

Context in Social Media Question Asking: A Practice Approach

Hasan Shahid Ferdous

School of IT, Deakin University, Australia

Email: hasan.ferdous@deakin.edu.au

Kazi Abir Adnan

School of CIS, The University of Melbourne, Australia

Email: kazi.adnan@unimelb.edu.au

Abstract—Recent works have highlighted the complementary advantages and disadvantages of web search and social media question asking (SMQA), and there has been a growing interest in merging these two. To make this idea a success, understanding the context of questions in social media is a prime concern. Yet existing works give us a hint that the positivist approach towards the context of questions is not sufficient in achieving this, and we need to revisit the concept of context regarding SMQA. In this paper, we took a *practice* based lens [1] to examine existing works along with real-life data of queries in a popular social network to see if it explains the broader contextual factors and associated values. Our data highlighted the importance of understanding the complex social relationship that people make within the structure of social networks, and suggest design strategies to support this through iteratively progressing from phenomenological to representational view of *practice*.

Keywords—Social Media, Search Engine, Social Query.

I. INTRODUCTION

Social media question asking (SMQA) has emerged as an area of research that deals with users' information-seeking behaviour through popular social networking sites (e.g., Facebook, Twitter, etc.) instead of traditional search engines. With the emergence of social networking sites (SNS) over the past two decades, it took little or no time for both researchers and users to understand its potential in information retrieval. Many issues have intrigued the researchers over this period, for example, what types of questions swamp the ever-increasing arena of social networks [2], why do we use it for question-asking purpose [2] (and why we do not [3]), who answers these queries [4] and why [5], social-bonding [6] and social costs [7] associated with SMQA, how to take advantage of search engines (SE) with this social search ([8], [9]), etc. In this work, we aim to take a look back to understand the context of the questions in social networks.

Understanding the context of a question in SMQA can provide us with a multitude of advantages. SMQA and search engines have complementary benefits ([10], [9]) and major SEs are exploring this field to incorporate information from a social network into their search results [11]. Doing this in the opposite way, i.e., bringing web search results to SMQA has substantial advantages too, supporting our natural information-seeking behavior [12]. This will harness the benefits of SMQA (e.g., trustworthiness [2]) while removing its drawbacks (e.g., delay [13], lack of response [14]). Same as traditional SEs,

understanding the context of the query is of utmost importance here too.

Some of the issues could be identified on the very outset, such as hidden context, irrelevant keywords, etc. Yet the lack of ability to address this issue in some related projects ([8], [15]) brings the inherent complexity of the problem to attention. It would appear from the data gathered and our discussion later that this well-known problem in natural language processing and information searching has additional facets [16] when it comes to SMQA, and careful consideration of the problem scope is required. In this work, we draw on some recent trends in Human Computer Interaction (HCI) to provide a pathway in addressing this issue. We want to probe on why the existing approaches do not work well and discuss a different perspective providing more in-depth perception. In understanding the context of questions, we focused on two aspects of SMQA. First, the *social* and *conversational* nature of SMQA that separates it from the traditional web search. Through examples, we will discuss how the meaning of a query in SMQA is established in real-time. And second, the explicit and implicit assumptions that users make based on their social relations and experience and thus *hide* or *embed* the query in the context of it.

Our contribution in this paper is to address the gap that the representational stance of context in association with social media question asking has created. We took a practice lens to understand the mundane details of the dynamic social experience and relations in our social networks that enable us to comprehend the background assumptions in SMQA. The practice model of context concerning SMQA enables us to see the continual negotiation and progressive nature of our social queries. We provide a pathway for understanding the scope of the problem associated and some directions for designers in this space.

To emphasize these aspects, we will discuss *anthropomorphism* and *technomorphism* as we experience through the use of technology and relate these with social media question asking. Then we discuss some contemporary works in the field of SMQA before we delve further to explore the existing works that tried to understand and utilize the *context* of SMQA in delivering assistance to the users. Then we visit the alternate model of context based on *practice* as discussed by Dourish [1] and other researchers. Real-life examples drawn from a popular social networking site (Facebook) are presented next,

relating the discussion with previous works and how the practice model fits to address the issues.

