QWiki: Need for QnA & Wiki to Co-exist

Simran Setia
Indian Institute of Technology Ropar
Punjab, India
2017csz0001@iitrpr.ac.in

S.R.S. Iyengar
Indian Institute of Technology Ropar
Punjab, India
sudarshan@iitrpr.ac.in

Amit Arjun Verma
Indian Institute of Technology Ropar
Punjab, India
2016csz0003@iitrpr.ac.in

ABSTRACT

Access to knowledge has never been as easy and quick as it has been in the 21st century. With the advent of the Internet and crowd-sourced knowledge building portals such as Wikipedia, Stack Exchange, Quora, and GitHub, information is just a click away. It is interesting to observe that the crowd *builds* these information repositories and not the experts. Information accumulation on a wiki-like portal and discussions on the QnA forum function independently as collaborative knowledge building practices. There is a need to understand the best possible practices to acquire and maintain knowledge in such crowdsourced portals. In this paper, we introduce QWiki, a novel approach of integrating a wiki-like portal and a QnA forum, seeking the union of aforementioned independent collaborative practices. The experimental analysis demonstrates that QWiki helps in knowledge acquisition and knowledge building process. The proposed model highlights the importance of interaction between a wiki-like portal and a QnA forum.

CCS CONCEPTS

• **Human-centered computing** → **Collaborative and social computing**; **Collaborative and social computing systems and tools**; **Wikis**; *Activity centered design*;

KEYWORDS

crowd, knowledge building, Wiki Portal, QnA Forum, triggering

1 INTRODUCTION

The present age of the Internet has introduced a new dimension for collaboration. Virtual collaboration, aided by the Internet, has led to the widespread use of virtual collaborative practices like collaborative query solving and collaborative authoring of articles. These practices have resulted in the accumulation of community

knowledge in online Knowledge Building Portals (KBPs). The online KBPs like Quora[1], Reddit [2] are one-step solution to solve user's queries. One of the most popular among the QnA websites is Stack Exchange. Stack Exchange is a network of many QnA websites based on diverse topics like Stack Overflow [3], Super User, Ask Ubuntu. Being backed by one of the most active communities, 92% of the Stack Overflow questions are answered in a median time of 11 minutes [25]. On the other hand, the KBPs like Wikipedia[4] aid collaborative development of the articles based on various topics [11]. Wikipedia has gained immense popularity since its inception [2]. It is because it comprises of 40 million articles in 293 languages [45]. A wiki portal and a QnA forum are contrasting in nature in terms of the knowledge building process. The former builds knowledge as wiki articles, and the latter builds knowledge as question and answer threads. Though the nature of knowledge building practices in these two portals is contrasting, they serve as one of the prime platforms for realizing the knowledge building practices. However, prior research has pointed out many limitations of existing knowledge building practices. The discussions on online QnA forums grow to thousands of posts, which becomes difficult for the reader to comprehend [29, 47]. Also, many users encounter queries while reading articles on Wikipedia, indicating the existence of knowledge deficiencies in the articles [3, 4]. With a QnA forum in interaction with a wiki portal, the questions encountered in any wiki article can be discussed, and with a wiki portal in interaction with a QnA forum, the knowledge scattered in different threads can be archived. Discussion and archival, being the two modes of knowledge building, should go hand in hand. Any discussion can be archived, and any archive can be discussed. An ideal knowledge building practice should comprise of discussions as well as archives. Therefore, we propose the integration of a wiki portal and a QnA forum called QWiki, which is a combination of archives and discussions. In this paper, we show how a QnA forum and a wiki portal mutually benefit each other. We present an empirical study based on a QWiki portal, deployed for a MOOC. According to the results obtained, QWiki doesn't only aid the readers in knowledge acquisition but also aids the editors in building knowledge. The results and findings of the experiment are explained in the forthcoming sections.

The remainder of the paper is organized as follows. Section 2 reviews the literature supporting the integration of a QnA forum and wiki portal. Section 3 covers the details of the dataset collected from the QWiki portal. The hypotheses, the literature supporting the hypotheses, and the validation of the hypotheses using the

---

[1]https://www.quora.com
[2]https://www.reddit.com
[3]https://stackoverflow.com
[4]https://www.wikipedia.org

dataset are discussed in Section 4. Section 5 describes the coherence observed in discussion environments. Section 6 reviews the theoretical foundations of this study.

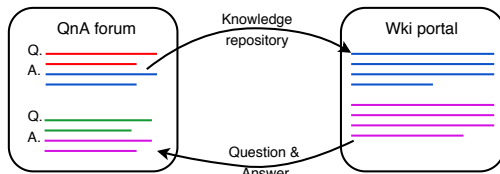## 2 IMPORTANCE OF INTEGRATION OF WIKI PORTAL AND QNA FORUM

Figure 1: Interplay between wiki Portal and QnA Forum

We call the integration of a QnA forum and a Wiki portal as "QWiki". In a QWiki portal, every wiki page has a corresponding QnA page. Since the QnA page is just a click away from the wiki page, whenever a user encounters a question in a wiki article, he can ask the question on the corresponding QnA page (as depicted in Fig. 1). Another advantage of a QWiki portal is that the knowledge generated in question and answer threads can also be easily archived to the respective wiki articles by the users (as depicted in Fig. 1). The pointers in the literature stating the importance of the integration are stated in the forthcoming subsections.

2.1 Wiki Portal in interaction with QnA Forum

2.1.1 Wiki for Knowledge Archival: In [38], authors prove that the appropriate answer to a question on Twitter is the sum total of discrete answers which are posted by a group of relevant users on that thread. Also, Zhang et al. developed a summarising tool called *Wikum*, which helps to summarise the discussions on a QnA forum [47]. The discussions on a QnA forum grow to thousands of posts, which becomes difficult for a reader to comprehend. So, a summarising tool like Wikum helps to bridge the gap between Wikis and discussion forums. Arkose is also a tool designed for online discussion forums which is used to filter, structure, and organize the information present in the discussion forums [29]. In addition to this, Answer-wikis help to summarise the discussions on a particular question on Quora [32]. It allows all the collaborators to agree on an aggregate answer to the question. The research mentioned above highlights the importance of presenting knowledge in an organized way. Knowledge archival in wiki articles leads to the addition of new knowledge to the wiki portal. This process of curation of new knowledge artifacts as a result of discussions among individuals is called *Knowledge Building*. Also, knowledge archival to respective wiki articles helps the users to revisit the relevant information efficiently. The knowledge scattered in various question and answer threads in a QnA forum should also be stored in an organized way. So, we claim that a dedicated wiki page in interaction with the QnA forum would serve as an apt platform for archiving the knowledge present in the QnA forum in an organized way.

2.1.2 Wiki as Knowledge Base: Researchers have exhibited the possibility of conversational agents as a human-computer interaction tool to access the knowledge present on Wikipedia for question-answering task [7]. In [37], authors use Wikipedia as a knowledge base to find potential answers to questions asked on Yahoo! Answers. Also, many users from Stack Exchange [42], Quora[5] point to Wikipedia articles in their responses to some questions. So, a wiki portal acts as a knowledge base to answer questions present on a QnA forum.

Apart from the points mentioned above, there are many other ways through which a wiki portal proves to be beneficial to a QnA forum [19]. CSS-D is a public mailing list devoted to discussions on CSS(Cascading Style Sheet). Also, a wiki portal is deployed by the administrators which act as a shared repository, a user-defined FAQ, and dumping ground for CSS tricks for the CSS-D community. According to the authors, the wiki portal witnessed 10 times more visitors than the subscribers of the mailing list. The wiki portal led to fewer basic, repeat questions in the mailing list. It even encouraged the list members to summarise their discussions on mailing lists. Also, referring to the wikis in the mailing list improved the efficiency of the question-answering task. Hence, a wiki portal in interaction with a QnA forum is essential for the maintenance of a QnA forum.

