists,30 and experts working with these packages might prefer not to migrate to StackExchange. However, r-help, StackOverflow and CrossValidated are the biggest platforms of their kind. We also knowingly omitted other venues where knowledge exchange happens, e.g., Google+ groups or individual blogs. Our results are very strong, and give us confidence that StackExchange is a good representative of the missing Q&A communities. Moreover, the data extracted is subject to potential reliability threats arising from missing messages from mail archives, and questions and answers being deleted from StackExchange. Indeed, as indicated by one of the survey participants, StackExchange sites encourage the participants “to ask well-formed questions, leading to well-formed answers”, while ill-formed questions are being removed. Question or answer removal might lead to underestimating the activity of the StackExchange contributors, as well as the overlap between StackExchange and r-help. Similarly, removal of StackExchange accounts might lead to underestimation of the overlap. Another reliability threat is related to multiple system representations of a single individual, which again might incur an underestimate of the overlap between StackExchange and r-help. We have explicitly addressed this threat when discussing identity merging in the methodology section: while no data gathering is perfect, our efforts are particularly detailed, and have been shown elsewhere to work satisfactorily. Finally, reliability might be threatened by noise in the data, such as spam messages stored in the mail archives. The next group of threats to validity refers to representation of the data extracted in the model of one question followed by multiple answers. While the distinction and relation between questions and answers on StackExchange is explicit in the data organisation, the natural counterpart in the mailing lists is the discussion thread (see “Comparing apples and oranges” in the methodology section). Recognition of the discussion thread starter as the asker, and other thread participants as answerers can, however, be threatened by the thread starter posting an announcement rather than a question, as well as by thread participants trying to clarify the intention of the thread starter rather than answering her questions. We have performed an informal evaluation of the discussion threads in r-help and observed that the lion’s share of threads adheres to the “question and answers” model similar to StackExchange. Temporal aggregation threats arise when aggregating events that occur at different points in time (cf. Figures 1 and 4). Indeed, inappropriate choice of the time granularity might have made our work subject to ecological fallacy. However, our choice for month as the basic time unit stems from the way traditional mail archives are presented (per month) as well as the literature.

**3. Overview of the Proposed System**

**3.1 Introduction**

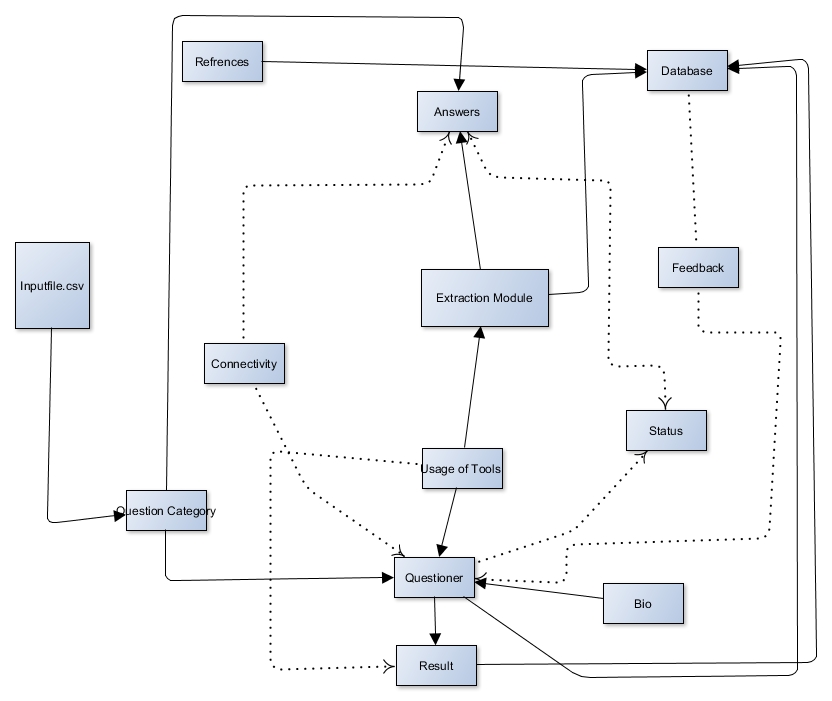
As case for our case study we selected Python, a popular data analysis software, for several reasons: First, Python is a typical example of an open-source software ecosystem, comprising a (relatively) closed core of developers providing the basic functionality and coordinating new releases, various developers contributing patches and bug fixes, numerous developers contributing packages (plugins) that extend the functionality way beyond that provided with each release, and a plethora of users. Other examples of such ecosystems include the Eclipse IDE and its third-party plugins, or the R/Ruby/LaTeX programming languages and their various contributed packages/gems. Second, R has been evolving for almost 20 years, and its entire history of mailing list communication is archived and publicly available. Third, Python has been the recent subject of an extensive study of its evolution. Fourth, Python promises to provide broader relevance outside the software developers’ community, as many users and contributors are data analysts from different domains such as economics or biology, with none or limited software engineering experience