II. SMQA: ANTHROPOMORPHISM VS. TECHNOMORPHISM

Anthropomorphism is attributing human form or other characteristics to anything else. For example, when someone says “Mr Obama’s statue of wax at Madame Tussauds looks very lively”, s/he is ascribing some *human* aspects of Mr Obama to that statue. The question that relates this concept with SMQA is: when users search for answers through search engines or social networks, does anthropomorphism play any role?

While using search engines, we often keep in mind that we are *asking a machine* which will find keywords in our query and best match with its database using advanced algorithms. Users *know* that they are communicating with a machine, which has reflexive consequences on the users’ behavior - they act as the machine expects them to act. Explained by Vertesi in her article *Seeing like a rover* [17], this is opposite to anthropomorphism. She termed it to be *technomorphism*, that is when people try to think or act like a machine. They try to consider themselves in the position of those machines. It is only in this way that they can accurately understand what this machine can (or can’t) do for them. It may happen consciously or subconsciously every time users do a query in search engines as they inevitably make the query keyword rich. One evidence to this claim is the average query length - 2.4 in 2011 [18], or 3.08 in 2012 [19], thus indicating that the queries in search engines are keyword rich.

We argue that the opposite thing happens when we ask questions in our social networks. Users remain aware that their questions will be seen and answered by other persons in that network, maybe by their friends. They remain aware of human intervention and frame the question accordingly. In prototypical cases, when someone gets an autonomous reply (for example in case studies [8]), that user considers it as spam.

Now, that users expect their query to reach a human being has a significant impact on how they pose that query. It is evident, for example, through the measure of average query length of 21.45 words (standard deviation: 11.39) in our data. Evidence of unnecessary descriptions, implied information, humor, etc. is widespread in SMQA. After all, social media is only a medium to spread the query to other persons. This communication is computer mediated human-to-human; different from the human-to-computer interaction with search engines. With SMQA, generally, there is no question of either being anthropomorphic or technomorphic - as users neither pose their query to a machine nor expect a machine to answer.

As we consider understanding the context of questions, especially for the integration of search engines with social media question asking, we need to consider this. The query in SMQA is made with the human observer in mind, and users are *not* being technomorphic in posing the query. But a SE intervention requires to understand and answer that query. When it does (as in [8]), users do no longer remember the non-technomorphism in their query and judge the automated reply from anthropomorphic mindset (as also noted by [8]).

Two notable reactions from the users of SearchBuddies [8] were: “*I might need to do something about [SearchBuddies], it’s acting like the annoying friend you silently delete*” and “*Oh hell. Die...bot*”. We can see the notion of anthropomorphism in these two feedbacks. Algorithmic search engines often fail too, but rarely are they publicly told to “*die*”.

The language in which users pose their query to other humans can be much harder for a machine to interpret using the keywords it is ready to process. And as we have learned from previous failures ([8], [15]), a successful understanding of the context would require not only understanding the query itself, but also the social connections to which it must conform.

III. RELATED WORKS IN SOCIAL SEARCH

The history of research on social search precedes the age of social networks ([20], [21]). The use of user-forums is almost as old as the Internet itself, being one of the first uses of this wide-area connectivity. As the Internet started to connect the vast population around the world, user groups and forums became ubiquitous too. People can post their queries and get answers from others using these crowd-sourced forums. They still do exist and is very useful, but as the Internet paved the way for social networking sites, a new era of social search appeared before us. Question asking through these social networks are often termed as *friend-sourced* compared to the existing *crowd-sourced* social search methods [2].