2.2 QnA Forum in interaction with wiki Portal

2.2.1 QnA Forum for Knowledge Acquisition: The information present on Wikipedia is as exhaustive as any other encyclopedia [14], and the majority of the users are involved in reading the information present on Wikipedia [40]. But, there are significant concerns about the readability of articles on Wikipedia [23]. The readability of a text is associated with the ease with which a reader can understand the text. Since wiki articles score low on readability metrics [23], the readers are unable to acquire the knowledge given in the wiki articles. This process of understanding the knowledge present in the wiki articles is called *Knowledge Acquisition*. In order to tackle the issue of poor knowledge acquisition due to low readability, a QnA forum in interaction with a wiki article would serve as an apt platform for knowledge acquisition. In history, we find instances of people relying on questioning as an effective way to learn a concept. Socratic questioning is about discovering answers by asking questions from the students [31]. Socratic questioning probes thinking at a deep level and promotes the understanding of a concept or idea. Similarly, we claim that a QnA forum helps to develop a better understanding of the knowledge present in the wiki article.

2.2.2 QnA Interface vs. Talk Pages Interface: The evaluation results reported by Schneider et al. term the talk page interface as confusing [35]. The QnA interface does affect the conversational behavior of the users [43], which calls for an alternative interface, more sophisticated than talk pages for discussions in a wiki portal. The QnA interface used in our QWiki portal supported structured discussions, having dedicated discussion thread for each of the questions.

---

[5]https://www.quora.com/Why-does-science-not-give-people-a-cure-for-cancer
This is one of the many instances where crowd points to Wikipedia article
as an answer to a question in Quora.

2.3 Importance of a Discussion Environment

Every Wikipedia article is accompanied by a talk page. According to the Wikipedia guidelines, a talk page is a discussion space that can be used by the editors of the article to discuss the improvements related to the article [44]. Even after the limited scope of discussions on the talk page, high-quality articles on Wikipedia have dense talk pages [39]. It is observed that not only is the length of the featured articles ten times larger than its counterparts, but they are also well organized and easy to comprehend [39]. Moreover, the average number of edits on Wikipedia articles is 15, and the articles with more than 100 edits have talk pages. Conversely, the articles with talk pages have more number of edits and users as compared to the articles with no talk pages [41]. It is apparent from the above discussion that the talk pages serve instrumental in the knowledge growth in Wikipedia articles. Also, it has been observed that heavily edited articles and talk pages go hand in hand [41]. We call this phenomenon as coherence between Wikipedia articles and talk pages, discussed in the next section.

2.4 Coherence in Wikipedia

In Wikipedia, this coherent activity between articles and the corresponding talk pages has been observed by Viegas et al [41]. In order to prove this fact, we analyze 100 random Featured Articles of Wikipedia and their corresponding talk pages for a period of 2000 days. The dataset of 100 Featured Articles along with their talk pages is downloaded with the help of WikiTalker library [6]. As we can observe in Fig. 2, the activity in Wikipedia articles conform to the activity observed in talk pages. After the inclusion of the talk page, the discussions among the editors related to the article can be realized. The finer details of the coherent activity across Wikipedia articles and respective talk pages can be observed in Fig. 3. The coherence in the activity of Wikipedia articles and respective talk pages is measured with the help of the Wilcoxon rank-sum test [26]. It is used to test whether two samples follow the same distribution or not. This test is applied to the samples obtained from Wikipedia articles and their respective talk pages. The two samples in consideration are the array of the number of sentences added to the 100 Featured Articles, 1000 days post to the inclusion of the talk page, and the number of sentences added to their respective talk pages, 1000 days after the creation of talk page. According to the results obtained, we observed a higher p-value for most of the samples obtained from the Featured Articles and their corresponding talk pages. The average of the array of p-values is 0.80 with standard deviation of 0.36. A higher p-value in Wilcoxon rank-sum test suggests that there is no difference in the distribution of the samples obtained. This, in turn, suggests that there is no much difference observed in the Wikipedia article and talk page activity. Hence, the coherence in the activity of Wikipedia articles and their respective talk pages establishes the importance of discussion forums-like talk page for a Wikipedia article. But, it should be noted that only 27% of the Wikipedia articles have talk pages [20]. For the ones which have a talk page, the talk page is included at a very later stage after the first article edit (as depicted in Fig. 4). The difference in the first talk page edit and first article edit ranges from 1000 to 3000 days.

This huge time-lapse in the inclusion of the talk page represents the downside of having a restricted discussion forum like talk pages.

It is clear from the above observations that the potential of a discussion platform like talk pages hasn't been harnessed properly. That is because of the limited scope of discussions on talk pages. In contrast, a QnA forum can be used for any sort of discussion ranging from conceptual doubts to discussion about the structure of the article. Thus, a platform like QWiki would materialize as a great addendum to the collaborative dynamics on Wikipedia by revamping the discussion space.

According to the above-stated points present in the literature, a QnA forum helps the users to acquire knowledge present in the wiki article, whereas a wiki portal helps to archive the knowledge present in the QnA forum. In this respect, to prove the claim that the integration of the Wiki portal and QnA forum helps to reap the benefits mentioned above, we deployed a QWiki portal on a MOOC. The dataset collected in the form of wiki articles and question and answer threads is used to prove or disprove this claim.

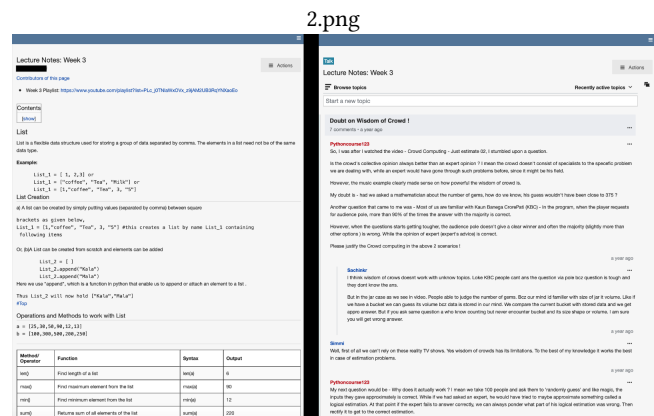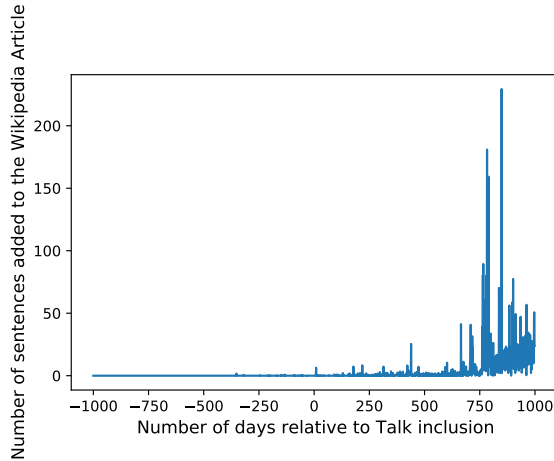## 3 JOCWIKI: THE QWIKI PORTAL

Figure 5: Screenshot of JOCWiki2 Week 3 article and its corresponding QnA forum. The left-hand side of the image represents the Week 3 wiki article, and the right-hand side represents the QnA forum of Week 3. The QnA forum is directly accessible from its corresponding wiki article in JOCWiki. (The picture is blurred respecting the blind review policy.)
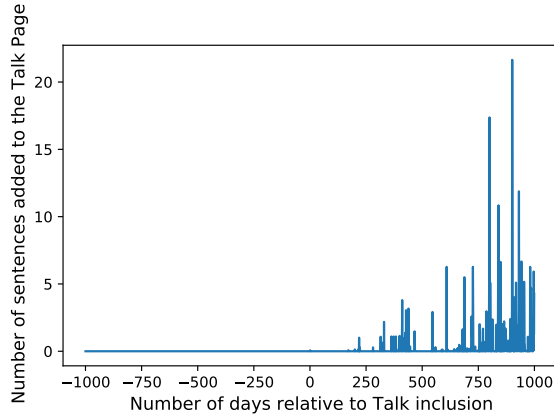
A QWiki portal named JOCWiki is deployed for a MOOC "The Joy of Computing using Python" in the field of Computer Science. The MOOC is about learning Python programming. The course contents are divided into 12 weeks. The JOCWiki portal is deployed using Mediawiki Software [27]. Mediawiki is an open-source wiki engine that supports the collaborative editing of the web pages. The wiki website deployed by the Mediawiki software has the conventional talk page in interaction with every wiki page. We use Mediawiki extension, Structured Discussions, which converts conventional talk pages to QnA pages [28]. Like a talk page in Wikipedia, every wiki article is accompanied by a QnA page (as shown in Fig. 5) in JOCWiki. The dataset in consideration is from the JOCWiki portal[7].