**3.2 Architecture for the Proposed System**

The architecture proposed for the project, resembles the model for any other search engine. Fig. 3.2.1 shows an overview of the different components that exists in our search engine design. As a one-time offline process, the entire wiki corpus is fed into the indexer for the generation of the index table. A trained ‘Learning to Rank’ system serves as our ranking engine for ordering the results with respect to their relevance with the search query. The queries submitted by the users are sent to the search module for processing which communicates with the index database to pick the best possible matching results. These results are then sent to the ranking system to obtain an optimal order of display for them. The results are then displayed to the user in an interactive fashion.

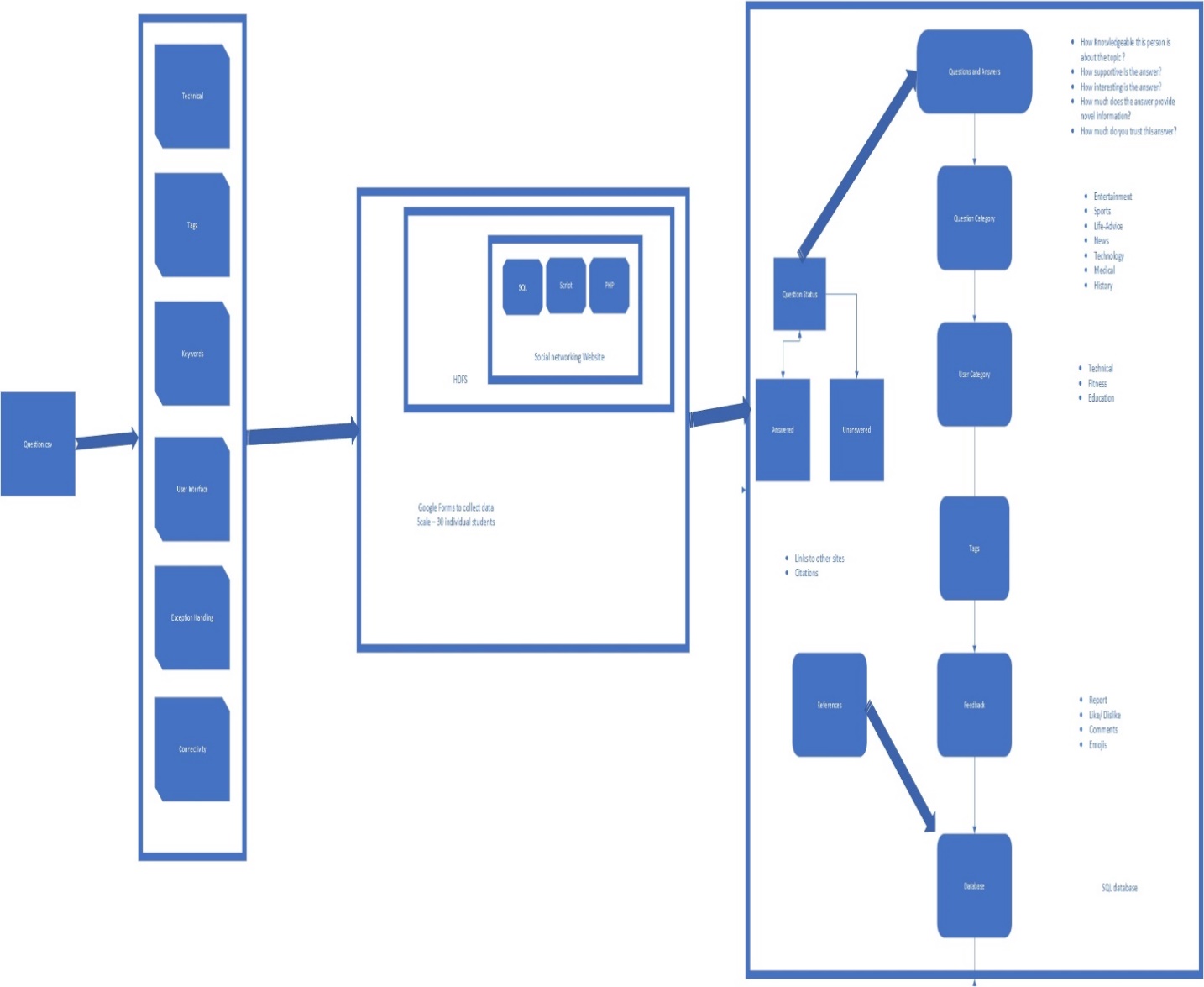
Databases are maintained to store the indexed data as well as features corresponding to different documents which serves as an input to the ranking algorithm. Our architecture supports the nature of an ideal search engine by returning query results in a very short period of time without iterating over the entire corpus.