One of the first studies about SNS based information search is done by Efron et al. [22], who identified that micro-blogging services like *www.twitter.com* are gradually becoming a popular venue for informal information exchange. They showed that question-asking in micro-blogs is strongly tied to peoples’ naturalistic interactions, which helped them to offer a taxonomy of questions in micro-blogs. They also showed that the act of asking questions in Twitter is not analogous to information seeking in more traditional information retrieval environments, and contextualized these articulations through analysis of a large body of tweets. Teevan et al. [23] presented the systematic overview of search behavior on Twitter and its differences with the web search using questionnaire data and analysis of query logs. They found that Twitter results included more social content and events, while web results contained more facts and navigation. Their study recommended that search engines could use trending Twitter queries to discover additional queries that have strong temporal components.

Lampe et al. [24] investigated the Facebook users’ characteristics based on a survey of 614 people who used it to ask something. They identified that the perception of the relationships within members acts as significant predictors of the information-seeking approach. They did not show any comparison between SNS and SE regarding obtaining any particular type of information. This question is addressed by Morris et al. [25], where they explored the pros and cons of using SNS as an information source and compared user interaction when they search anything either on SNS or SE. There are many motivations for asking questions in SNS. Among those the most important reason was the belief

that people in our social network knows our *context* better, therefore may provide more relevant answers [26].

All these studies establish that social search has essential differences and complementary benefits (and drawbacks) with search engines. While users often believe that SMQA can provide more relevant and trustworthy answers [2], even when the information is not available through search engines [13], many of the queries remain unanswered [14]. On the other hand, SEs can provide us with many search results almost instantaneously while one might require to wait hours to get an answer from social networks [13].

IV. THE CHALLENGE OF PROVIDING ASSISTANCE IN SOCIAL MEDIA QUESTION ASKING

Hecht et al. [8] took an initiative of integrating traditional search engines with social media to provide an algorithmically generated reply to users' queries on Facebook. Their project *SearchBuddies* had two components - *Investigator* and *Social Butterfly*. Investigator used a white-list of 31 web domains, developed using the data set of status message questions from Morris et al. [2] and determined by an experienced user. In their 67 day deployment with 122 people, SearchBuddies identified 262 questions, based on the presence of "?" symbol in the status message, of which 72 were later determined to be false positives (most of these were rhetorical questions). Investigator sent all these queries to a traditional search engine API, and if any of the top three results came from the white-listed domains, it posted a short link as a comment to the original query. They tried to minimize unreliable and irrelevant posts forwarded by the Investigator through human intervention in white-listing the web domains, making only 58 replies (22%) to those 262 queries. Still, many of them could not address the actual question, while some were providing irrelevant answers to provoke humor among the audience and anger in some others (e.g., "*oh hell, die ... bo!*").

Social Butterfly part of SearchBuddies tried to identify other persons from the asker's social network who might have some knowledge about the query. They used people's interest and places from their Facebook profile to filter which of the asker's friends may provide a useful pointer to the query and tagged them in that question. Feedback from the users provided some insights about the lack of social ties in considering the list, failure to understand the context of the question, and of course, some success of the initiative. This study provided an important direction for SE integration with SMQA - we need a high relevance threshold to provide an automated reply and avoid false positives. With search engines, returning no answer is annoying while we are willing to forgive the *dumb algorithms* for not understanding our query for providing irrelevant answers. But as we have learned from this study, with SMQA, providing irrelevant answer could outrage people to the point of blocking the system from further intervention.

Now if Twitter was a too generic platform, let's focus our attention on a restricted case - the formal environment in the intranet of an IT company - Microsoft Corporation. White et al. [15] tried to keep a balance between time latency

and interruption costs when they developed their synchronous social Q&A system *IM-an-Expert* to seek professional assistance from within the community. Their project had two parts - recognizing the expertise and an Instant Messenger (IM) interface to communicate with them. The process involved creating an explicit self-reported knowledge-oriented profile, where users provided keywords and URLs to personal websites describing their expertise. They also analyzed the mailing lists of their 30000 employees, accruing over 0.3 million emails to crawl and index. They invited random people from one of the campuses in the organization, recruiting 260 participants.