---

[6]https://github.com/descentis/WikiTalker

[7]https://github.com/SimranIITRpr/QWiki-Dataset-

(a)

(b)

Figure 2: The above figure is plotted for 100 featured articles. The analysis is done for 1000 days prior to Talk Page inclusion and 1000 days post Talk Page inclusion. The average number of sentences added over 100 articles on each day is plotted against the number of days relative to Talk Page inclusion in (a). The average number of sentences added over 100 talk page articles on each day is plotted against the number of days relative to Talk Page inclusion in (b). The average of the number of sentences is taken as zero prior to the talk page inclusion.

These wiki articles are developed by the course students based on the course videos of a week. Also, the QnA page of a particular week's article is used by the students for conceptual queries related to the course. With over 2,000 students (age group of 15 years to 45 years) registered on the portal, the Course Instructor and the Teaching Assistants of the course monitor the portal for active participation of students. We deploy the JOCWiki portal for the fall semester of 2018 (JOCWiki1) and the spring semester of 2019 (JOCWiki2). In the latter, we start with a new JOCWiki portal with no data from the previous run.
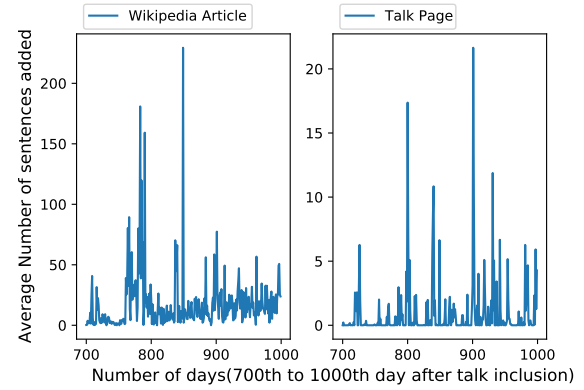

Figure 3: The above figure is plotted for 100 featured Wikipedia articles and their respective talk pages.
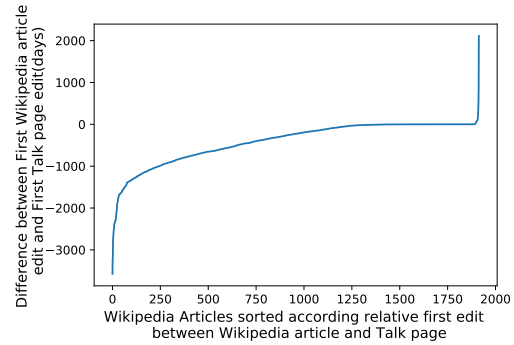

Figure 4: The time difference (First wiki article edit- First Talk page edit) is negative for most of the articles suggesting the late inclusion of talk pages to Wikipedia pages. The above figure is plotted for around 1900 randomly chosen Featured articles.

4 COEXISTENCE OF A WIKI PORTAL AND A QNA FORUM

The poor readability of Wikipedia articles makes Wikipedia more of a knowledge building tool and less of a knowledge acquisition tool [23]. On the other hand, the QnA forum helps the users to acquire knowledge quickly with no provision to archive the knowledge. We analyze the JOCWiki data from knowledge acquisition perspective as well as knowledge building perspective in the forthcoming sections.

4.1 QWiki: The Knowledge Acquisition Perspective

The poor readability of Wikipedia articles is an obstacle to knowledge acquisition. The commonly used readability metric, Flesch-Kincaid readability test [15], outputs the number of education years required to understand a piece of text. A piece of information that is easy to understand, say, for a college graduate may be difficult for a class 9 student. But, adhering to the nature and scope of Wikipedia

articles, it isn't very easy to maintain desired readability for all users with different educational backgrounds. Reading being the most common activity on Wikipedia establishes the need for curating and maintaining Wikipedia articles with desired readability scores [40]. One possible solution to seek information and understand the underlying piece of text is using a QnA forum. Wikipedia provides a discussion platform in the form of talk pages. However, only 10% of the total posts on talk pages are by the readers seeking information [41]. We analyze the data of JOCWiki portal to show that integrating a QnA forum with a wiki portal leads to knowledge acquisition.
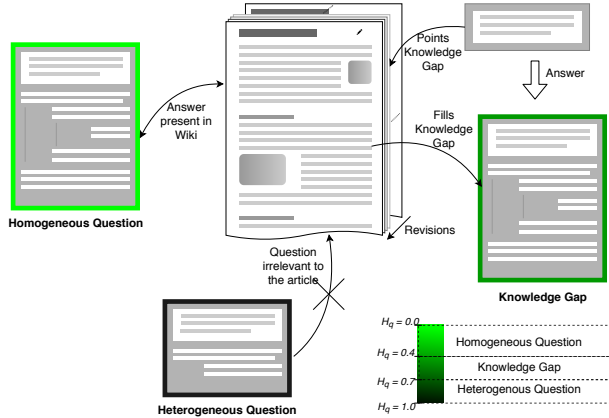

Figure 6: Classification of questions in JOCWiki.

The questions on QnA forum on JOCWiki are classified as *Homogeneous questions, Knowledge gaps*, and *Heterogeneous questions* (as depicted in Fig. 6). Homogeneous questions are those questions that are based on the contents of the article, such that the answers can be directly pointed out in the wiki articles. An instance of homogeneous questions is shown in Fig. 7. On the other hand, questions for which answers can not be pointed out in the articles, despite them being related to the contents of the article, are termed as Knowledge gaps. An instance of the knowledge gap is shown in Fig. 8. The third category of questions that are not associated with the contents of wiki articles is termed as Heterogeneous questions. The examples of heterogeneous questions include those related to assignments and admin related queries such as, "Can you please explain programming assignment?", "What is the answer to the quiz question 6?", and "Is the quiz question 5 correct?".

In order to determine the homogeneous questions, knowledge gaps, and heterogeneous questions, the wiki articles in JOCWiki are divided into segments using semantic word embeddings [1]. Text segmentation is used to split the text in wiki articles into contiguous coherent sections. We used LDA (Latent Dirichlet Allocation) to find *topics* of each question and answer threads in the QnA forum and each *segment* in the corresponding wiki article [6]. For a piece of text, LDA outputs a set of keywords along with their probability distribution. This set of keywords, along with probability distribution, is a *topic* inferred through LDA. The LDA topic model is implemented with the help of the Gensim python library [33]. It is

important to note that LDA is not an apt topic modeling technique for texts shorter than 30 words [46]. The underlying segments of the wiki articles are at least 6 sentences long, which roughly amounts to 100 words per segment [12]. Also, the average number of words across all question and answer threads of JOCWiki1 is 1366.1 and that of JOCWiki2 is 1140.4. Therefore, each question and answer thread in the QnA forum and each segment in the corresponding wiki article is associated with the probability distribution of the keywords. Hellinger distance [5] is the probabilistic analog of Euclidean distance and is used to calculate the similarity between the probability distributions of a question and answer thread and each text segment of the corresponding wiki article. To classify a question and answer thread q to one of the classes mentioned above, the average Hellinger distance H_q of the question and answer thread is calculated across all the segments present in the wiki article as follows :

$$H_q = \frac{1}{t_s} \sum_{j=1}^{t_s} h_d(P_q, Q_{s_j}) \; ; \; h_d(P_q, Q_{s_j}) = \frac{1}{\sqrt{2}} \| \sqrt{P_q} - \sqrt{Q_{s_j}} \|_2$$

where $h_d(P_q, Q_{s_j})$ is the Hellinger distance between probability distribution P_q of question and answer thread q and probability distribution Q_{s_j} of jth segment of the wiki article, t_s is the total number of segments of the wiki article. The smaller the Hellinger distance, the more similar the probability distributions are.