* 1. **Proposed System Model**
     1. **Low Level Diagram**



**Figure 3.3.1 Low Level Diagram**

* + 1. **High Level Design**



**Figure 3.2.1 Architecture diagram**

**4. Proposed System Analysis and Design**

**4.1 Introduction**

We picked three social Q&A sites to study that offer similar Q&A interfaces, but that differ in the volume of contribution and membership: Yahoo Answers, Answerbag, and Ask Metafilter.\* These sites each offer an opportunity for users to ask questions on any topic for the community to answer. While there are other types of online forums for online question asking and answering, such as digital reference services and “ask an expert” sites, we do not consider these sites in this analysis, as they have a more restrictive Q&A process – relying on single “experts” to answer questions – and experience empirically different quality problems than social Q&A sites. For each of the three sites, we collected information over a range of dates, including full text, user identifiers, category names and identifiers, and timestamps. See Table 1 for summary statistics of this dataset. Yahoo Answers (answers.yahoo.com) is the largest Q&A site in the United States, claiming 74% of U.S. Q&A traffic . We downloaded data using the Yahoo Answers web API for a period of seven weeks, resulting in a data set of over 1 million users, 4 million questions, and 24 million answers. Remarkably, the Yahoo Answers community asked an average of 88,122 questions per day over the period of our data collection. Notable features of the Yahoo interface include questions that accept new answers for just 4 days (or at the asker's request, 8 days), browsing organized around categories, and a prominent system of rewarding users with “points” for answering questions. Answerbag (www.answerbag.com) is a much smaller Q&A site, with an estimated 4.5% market share. We downloaded data using the Answerbag web API, for a data set that spans 6 months of activity. We chose to collect data on questions asked in the first half of 2007, as Answerbag questions remain “open” indefinitely, and many questions continue to receive answers long after they are asked (in contrast with the short lifespan of questions at Yahoo). Answerbag distinguishes itself by sorting questions either by a user-rated “interestingness” metric, or by the last answer received. Also, Answerbag questions do not have a separate subject field, and are limited to just 255 characters. Finally, Answerbag answers may themselves be commented on – we exclude this data from our analysis to ensure consistency across the data sets. Ask Metafilter (ask.metafilter.com) is the smallest of our Q&A sites, with an estimated 1.8% share of U.S. Q&A traffic, and an average of 56 questions per day over the course of our study. We collected over 2 years of data from Ask Metafilter using a custom scraper. Questions on this site may receive answers for a year after they are asked; we stopped scraping questions newer than May 2007 to ensure completeness. Ask Metafilter requires participants to pay $5 to join the community, which has led to a much lower volume of contributions. Also notable is the fact that nearly every question (99.7%) receives at least one answer.