They used the TF-IDF [27], an established ranking function used extensively in information retrieval (IR), normalized based on profile length. When a user posted a query, it was analyzed based on keywords using TF-IDF to identify top 5 *experts* based on their existing profiles. The query was forwarded to either top two or all five of them (2 control groups to compare), and when one receiver agrees to answer the query, the negotiation process ends. Otherwise, the system would keep forwarding the query to another 2 or 5 people from its expert list until someone agrees to answer the query. This is somewhat a restricted environment - we are considering the employee network of a corporate office, thus can reasonably assume the absence of ambiguous and rhetorical questions. Yet, their findings showed us that many, if not most, of their users (45% - 55%) provided feedback that more than 90% of the question asked to them were not relevant to their expertise.

We do not criticize the work or approach taken by White et al. [15] or Hecht et al. [8]. Their systems considered many design aspects that are required to be present and provide essential insights into the problem space. What all these works inherently lack is understanding the complexity of the communication when humans communicate with other human beings, as we shall show with some examples later.

Now, as we have seen some apparent failures of machines to understand human-to-human queries, we do expect that human beings might have done better. So we now focus into another study, which is a *crowd-sourced* approach to the *friend-sourced* world of social media question asking.

Jeong et al. [28] experimented with a paid crowd-sourced system to provide answers in reply to queries made on Twitter. They tried to overcome the criticisms faced by SearchBuddies [8] by incorporating humans in the loop, as crowd-workers can identify questions that may be inappropriate to answer (such as humorous or rhetorical) or require special consideration (e.g., questions about sensitive topics). They identified questions that require additional *context* (22% of all non-rhetorical queries) that could not be extracted from the asker's Twitter profile. Their view on context is, what we can call the representational stance. After carefully avoiding all these challenges, they found that the crowd-sourced answers were similar in nature and quality to friend-sourced answers, also verified by the unsolicited positive feedback by the users.

So, involving humans in the task could potentially solve the dilemma. But it leaves us with the sheer amount of task required by humans to answer all those 85,739 queries Jeong

et al. [28] identified in a month, that we can assume will be increasing over time. Also, we need to consider those 18.84% of all the non-rhetorical questions that the crowdworkers considered as lacking context and did not reply but got some response from asker's followers on Twitter. Do our friends know our context better than some random crowdworker? Probably does, as also reflected by a few cases where their system did not work fine. We are not criticizing their approach, but want to broach the fact that understanding the context of a human query to another human is non-trivial.

V. CONTEXT AND PRACTICE

In his work [1], Paul Dourish explored the concept of context and suggested that representational or positivist stance for encoding context found in many of the HCI works misinterpret the role of context in everyday human activity. Dourish [1] proposed an alternate model for it based on phenomenological reasoning. In identifying the role of context in understanding human-to-human queries, we support his view. We want to explore how this model fits into our problem of understanding context in SMQA. Many works in HCI has considered context as a representational affair to make the computation sensitive and responsive to the physical or social settings it is used. In short, these approaches can be called positivist reasoning that seeks objective, independent descriptions of social phenomena, often in favor of broad statistical trends and idealized models, and are often quantitative in nature. Dourish [1] summarized these works to have four assumptions.

First, *context is a form of information that can be encoded and represented as any other information*. For example, Jeong et al. [28] took the context of the query as solid and distinguished pieces of information that we can extract from the query itself or from the asker's profile in the social network, namely username, short bio, location, photo, and recent tweets.

This approach, while practical and feasible, skipped those 18.84% queries that the asker's acquaintance could provide an answer with, but their crowd-workers could not; because of the lack of "contextual information". To successfully integrate SE with SMQA, we need to be prepared to address this discrepancy. The other two works we discussed in details also followed the same pathway. SearchBuddies [8] used places and interests mentioned in the user's profile to identify relevant persons in that asker's social network. White et al. [15] also used keywords mentioned by the users to select *expert*.