The classes of question and answer threads are identified as follows:

$$\begin{cases} 0 < H_q < 0.40 \implies q \text{ is Homogeneous question} \\ 0.40 \leq H_q \leq 0.70 \implies q \text{ is Knowledge gap} \\ 1 > H_q > 0.70 \implies q \text{ is Heterogeneous question} \end{cases}$$

All the above boundary values of H_q are defined empirically. Low Hellinger distance implies similar probability distributions, which in turn implies a higher degree of similarity between LDA topics. Similar LDA topics highlight the fact that the underlying question and answer thread in the QnA forum is similar to the segment in the wiki article. Therefore, the question is classified as a homogeneous question. Similarly, if there is a higher Hellinger distance between probability distributions, the question is classified as a heterogeneous question. The Hellinger distance ranging from 0.40 to 0.70 is used to classify the questions as knowledge gaps. This range implies that the probability distributions are not as similar as homogeneous questions and not as dissimilar as heterogeneous questions. The above range suggests that the underlying question and answer thread is based on the underlying segment of the wiki article but is not entirely unrelated to the segment. The question is thus, classified as a knowledge gap.

Considering the question and answer threads in the QnA forums for 12 weeks of JOCWiki1 and JOCWiki2, we observe that the number of homogeneous questions across JOCWiki1 and JOCWiki2 (1.27 %) is way lesser than heterogeneous questions and knowledge gaps (as depicted in Fig. 9). Knowledge gaps across JOCWiki1 and JOCWiki2 (87.83 %) followed by heterogeneous questions contribute as the major chunk of question and answer threads.

```
Complete Code:
    import speech_recognition as sr
    audio_file = ('C:/test_file.wav')
    r = sr.Recognizer()
    with sr.AudioFile(audio_file) as source:
        audio = r.record(source)
    try:
        print("audio file contain" + r.recognize_google(audio))
    except sr.UnknownValueError:
        print("Google speech recognizer cant understand your audio")
    except sr.RequestError:
        print("Couldnt get the result from Google Speach Recognition")
```

Speech Recognition
2 comments · a year ago

Vasik
My program is import speech_recognition as sr AUDIO_FILE={"test.wav"} #Use audio file as source r=sr.Rec
with sr.AudioFile(AUDIO_FILE) as source: audio=r.record(source) #reads audio file try: print("Audio file contai
r.recognize_google(audio)) except sr.UnknownValueError: print("Google Speech Recognizer cound not unde
sr.RequestError: print("Couldn't get results from Google Speech Recognizer") The error displayed after exec
sr.AudioFile(AUDIO_FILE) as source: File "C:\Users\admin\Miniconda3\lib\site-packages\speech_recognition
__init__ assert isinstance(filename_or_fileobject, (type(""), type(u""))) or hasattr(filename_or_fileobject, "read"
filename string or a file-like object" AssertionError: Given audio file must be a filename string or a file-like obj

Figure 7: The above figure shows an example of a homogeneous question from JOCWiki1. In this example, the user is asking about an error in a program that has been clearly stated in the wiki article. Although the program posted by a user is exactly the same as the one in the wiki article, the user is unable to figure out the curly braces error in his program. (The usernames have been anonymized.)


```
Basic Programs:
Temp_Dictionary = {}
    for i in range (10):
        Temp_Dictionary[i] = i**2
    print(Temp_Dictionary)
Output: {0: 0, 1: 1, 2: 4, 3: 9, 4: 16, 5: 25, 6: 36, 7: 49, 8: 64, 9: 81}

Temp_Dictionary = {}
Text = "Dictonary"
Text = list(Text)
for i in range (len(Text)):
    Temp_Dictionary[i] = Text[i]
print(Temp_Dictionary)
Output: {0: 'D', 1: 'i', 2: 'c', 3: 't', 4: 'o', 5: 'n', 6: 'a', 7: 'r', 8: 'y'}
```

Regarding dictionary to start with index 1 rather than 0
5 comments · a year ago

Vikram
In this code if we want output as {1: 1, 2: 4, 3: 9, 4: 16, 5: 25, 6: 36, 7: 49, 8: 64, 9: 81}

then what should be done?

```
Temp_Dictionary = {}
    for i in range (10):
    Temp_Dictionary[i] = i**2
    print(Temp_Dictionary)
Output: {0: 0, 1: 1, 2: 4, 3: 9, 4: 16, 5: 25, 6: 36, 7: 49, 8: 64, 9: 81}
```

Figure 8: The above figure represents an example of a knowledge gap from JOCWiki1. In the wiki article, a program to initialize a dictionary starting from 0 has been added. But, the user wants to initialize the dictionary starting from 1 rather than 0.
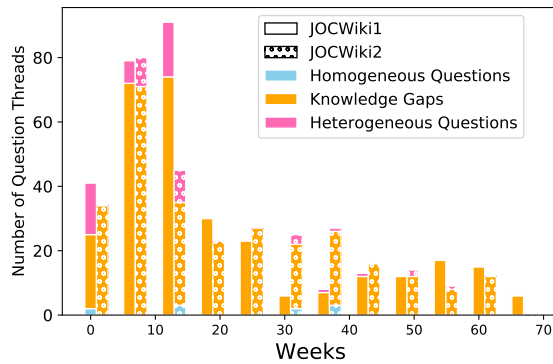

Figure 9: Distribution of question and answer threads in QnA forum of JOCWiki1 and JOCWiki2.

4.2 QWiki and Wikipedia + Stack Exchange

In order to capture the interaction between wiki articles and QnA forum in QWiki, three sets of keywords are identified. The keywords are extracted from LDA topics of week 4 wiki article and QnA forum (as depicted in Fig. 10a). The sets are as follows:

- Set of keywords present in the wiki article. (A)
- Set of keywords present in the concerned QnA forum. (Q)
- Set of keywords present in both wiki article and QnA forum. $(A \cap Q)$

In the same context, we analyzed the Wikipedia articles and Stack Exchange question and answer threads of the same topic to determine the similar and dissimilar keywords. It is very much possible for the users to ask questions encountered in Wikipedia articles on QnA forums like Stack Exchange. In order to capture the interaction between QnA forums like Stack Exchange and Wikipedia articles, we analyzed the Wikipedia articles and Stack Exchange question and answer threads on the topics of week 4 articles. Week 4 in JOCWiki comprised of Birthday Paradox, Guess the movie name, Magic Square, and Dobble game. Since there is no Wikipedia article on Dobble game, LDA keywords are obtained for Wikipedia articles on Birthday Paradox, Guess the movie name, and Magic Square. Apart from the Wikipedia articles mentioned above, LDA is applied to the question and answer threads found in Stack Exchange based on Birthday Paradox, Guess the movie name, and Magic Square. The above question and answer threads are extracted from Stack Exchange data dump using the tags (based on the topics mentioned above) like, Birthday Paradox, Guess the movie name, Hangman, Magic Square. The Week 4 wiki article entry, as well as the question and answer threads on Dobble game, are deleted from the respective wiki article and QnA forum. LDA is then applied to the Week 4 article and QnA forum for a fair comparison. The average proportion of intersecting keywords across 3 Wikipedia articles and related question and answer threads on Stack Exchange is 0.032. The average proportion of intersecting keywords across Week 4 articles and their respective QnA forums (excluding the entries related to Dobble game) in JOCWiki1 and JOCWiki2 is 0.17. This is because unlike the keywords encountered in Week 4 in JOCWiki

(as depicted in Fig. 10b), the common keywords in Wikipedia and Stack Exchange are mostly the topics mentioned above on which articles and question and answer threads are based such as magic, square, birthday, guess, and letter (as depicted in Fig. 10a, Fig. 10b).
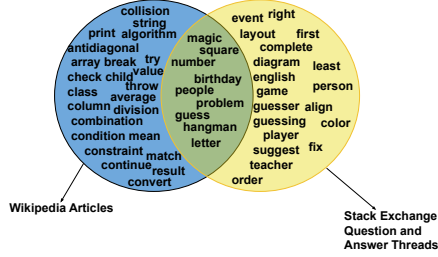
(a) The figure above represents A i.e. topics present in the Article, Q i.e. topics present in QnA forum and $A \cap Q$ for Week 4 Article and QnA page in JOCWiki
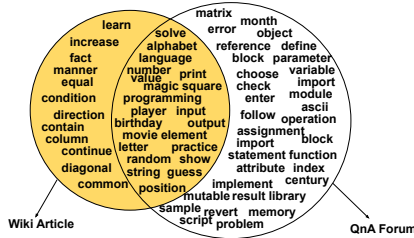
(b) The keywords represented here are obtained through LDA on question and answer threads and articles related to Birthday Paradox, Guess the movie name, and Magic Square in Stack Exchange and Wikipedia respectively.