**4.2 Requirement Analysis**

**4.2.4 System Requirements**

**4.2.4.1 H/W Requirements**

* + Laptop with Broadband Connection
  + Laptop with at least 8 GB RAM and new generation processor

**4.2.4.2 S/W Requirements**

**-** SK Learn Library

- IDE for development of code

- Matplot Library for Plotting Graphs

**5. Library Used –**

**5.1 SKLearn –**

The scikit-learn project started as scikits.learn, a Google Summer of Code project by David Cournapeau. Its name stems from the notion that it is a "SciKit" (SciPy Toolkit), a separately-developed and distributed third-party extension to SciPy. The original codebase was later rewritten by other developers. In 2010 Fabian Pedregosa, Gael Varoquaux, Alexandre Gramfort and Vincent Michel, all from INRIA took leadership of the project and made the first public release on February the 1st 2010 Of the various scikits, scikit-learn as well as scikit-image were described as "well-maintained and popular" in November 2012.

**5.2 Numpy -**

NumPy is a library for the Python programming language, adding support for large, multi-dimensional arrays and matrices, along with a large collection of high-level mathematical functions to operate on these arrays. The ancestor of NumPy, Numeric, was originally created by Jim Hugunin with contributions from several other developers. In 2005, Travis Oliphant created NumPy by incorporating features of the competing Numarray into Numeric, with extensive modifications. NumPy is open-source software and has many contributors.

**5.3 Panda -**

In computer programming, pandas is a software library written for the Python programming language for data manipulation and analysis. In particular, it offers data structures and operations for manipulating numerical tables and time series. It is free software released under the three-clause BSD license. The name is derived from the term "panel data", an econometrics term for data sets that include observations over multiple time periods for the same individuals.

**5.4 Networkx -**

NetworkX is a Python package for the creation, manipulation, and study of the structure, dynamics, and functions of complex networks. Data structures for graphs, digraphs, and multigraphs. Many standard graph algorithms Network structure and analysis measures Generators for classic graphs, random graphs, and synthetic networks Nodes can be "anything" (e.g., text, images, XML records) Edges can hold arbitrary data (e.g., weights, time-series) Open source 3-clause BSD license Well tested with over 90% code coverage Additional benefits from Python include fast prototyping, easy to teach, and multi-platform

**5.5 Matplot Library –**

matplotlib is a plotting library for the Python programming language and its numerical mathematics extension NumPy. It provides an object-oriented API for embedding plots into applications using general-purpose GUI toolkits like Tkinter, wxPython, Qt, or GTK+. There is also a procedural "pylab" interface based on a state machine (like OpenGL), designed to closely resemble that of MATLAB, though its use is discouraged.[3] SciPy makes use of matplotlib.

**6. Results and Discussion**

**6. Conclusion, Limitations and Scope for Future Work**

**6.1 Conclusion**

In this undertaking, we present a novel way to deal with figuring a web index sans preparation, being able to transform it's positioning calculation as indicated by the model of the basic wiki information by utilizing 'Figuring out how to Rank' frameworks. We make utilization of the Coordinate Ascent calculation for preparing our model over the Wikipedia produced information, streamlining on the Expected Reciprocal Rank (ERR) metric, utilizing the RankLib library. The whole stage was produced on a dispersed frameworks engineering, taking into the thought the substantial size of the wikis. Through the course of this venture, we utilized different Amazon Web Services to satisfy our necessities. The Amazon S3 benefit was used for capacity of the Wikipedia information dump and other supporting antiques. The Hadoop employments were executed on the disseminated Amazon Elastic MapReduce (EMR) stage, utilizing 4 c3.2xlarge cases. The general runtime for the ordering and positioning occupations consolidated was around 32 hours. The execution of the created framework was tasteful as far as the outcomes delivered which bolsters our thought introduced in the venture of building a malleable internet searcher stage for inner wikis.