Second, according to the positivist approach, *we can define in advance which information counts as a context*. Clearly, all the three aforementioned works did so. We will see that in the open and diverse arena of social networks, where avid users share their real-life emotional or professional issues with people ranging from their family members to strangers, this assumption may not always hold.

Third, all three works discussed in the last section assumed that *through the life-cycle of an application, the pieces of information that we consider as context are stable*. They pre-determined the relevance of any potential contextual element which remained static throughout the study. SearchBuddies

explicitly acknowledged the pitfalls of this approach and called for a more dynamic scheme in their discussions.

Finally, *context and activity are considered separable*. This is an assumption of positivist reasoning followed by all these three works. For example, when a user looks for a suitable restaurant in Paris (only mentioned as an example from [28]) for a dinner date, we can consider the location (Paris) and time (dinner time) as the associated context (in positivist approach) and dating as the associated activity. However, if we consider activity and context as separate (ignoring the bidirectional relationship among them), we might end up like the crowdworkers in [28], suggesting a restaurant more suitable for teenagers than adults.

If we consider the phenomenological reasoning of context, we can come with a different stance for each of the four assumptions above. First, we will require to consider context as not the information, but the *relation* between objects or activities. Everything can be part of the context, but then the question is whether it is contextually relevant to the activity/object under consideration or not. And this property of contextual relevancy needs to be resolved dynamically. Because in phenomenological perspective, context is not a stable property but evolves dynamically as activity appears and progresses, in relation with particular settings, actions, and parties involved with that action. Context and activity are inseparable. As Dourish has put, "*Context isn't just 'there', but actively produced, maintained and enacted in the course of the activity at hand*" ([1], p. 5).

So if we follow the phenomenological view of the context, where does it lead us to *solve* the aforementioned problem of algorithmically understanding human-to-human query by computers? Dourish [1] took the help of a larger frame called *practice* to investigate how this alternative model of context fit into the realm of interactive system design, and so do we. Before going into real-world examples of how the concept of practice conforms to our current problem with SMQA in the next section, we elaborate into the concept of practice and its applications in HCI research first.

The informal understanding of practice is the detail of what people do, in contrast to what rule-books suggest what people ought to do. Wenger in his book *Communities of Practice* ([29], p. 51) explained, "*Practice is first and foremost a process by which we can experience the world and our engagement with it as meaningful*". Thus the concept of practice is the one that unites actions with meaning. It explains how the meaning of particular actions dynamically evolves and adapts with our participation in the communities.

Pierce et al. [30] drew on fields such as anthropology, sociology, cultural studies, philosophy, and geography to locate practice within the routine and seemingly mundane activities of everyday life. They incorporated consideration of the social, cultural, and material contexts in which those activities are situated. A key reference point is Reckwitz's framing of practices, consisting of relationships between various types of elements including "*forms of bodily activities, forms of mental activities, things and their use, a background knowledge in*

the form of understanding, know-how, states of emotions and motivational knowledge" ([31], p. 249). Shove et al. proposed a simplified three-part framework for conceptualizing practice - i) entities as consisting of "*things, technologies, tangible physical entities, and the stuff of which objects are made*", ii) competences, which "*encompass skill[s], know-how and technique[s]*", and iii) meanings, which include "*symbolic meanings, ideas and aspirations*" ([32], p. 14). Other sociologists, for example, Schmidt [33], argued that practice is not mere *doing*, or simple execution of pre-defined rules, but crucially involves reasoning in the *doing*. The work that people perform is not simply the following of preordained rules, but necessarily involves the local interpretation of these rules in the light of the evolving situation.

User studies in HCI often focus on studying current practices about what people currently do, want, or wish for. But, phenomenologist put the concept of practice with a much larger view. We cannot ignore the dynamics of practice, that is how current practice conforms into future ones and keep changing substantially over time. The Practice paradigm, as explained by Kuutti et al. [34], takes a longitudinal view to examine the process and performance of users longer-term actions that persist over time. It situates users' activities not only in time and space but also considers the surrounding physical and cultural environment interwoven within the practice. From a practice perspective, our words and activities can be seamlessly integrated and understood by others because they are not discrete *keywords* but has an inherent meaning that evolves through our minds, physique, artifacts, etc. These are connected so that the result of performing one activity serves as a resource for another. So it is unsurprising then that the search engines looking for *keywords* in SMQA failed to address the *practice* of it and hence the context of the query was not often appropriately recognised.