Figure 10: Interaction in QWiki and between Wikipedia and Stack Exchange

As QWiki provides a platform in the form of a QnA forum to ask questions encountered in the corresponding Wiki articles, there is an interaction between the QnA forum and the Wiki portal. Due to no interaction between Stack Exchange and Wikipedia, the ratio of common keywords encountered in Stack Exchange and Wikipedia setup is lesser than that of the QWiki setup. This is because the questions encountered in Wikipedia articles are not posted by the users on an independent QnA forum like Stack Exchange. This establishes the fact that a QnA forum in interaction with Wiki portal helps a user to ask questions encountered in wiki articles which in turn helps to acquire the knowledge present in Wiki articles.

As per the discussion above on knowledge acquisition perspective of QWiki, homogeneous questions and knowledge gaps in a QnA forum help the user to acquire knowledge in the corresponding wiki article. Hence, a QnA forum in interaction with the wiki portal makes an effective knowledge acquisition tool.

4.3 QWiki: The Knowledge Building Perspective

Knowledge building is the pooling of ideas through collaboration among individuals. New knowledge artifacts are created as a result of the pooling of ideas [18]. Our focus is on one of the prime platforms for knowledge building, Wikipedia. One of the reasons behind knowledge evolution on Wikipedia is triggering [9]. As an instance of triggering, in the "New York City" article of Wikipedia, the wiki link [8] of *financial center* is created in the seventh revision, and subsequently, in the twelfth revision, wiki link of *New York Stock Exchange* is added. These two wiki links added in neighboring revisions are semantically similar. The inclusion of one wiki link triggered the addition of another wiki link. This process by which an idea or a piece of information cascades the generation of more ideas is called Triggering [24, 30]. Similar traces of triggering in a Wikipedia article are also found by Rezgui et. al [34]. Triggering claims that an idea is a backdrop of the next idea encountered in a knowledge building setup.

The triggering phenomenon is also observed in the QWiki environment. The instances of triggers are observed from the *QnA forum to wiki article* (Q→A triggers) and *the wiki article to QnA forum* (A→Q triggers). Based on the semantic similarity of knowledge units[9] added across the QnA forum and wiki article, we have further classified the triggers as *Primary Triggers* and *Secondary Triggers*. To understand the effect of triggering on knowledge building in JOCWiki, we have considered sentences as knowledge units. In order to quantify the triggers, the semantic similarity of each sentence present in the wiki article is measured with each sentence of question and answer threads present in the QnA forum. The semantic similarity is measured using sentence2vec [22]. The timestamp of sentence addition is used to infer if the trigger is Q→A trigger or A→Q trigger. Since the cosine similarity is used to find the similarity between two sentence vectors, the semantic similarity score varies from 0 to 1. We are considering highly similar sentences as triggers, having a similarity score ranging from 0.80 to 1. Following are the details of the semantic similarity score of Primary Triggers and Secondary Triggers:

- **Primary Triggers:** The triggers across the QnA forum and wiki portal, which have cosine similarity ranging from 0.90 to 1, are termed as primary triggers. If a knowledge unit is added in one of the platforms of QWiki (either QnA forum or wiki article), and a similar knowledge unit is added in the other platform, then it is a Primary Trigger. Examples of primary triggers in JOCWiki1 and JOCWiki2 are shown in Table 1.

- **Secondary Triggers:** The triggers across the QnA forum and wiki portal, which have cosine similarity ranging from 0.80-0.90, are termed as Secondary Triggers. If a knowledge unit is added to one of the platforms (either QnA forum or wiki article) and another knowledge unit similar to the previously added knowledge unit but with additional information, is added to the other platform, then it is a Secondary Trigger. Examples of secondary triggers in JOCWiki1 and JOCWiki2 are shown in Table 2.

[8]Wiki links are the links which lead to other Wikipedia articles.
[9]A word, phrase, sentence, paragraph or a document is considered to be a knowledge unit.

| Wiki Articles	QnA Forum	Type	Cosine Similarity
There are no prerequisites required for the course.	What is the meaning of word 'prerequisites'?	A->Q	0.98
Troubleshooting Spyder: In most of the cases, it happens that Spyder is already running and you have initiated one more instance of it.	But when i try to launch spyder, it is not perfectly launching and it appears as black screen.	Q->A	0.99
Slicing the list with negative values. In [7]:student[-1], Out[7]: 'Janani'	How to print list with negative indices?	Q->A	0.97
Sorting a list: Example,ages = [11,23,12,23,45,43,22,24,23],ages.sort(), print (ages).	print(ages.sort()) is not working	Q->A	0.90
The above code on Magic Square shows error list out of index.	Magic Square Error: IndexError: list index out of range.	Q->A	0.99

Table 1: Examples of Primary triggers in JOCWiki1 and JOCWiki2.

Wiki Articles	QnA Forum	Type	Cosine Similarity
Spyder IDE is divided into two screens, one is console and other one is where program is written.	I am getting a black screen after opening spyder, can u help me with the solution?	Q->A	0.81
For loop and while loop are used extensively to execute a set of looping instructions.	If a piece of code has a while loop in it and when an if loop is entered and a syntax error is thrown	A->Q	0.86
To start if we assume a 3x3 square then the sum of the elements must be equal to 15 in any direction - either in the vertical horizontal direction or even diagonally.	Can someone please explain how to make a matrix in a Python program?	Q->A	0.85
This code works for any number of doors.	There are variations of Monty hall problem for n doors.	Q->A	0.81
In anagrams, there is one root but there are many variations of the same word.	If the sum of ASCII values of two strings is then these two strings are anagrams.	A->Q	0.87

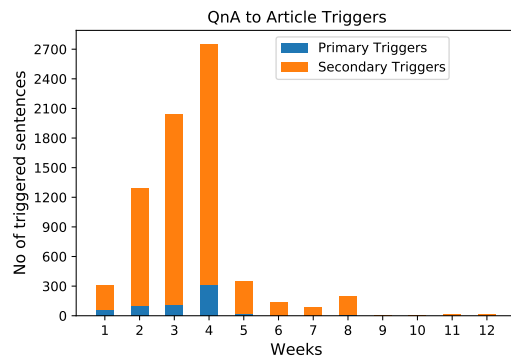Table 2: Examples of Secondary triggers in JOCWiki1 and JOCWiki2.

(a)
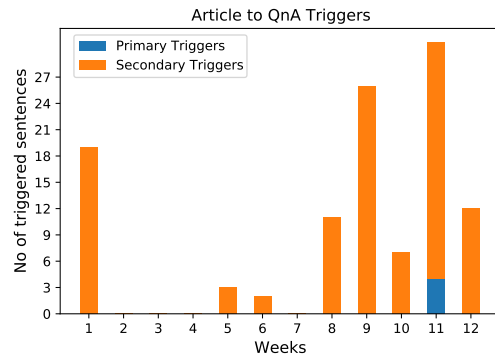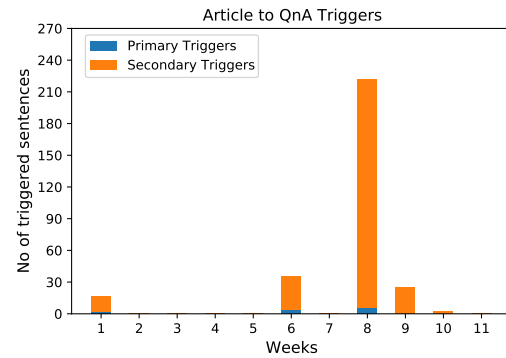
(b)

Figure 11: The total number of triggered sentences from article to QnA forum (a) as well as from QnA to article (b) due to primary ans secondary triggers in JOCWiki1.
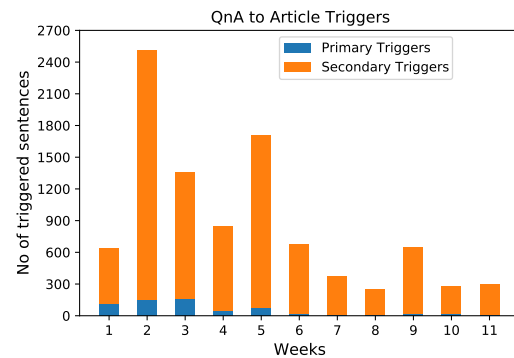

(a)

(b)

Figure 12: The total number of triggered sentences from article to QnA forum (a) as well as from QnA to article (b) due to primary ans secondary triggers in JOCWiki2.