**6.3 Scope for Future Work**

In future, we intend to enhance the general precision of the framework by upgrading our list of capabilities with more number of highlights, for example, score, click tally and substance quality score. Having built up its viability, we additionally mean to coordinate our proposed framework with the current publicly released web search tools, for example, Apache Lucene and Sphinx.

**7. References**

**Journals:**

1. Anderson, A., Huttenlocher, D. P., Kleinberg, J. M., and Leskovec, J. Discovering value from community activity on focused question answering sites: a case study of Stack Overflow. In Proc. KDD, ACM (2012), 850–858.
2. Antin, J., and Churchill, E. F. Badges in Social Media: A Social Psychological Perspective. In Proc. CHI, ACM (2011)
3. Bagozzi, R. P., and Dholakia, U. M. Open source software user communities: A study of participation in Linux user groups. Management Science 52, 7 (2006), 1099–1115
4. Begel, A., Bosch, J., and Storey, M.-A. Social networking meets software development: Perspectives from GitHub, MSDN, Stack Exchange, and TopCoder. IEEE Software 30, 1 (2013), 52–66.
5. Bettenburg, N., Shihab, E., and Hassan, A. E. An empirical study on the risks of using off-the-shelf techniques for processing mailing list data. In Proc. ICSM, IEEE (2009), 539–542.
6. Bird, C., Gourley, A., Devanbu, P. T., Gertz, M., and Swaminathan, A. Mining email social networks. In Proc. MSR, ACM (2006), 137–143.
7. . Bird, C., Gourley, A., Devanbu, P. T., Swaminathan, A., and Hsu, G. Open borders? Immigration in open source projects. In Proc. MSR, IEEE (2007), 6.
8. Dabbish, L. A., Stuart, H. C., Tsay, J., and Herbsleb, J. D. Social coding in GitHub: transparency and collaboration in an open software repository. In Proc. CSCW, ACM (2012), 1277–1286.
9. Deterding, S. Gamification: designing for motivation. Interactions 19, 4 (2012), 14–17.
10. Deterding, S., Sicart, M., Nacke, L., O’Hara, K., and Dixon, D. Gamification: Using game-design elements in non-gaming contexts. In Proc. CHI, ACM (2011), 2425–2428.

**Weblinks:**

1. https://faculty.ist.psu.edu/jjansen/academic/jansen\_qanda\_wenwo.pdf
2. http://groups.csail.mit.edu/up/other-pubs/cscw2012-tiestrength.pdf
3. <http://ijesc.org/upload/5d637e6b6598ac8d0b207a3ded27ad85.Knowledge%20Sha> ring%20in%20the%20Online%20Social%20Network%20of%20Yahoo!%20Answers%20and%20its%20Implications.pdf
4. <https://bmcfee.github.io/papers/mlr.pdf>
5. <https://en.wikipedia.org/wiki/Non-functional_requirement>
6. <http://www2013.w3c.br/proceedings/p1341.pdf>
7. <http://web.cs.wpi.edu/~kmlee/pubs/patil16snam.pdf>
8. <http://www.cs.cmu.edu/~dmovshov/papers/asonam_2013.pdf>
9. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.330.9946&rep=rep1&type=pdf>
10. <https://pdfs.semanticscholar.org/18eb/866d28d59d60314122c15682f8457cce4f27.pdf>
11. <http://www.firstmonday.dk/ojs/index.php/fm/article/view/2182/2028>.