The *turn to practice* paradigm ([35], [36]) in social sciences has been slowly gaining its position in HCI, though the formal agenda for *computer supported practice* has not been established yet [34]. In the existing works that dealt with the *context* of interaction in SMQA, the input query in the social network of the user is considered at the *center*, and everything else is considered as a *context* to this primary interaction and is treated independently. Alternatively, the practice paradigm requires us to *de-center* this privileged position of interaction, as noted in [30]. It requires us to consider the query text as only one factor among several that are relevant and important. The practice paradigm views the current query as the momentary result of evolution, which is constantly changing, rather than being static. Practice can be interpreted as the ultimate context. Then we can no longer consider the query as the foreground and context as the background. It can be studied and understood only through the whole performance, not separately.

VI. DATA COLLECTION AND DISCUSSION

The data presented here are collected from real Facebook posts of people. As argued by Kuutti et al. [34], practices

are contingent, mediated, and cannot be understood without reference to the particular place, time, and concrete historical context where they occur, they can only be studied close-up. Also, as we are interested in real-life practices, they must be studied where they occur, in their natural setting [35].

These data are part of another related project by the authors [37]. We posted a request for response through our informal university mailing list and Facebook group, through which we could reach about 20000 alumni and current students. We requested a sample questions they posted on Facebook over the past one-month period along with the responses received. We posted this request once for four successive weeks. In total, we could obtain 991 samples from that many unique users. Our data collection method is similar to existing research in this domain (e.g., [2], [15]). The ethical concerns related to collecting data from Facebook was carefully scrutinized.

This paper does not present a quantitative view of all these queries, nor discusses their qualitative aspects. Here we would like to present some of the queries that highlight the drawbacks of the representational view of context regarding SMQA and advocate in favor of the phenomenological reasoning of practice in understanding those queries. Some of the questions were in the native language (Bengali) and were translated into English. These were classified into types and topics as discussed in [2] and had been extensively analyzed in [37]. Many of the queries we collected can be answered quickly by search engines; also, many of those go beyond the capacity of search engines due to the lack of sufficient data on the web. Many of the questions could be identified as opinion, rhetorical, invitation, favor, or social connection, and either required human response only or has no definite answer. In this work, we exclude those and solely focus on queries that have some chance of being correctly responded by an algorithmic search engine.

A. Social Nature of SMQA

SMQA is not about information only; it has the term *social* in its very name and nature. Users know that they are asking the question to another fellow human user, so being conversational and courteous is a norm. The queries commonly found in SNS are not keyword-rich, instead contain a lot of explanation about why they need that information, and often gratitude expressed in advance. Let us consider the following query posted by one participant:

"Quebe [an ISP] has blocked YouTube again, I need to see some YouTube video for urgent purpose, please let me know how to do it. I will remove this status after learning the trick."

One might wonder how SEs will respond to this query. If we enter this quote as a search text in the web interface of a SE, we may see some erratic answers. We can instantly understand what the asker meant in this 32 word long query. Using a search engine with two keywords here "blocked youtube" can provide the asker with the answer s/he was looking for. But as we expect, traditional SEs will try to match as much text as it could with different sources found on the web. In general, we



Fig. 1: Implicit information in the context of question.

find that query texts in SMQA contain a lot of wordings that entirely makes sense to other human beings while making the task of the search engines extremely difficult.