As depicted in Fig. 11 and Fig. 12, it is observed that Q→A triggers are more than A→Q triggers for JOCWiki1 and JOCWiki2. This shows that a QnA forum facilitates the generation of more triggers and hence, helps in the process of knowledge building. In Fig. 11, we also observe that the number of triggered sentences from the QnA forum to wiki article for week 2 to week 4 is relatively more than other weeks, whereas the number of triggered sentences from wiki article to QnA forum is relatively lesser. If the Q→A triggers are relatively higher, this highlights that knowledge is triggered in a wiki article through discussions on the QnA forum. Also, if A→Q triggers are relatively lesser, we infer that there are fewer chances of the article being knowledge deficient. Hence, if there are sufficient Q→A triggers, then it leads to better knowledge building in the wiki article and lesser A→Q triggers. On the contrary, considering week 8 to week 12 triggers, as shown in Fig. 11, if there are relatively lesser number of Q→A triggers, this leads to knowledge deficiency in the article giving scope for more A→Q triggers.

4.3.1   Talk Page Triggers. In Wikipedia, a conceptual question cannot be asked on talk pages, while there is scope for suggestions regarding improvements of the existing article. While discussing the improvements related to the article, the information on talk pages is translated to the corresponding wiki article. This is regarded as *talk page to wiki article* trigger or T→A trigger. When some questionable information is added to the article, it triggers a discussion on the talk page. This is regarded as *wiki article to talk page* trigger or A→T trigger. In a nutshell, the interaction between wiki articles and talk pages generates triggers.

As shown in Table 3, instances of primary and secondary triggers are also found in Wikipedia. The topics "Area Calculation", "NetworkX", "Natural Language Processing", and "Six degrees of separation" are a part of a single article in week 9 of JOCWiki. Whereas, on Wikipedia, they exist as four exclusive articles. We compare the triggers across wiki article and QnA forum in the JOCWiki portal and across the wiki article and Talk pages of the corresponding topics on Wikipedia. We choose these topics in particular from our course JOC due to the common existence of the articles on JOCWiki and Wikipedia. Due to the exclusivity of the topics in course, we could not compare the rest of the weeks' articles.

As of March 2020, there are zero triggers across the NetworkX wiki article and talk page due to the absence of the talk page. For the rest of the articles, there are a considerable amount of sentences generated through triggers. Due to the varied size of the articles, we normalized the triggers as the ratio of the number of triggered sentences to the total number of sentences present in the article. According to the results obtained, the average ratio of the number of sentences triggered in the Wikipedia articles mentioned above is 0.239, and the number of triggered sentences in the corresponding talk page is 0.317 (as depicted in Table 3). On the other hand, the fraction of the number of sentences triggered in the week 9 article of JOCWiki is 0.31 and the number of triggered sentences in the corresponding QnA forum is 0.365 (as depicted in Table 3). Although the fraction of the number of triggered sentences in wiki

---

[10]Lifetime is the timestamp of the last edit of the article minus first edit of the article

article of Wikipedia and JOCWiki portal is approximately same, the lifetime[10] of the articles on Wikipedia is 9.5 years, and JOCWiki portal is 24 days. The ratio of triggered sentences in the JOCWiki article is achieved in a period of 24 days, whereas the ratio of triggered sentences in the Wikipedia article is achieved in 9.5 years. The difference in proportion of triggered sentences due to $A \rightarrow Q$ and $A \rightarrow T$ triggers is because the talk page is mostly used by the editors to discuss the improvements in the article whereas the QnA forum is used by all types of users (editors and readers) to discuss the concepts as well as improvements to the article. Also, the previous studies suggest that very few article edits are associated with the discussions on talk pages [34]. This, in turn, points towards the fact that there are very few triggers generated by talk pages in Wikipedia. Hence, a QnA forum integrated with a wiki portal provides a more conducive environment to induce triggers.

Triggers act as pointers for knowledge addition in a crowdsourced knowledge building portal. They play an essential role in identifying the information to be added to the article. We observed that integrating a wiki portal with a QnA forum helps in the generation of triggers, leading to better knowledge building prospects in a QWiki environment.
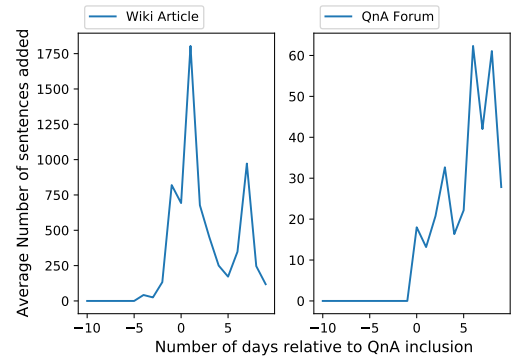
5   COHERENCE IN QWIKI

Figure 13: The above figure is plotted for JOCWiki1 data. The analysis is done for 10 days prior to QnA forum inclusion and 10 days post QnA forum inclusion. It is performed over all 12 articles of JOCWiki1 and the average number of bytes added on each day is plotted against the number of days relative to QnA forum inclusion.

As per the discussion in the previous section, it is observed in JOCWiki1 and JOCWiki2 that the knowledge generated in either of the platforms (either QnA forum or wiki portal) acts as a trigger for the generation of knowledge in the other platform. As an implication of triggering, *coherence* is observed between wiki article activities and QnA forum activities. As an instance, whenever there is an addition of sentences in the QnA forum, there is an addition of sentences in the wiki article or vice-versa. The coherent activity in Wikipedia has already been discussed in section 3. After the inclusion of the talk page, the discussions among the editors related to the article can be realized. These discussions act as triggers to add more knowledge to the Wikipedia article. The talk page aids

Article	A→T/Q Triggers		Proportion of triggered sentences	Q/T→A Triggers		Proportion of triggered sentences
	Primary	Secondary	due to A→T/Q Triggers	Primary	Secondary	due to Q/T→A Triggers
Area	4	70	0.41	1	14	0.096
NetworkX	0	0	0	0	0	0
NLP	13	69	0.36	1	124	0.59
Six Degrees of Separation	61	175	0.50	13	59	0.32
Average	19.5	78.5	0.31	3.75	49.25	0.25
Week 9 - JOCWiki1	0	26	0.21	4	4	0.06
Week 9 - JOCWiki2	0	25	0.51	23	625	0.56
Average	0	25.5	0.365	13.5	314.5	0.31

Table 3: Proportion of triggered sentences in Wikipedia and JOCWiki.

the users in discussing the content to be added to Wikipedia articles and hence, serves as an apt triggering platform. Hence, we observe an increase in the addition of the sentences to a Wikipedia article after the inclusion of a talk page. Since a discussion environment triggers knowledge addition in the corresponding article, liberty of discussions would aid better knowledge building prospects in Wikipedia articles.

In JOCWiki, the coherent activity across the wiki article and QnA forum is observed for a period of 10 days before and after the inclusion of the QnA forum (as shown in Fig. 13, Fig. 14). We observe a rise in the number of sentences added to the wiki articles after the inclusion of QnA forums. This is because of the triggers generated, which lead to the addition of sentences in wiki articles and their respective QnA forums. From Fig. 15, it is important to note that the QnA forums in JOCWiki1 and JOCWiki2 are included soon after the first wiki article edit. Due to the late inclusion of the talk page on Wikipedia, editors miss out on many triggers that could have been pointed out otherwise. Also, due to the absence of the QnA forum, readers miss out on many important facts that could have been gathered from Wikipedia otherwise.
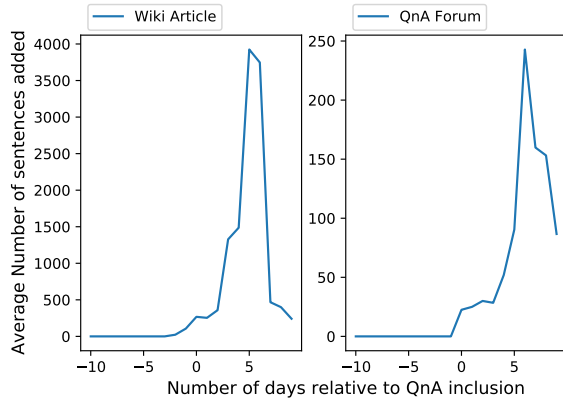

Figure 14: The above figure is plotted for JOCWiki2 data. The analysis is done for 10 days prior to QnA forum inclusion and 10 days post QnA forum inclusion. It is performed over all 12 articles of JOCWiki2 and the average number of bytes added on each day is plotted against the number of days relative to QnA forum inclusion.
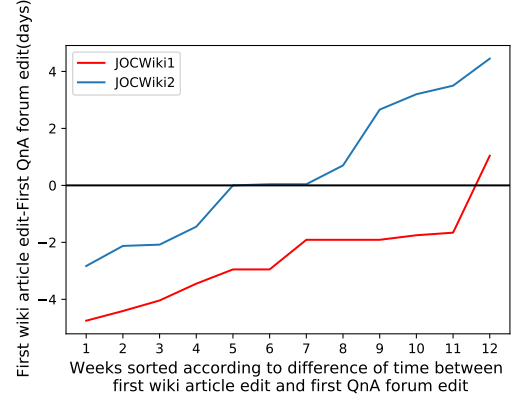

Figure 15: The above plot shows the early inclusion of the QnA forum to QWiki articles. The maximum time difference (First wiki article edit- First QnA forum edit) observed is 4 days for the week 1 JOCWiki1 article, and the time difference is positive for the majority of the weeks in JOCWiki2.

Therefore, the early inclusion of a QnA forum to a wiki portal aids the editors to develop the knowledge as well as the readers to acquire knowledge, thus portraying both the knowledge building as well as the knowledge acquisition perspective of QWiki.

6 LIMITATIONS AND FUTURE WORK

The QWiki dataset in consideration is from a portal named JOCWiki, specifically deployed for a MOOC. The crowd, responsible for the curation and maintenance of the JOCWiki portal comprises of course students. JOCWiki majorly targets the crowd interested to learn Computer Science topics whereas, the crowd in Wikipedia comprises of users from all over the world with varied interests. The participation in the JOCWiki portal is biased as it is driven by assignment evaluation and grades as a part of the course. However, participation in JOCWiki is not considered for final grading. But, students edited and referred to wiki articles/QnA forums to seek a better understanding of the concepts explained in the course. For Wikipedia, the participation of the users is unbiased as the users contribute to the articles with no selfish motives. In addition to this, our QWiki portal is for a limited duration of 12 weeks, unlike Wikipedia. It is out of scope to deploy the QWiki portal on a platform like Wikipedia. As a first step, we conduct a controlled

experiment by deploying it on MOOCs.

Since the QWiki portal in consideration (JOCWiki) is deployed on a MOOC course, there are concerns about the scalability of such a portal. Similar to JOCWiki portal where we have incentivization schemes like assignment scores, certification exams, it is a must to have incentivization schemes to attract crowd for the success of a portal. As a part of future work, we would like to implement a QWiki portal on a larger scale, like a QWiki portal for cities around the world. We would also like to explore the incentivization schemes on a QnA forum in the form of assigning badges to the users and voting the best answer among all the given answers. In addition to this, we would like to create a better interface for integrating the wiki portal and QnA forum. As of now, the QnA forum is a separate page integrated with each wiki article present on the QWiki portal. We would like to annotate each sentence of the wiki article with the option of asking questions so that it becomes easy for the users to find the answer to a question already asked by some other user.

7 RELATED WORK

7.1 Other wiki Integrated Projects

The literature pinpoints other portals which used a wiki-like interface and integrated other features according to their requirements. The Polymath project is a successful attempt to collaboratively solve some of the mathematical problems [10]. It is an amalgamation of wiki, blog, and various discussions in one place, which emphasizes the importance of discussions along with the wiki articles. In addition to this, there are many other wiki integration instances like qwikWeb, which is an integration of wiki and mailing list [13]. Twitter is also found to be a useful extension to a collaborative workspace like wiki [48]. Trac is a bug tracking system developed by Edgewall Software [36]. It is a combination of wiki and a bug tracking system. Users can create tickets for bug tracking, and they can organize information using the wiki system. Studies claim that direct work (editing in articles) is decreasing while indirect work such as questions, maintenance activity is increasing [21]. This demands the integration of a QnA forum with a wiki portal.

The collaboration of Wikipedia and Quora has also been proposed [8]. In a survey done by the authors, it is found that 83.6% of Wikipedia users are not satisfied with the content of Wikipedia articles. The authors claim that the users are not able to find answers to their queries in a Wikipedia article. Therefore, the authors propose to integrate a QnA facility like Quora which helps the users to find answers to their queries. We prove the above stated claim of effective knowledge acquisition through the data analysis of the JOCWiki portal. We also show the coexistence of the QnA forum and wiki portal aids the knowledge building process.

7.2 Triggering Theories

The addition of new knowledge to a knowledge building setup needs to be investigated to aid the acceleration of knowledge building in any knowledge building setup. There is extensive literature that supports triggering as one of the reasons behind the generation of new knowledge. The classical theories conform to the presence of triggering in various setups. Luhmann's theory of the Autopoietic system describes a social system such as collaborative knowledge building as an Autopoietic system, which essentially means that new ideas are generated by the existing ideas present in the system [24]. The same has been proved through Wikipedia data [9]. The authors tested it by measuring the Normalised Google Distance of the wiki links involved in a particular article. Wiki links represent essential parts of Wikipedia articles. Normalized Google Distance is used to measure the semantic similarity between wiki links. Lesser the value of NGD, the more similar the underlying wiki links are. Results show semantically related wiki links being added to the wiki articles, which prove the triggering phenomenon in play in a Wikipedia article.

Another theory about triggering phenomenon is Piaget's theory of equilibration. According to Piaget's theory of equilibration [30], if there is some piece of information incongruent to the existing knowledge of any individual, it gives rise to a cognitive conflict, which in turn encourages the individual to add more knowledge to the system. This process continues as long as equilibrium is reached between the system's knowledge and knowledge of the concerned individual.

The traces of triggering can be found in a discussion environment also. In Practical Inquiry Model [16], four phases are identified in a discussion forum. The first phase being triggering serves as instrumental in knowledge building. It is followed by the exploration phase. Even in computer-mediated discussions, the process of the addition of new knowledge starts from Triggering [17]. In discussions, triggering can be defined as posing the problem encountered in understanding the existing knowledge.

Triggering of new ideas is a corollary of discussions and brainstorming on the prevailing views. In the QWiki environment, the presence of triggers helped in the generation of knowledge, as explained in the previous sections. Hence, a setup like QWiki has the potential to make any crowdsourced article self-explanatory, self-sufficient, and thus complete.

8 CONCLUSION

QWiki is a novel technical design proposed to overcome the shortcomings of existing knowledge building practices i.e. collaborative authoring (exhibited by wiki portal) and collaborative query solving (exhibited by QnA forum). The QWiki portal alters the collaborative behavior of the crowd which results in better knowledge acquisition and accelerated knowledge building in the knowledge building portal. The knowledge gaps identified in the articles of JOCWiki facilitate better knowledge acquisition when addressed on the respective QnA forums. With the inclusion of an unrestricted QnA forum, a QWiki portal helps to capture various discussions among users. These discussions act as triggers that help in the generation of knowledge in wiki articles as well as the QnA forums. These triggers lead to knowledge building in the QWiki portal. As a result of triggering, we observe coherent activity across the wiki articles and their respective QnA forums. Hence, this research demonstrates the need for the coexistence of a wiki portal and a QnA forum.

9 ACKNOWLEDGEMENT

REFERENCES

[1] Alexander A Alemi and Paul Ginsparg. 2015. Text segmentation based on semantic word embeddings. *arXiv preprint arXiv:1503.05543* (2015).
[2] Rodrigo B Almeida, Barzan Mozafari, and Junghoo Cho. 2007. On the Evolution of Wikipedia.. In *ICWSM*.
[3] Samy A Azer. 2015. Is Wikipedia a reliable learning resource for medical students? Evaluating respiratory topics. *Advances in physiology education* 39, 1 (2015), 5–14.
[4] Siddhartha Banerjee and Prasenjit Mitra. 2015. Filling the gaps: Improving wikipedia stubs. In *Proceedings of the 2015 ACM Symposium on Document Engineering*. 117–120.
[5] Rudolf Beran et al. 1977. Minimum Hellinger distance estimates for parametric models. *The annals of Statistics* 5, 3 (1977), 445–463.
[6] David M Blei, Andrew Y Ng, and Michael I Jordan. 2003. Latent dirichlet allocation. *Journal of machine Learning research* 3, Jan (2003), 993–1022.
[7] Davide Buscaldi and Paolo Rosso. 2006. Mining knowledge from wikipedia for the question answering task. In *Proceedings of the International Conference on Language Resources and Evaluation*. 727–730.
[8] Anamika Chhabra and SRS Iyengar. 2016. Should wikipedia and quora collaborate?. In *Communication Systems and Networks (COMSNETS), 2016 8th International Conference on*. IEEE, 1–2.
[9] Anamika Chhabra and SR Sudarshan Iyengar. 2018. Characterizing the Triggering Phenomenon in Wikipedia. In *Proceedings of the 14th International Symposium on Open Collaboration*. ACM, 11.
[10] Justin Cranshaw and Aniket Kittur. 2011. The polymath project: lessons from a successful online collaboration in mathematics. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1865–1874.
[11] Ulrike Cress and Joachim Kimmerle. 2008. A systemic and cognitive view on collaborative knowledge building with wikis. *International Journal of Computer-Supported Collaborative Learning* 3, 2 (2008), 105.
[12] Pieter De Haan and Kees Van Esch. 2004. Towards an instrument for the assessment of the development of writing skills. In *Applied Corpus Linguistics*. Brill Rodopi, 267–279.
[13] Kouichirou Eto, Satoru Takabayashi, and Toshiyuki Masui. 2005. qwikWeb: integrating mailing list and WikiWikiWeb for group communication. In *Proceedings of the 2005 international symposium on Wikis*. ACM, 17–23.
[14] Andrew J Flanagin and Miriam J Metzger. 2011. From Encyclopaedia Britannica to Wikipedia: Generational differences in the perceived credibility of online encyclopedia information. *Information, Communication & Society* 14, 3 (2011), 355–374.
[15] Rudolph Flesch. 1948. A new readability yardstick. *Journal of applied psychology* 32, 3 (1948), 221.
[16] D Randy Garrison, Terry Anderson, and Walter Archer. 1999. Critical inquiry in a text-based environment: Computer conferencing in higher education. *The internet and higher education* 2, 2-3 (1999), 87–105.
[17] D Randy Garrison, Terry Anderson, and Walter Archer. 2001. Critical thinking, cognitive presence, and computer conferencing in distance education. *American Journal of distance education* 15, 1 (2001), 7–23.
[18] James W Guthrie. 2003. *Encyclopedia of education*. Macmillan ref. USA.
[19] Derek L Hansen, Mark S Ackerman, Paul J Resnick, and Sean Munson. 2007. Virtual community maintenance with a collaborative repository. *Proceedings of the American Society for Information Science and Technology* 44, 1 (2007), 1–20.
[20] Andreas Kaltenbrunner and David Laniado. 2012. There is no deadline: time evolution of Wikipedia discussions. In *Proceedings of the Eighth Annual International Symposium on Wikis and Open Collaboration*. ACM, 6.
[21] Aniket Kittur, Bongwon Suh, Bryan A Pendleton, and Ed H Chi. 2007. He says, she says: conflict and coordination in Wikipedia. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 453–462.
[22] Quoc Le and Tomas Mikolov. 2014. Distributed representations of sentences and documents. In *International conference on machine learning*. 1188–1196.
[23] Teun Lucassen, Roald Dijkstra, and Jan Maarten Schraagen. 2012. Readability of Wikipedia. *First Monday* (2012).
[24] Niklas Luhmann. 1995. Social Systems. Translated by John Bednarz, Jr., with Dirk Baecker.
[25] Lena Mamykina, Bella Manoim, Manas Mittal, George Hripcsak, and Björn Hartmann. 2011. Design lessons from the fastest q&a site in the west. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 2857–2866.
[26] Henry B Mann and Donald R Whitney. 1947. On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics* (1947), 50–60.
[27] MediaWiki. [n.d.]. MediaWiki. https://www.mediawiki.org/wiki/MediaWiki; accessed 23-March-2020.
[28] MediaWiki. [n.d.]. Structured Discussions. https://www.mediawiki.org/wiki/Structured_Discussions; accessed 23-March-2020.
[29] Kevin K Nam and Mark S Ackerman. 2007. Arkose: reusing informal information from online discussions. In *Proceedings of the 2007 international ACM conference on Supporting group work*. 137–146.
[30] Jean Piaget. 1977. *The development of thought: Equilibration of cognitive structures.(Trans A. Rosin)*. Viking.
[31] Catherine Porter, Jacques Brunschwig, Geoffrey Ernest Richard Lloyd, Geoffrey ER Lloyd, and Pierre Pellegrin. 2003. *A Guide to Greek Thought: Major Figures and Trends*. Harvard University Press.
[32] Quora. [n.d.]. Answer-Wikis. https://www.quora.com/What-is-an-Answer-Wiki-on-Quora-1; accessed 29-March-2019.
[33] Radim Rehurek. [n.d.]. Gensim. https://en.wikipedia.org/wiki/Wiki; accessed 24-March-2020.
[34] Amira Rezgui and Kevin Crowston. 2018. Stigmergic coordination in Wikipedia. In *Proceedings of the 14th International Symposium on Open Collaboration*. 1–12.
[35] Jodi Schneider, Alexandre Passant, and John G Breslin. 2011. Understanding and improving Wikipedia article discussion spaces. In *Proceedings of the 2011 ACM Symposium on Applied Computing*. 808–813.
[36] Edgewall Software. [n.d.]. Trac. https://trac.edgewall.org; accessed 20-March-2019.
[37] Parikshit Sondhi and ChengXiang Zhai. 2014. Mining semi-structured online knowledge bases to answer natural language questions on community QA websites. In *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management*. ACM, 341–350.
[38] Laure Soulier, Lynda Tamine, and Gia-Hung Nguyen. 2016. Answering twitter questions: a model for recommending answerers through social collaboration. In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*. ACM, 267–276.
[39] Besiki Stvilia, Michael B Twidale, Les Gasser, and Linda C Smith. 2005. Information quality discussions in Wikipedia. In *Proceedings of the 2005 international conference on knowledge management*. Citeseer, 101–113.
[40] Aaron Swartz. 2006. Who writes wikipedia. *Raw thought* 4 (2006).
[41] Fernanda B Viegas, Martin Wattenberg, Jesse Kriss, and Frank Van Ham. 2007. Talk before you type: Coordination in Wikipedia. In *2007 40th Annual Hawaii International Conference on System Sciences (HICSS'07)*. IEEE, 78–78.
[42] Nicholas Vincent, Isaac Johnson, and Brent Hecht. 2018. Examining Wikipedia With a Broader Lens: Quantifying the Value of Wikipedia's Relationships with Other Large-Scale Online Communities. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 566.
[43] Yi-Chia Wang, Mahesh Joshi, and Carolyn Penstein Rosé. 2008. Investigating the effect of discussion forum interface affordances on patterns of conversational interactions. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work*. ACM, 555–558.
[44] Wikipedia. [n.d.]. Wiki Talk Pages. https://en.wikipedia.org/wiki/wikipedia:Talk_page_guidelines; accessed 12-December-2019.
[45] Wikipedia. [n.d.]. Wikipedia Statistics. https://en.wikipedia.org/wiki/Wikipedia:Statistics; accessed 4-October-2019.
[46] Guangxu Xun, Vishrawas Gopalakrishnan, Fenglong Ma, Yaliang Li, Jing Gao, and Aidong Zhang. 2016. Topic discovery for short texts using word embeddings. In *2016 IEEE 16th international conference on data mining (ICDM)*. IEEE, 1299–1304.
[47] Amy X Zhang, Lea Verou, and David Karger. 2017. Wikum: Bridging discussion forums and wikis using recursive summarization. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. ACM, 2082–2096.
[48] Xuan Zhao, Wenpeng Xiao, Changyan Chi, and Min Yang. 2011. Integrating Twitter into Wiki to support informal awareness. In *Proceedings of the ACM 2011 conference on Computer supported cooperative work*. ACM, 733–736.