But this is the very nature of social media question asking. Spink et al. [18] determined that average web search queries are about 2.4 words long. After a decade of their study, the average length of queries has grown steadily over time (3.08 words as of 2012) and the average length of non-English languages queries had increased more than English queries [19]. Google implemented the hummingbird update in August 2014 to handle longer search queries as conversational queries are growing fast. Average query length in our obtained data is 21.45 words (standard deviation: 11.39). We do not consider our data truly reflecting every possible case, but we expect this trend to continue [12], and if that happens, extracting keywords from the query may not be a good option any more.

Let us consider another example: “Does anyone know the procedures of washing blankets? Quickly inform me if you know it... Need to do this task within weekend.” Does the phrase “quickly inform” make any sense to a search engine, as it always tries its best to answer any query quickly? Does a search engine require to know that someone wants the information regarding washing his/her blanket over the looming weekend? Is “weekend” a part of the context, as it is viewed by the representationalist norm? Or, the time frame is a part of the ordinariness, as Paul Dourish has associated context with ordinariness (for example, in our conversation)? He noted that “Context is an occasioned property of action in just the same way as ordinariness” ([1], p. 9). If we place context in SMQA within the frame of practice, we need to support this conversational nature of queries out of which our actions and meaning do evolve.

B. Implicit Information within Queries

Human conversations are not only words but also an uncountable number of assumptions and background-understandings between the speaker and the listener play crucial roles in conveying the message. It is unsurprising then



Fig. 2: The question translated from Bangla reads: “Live streaming link of the game please...”

that so many of the queries in the social network have implicit information that the query will render meaningless (or worse, misleading) if considered in an isolated way. For example, we can consider this subtle example query from a participant: “Is there any place at all in Dhaka that sells Starbucks coffee?” (Figure 1).

The query was about finding some Starbucks coffee in Dhaka, the capital of Bangladesh. The international coffee-house chain Starbucks has outlets in 79 countries (as of May 2020), but Bangladesh is not one of them. The first response to that query provided an algorithmically correct answer. But as it appeared from the subsequent conversation, the asker already knew that there is no Starbucks coffeehouse in Bangladesh, and what he was looking for was if he could buy some coffee beans of the brand, which though not typical, was available in Dhaka. So, how does the second person could distinguish the real query out of the ordinary? Also, do we expect our search engines to discern between the explicit and the real but implicit query? It is, we argue, possible under the practice paradigm that requires us to go beyond the action and crucially involves comprehending the reasoning of that action.

C. Context of Queries in Community of Practice

To the best of our knowledge, none of the works that dealt with the context in social queries has focused on the community of the associated person, although previous works by Lampe et al. [24] identified that the perceptions of relationships within network members are “significant predictors” of the information-seeking approach. Indeed, they did not take a practice lens to examine the affair, but when we focus our attention on the following two sets of queries, we can try to understand why the community of practice [29] should be considered. Set one involved the following five queries made through posts in Facebook in close proximity of time (in about three hours interval) (Figure 2, Figure 3): i) “Live streaming link of the game please...”, ii) “Can anyone please give me the live streaming link of Ban vs SL?”, iii) “Do you have online streaming link for Bangladesh V Sri Lanka Test?”, iv) “YouTube link of BD match?”, v) “What happened to us?? Why we threw it away?”



Fig. 3: Asking for streaming link of Australian Open match

The last one is probably rhetorical; it is included as it is related to the same event - a cricket match between Bangladesh and Sri Lanka. Two of these queries do not mention the name of the teams or the game, but we can see from our data that people did not feel any ambiguity of the queries. How could people understand the context of those queries? Then answer, we argue lies in the concept of *community of practice*.

The community of practice, according to Wenger [29], is a group of people who share a common interest in a particular domain or area. It is through the process of sharing information and experiences with the group that the members learn from each other, and have an opportunity to develop themselves personally and professionally. In the realm of social media question asking, explicitly or implicitly, we are channelling our query to the members of our community of practice (and we need to remember that generally, we are members of multiple and overlapping communities of practice). So it will be unsurprising to have machines, that are not considering this community as a factor in its computation, fail miserably in providing the service that its members can obtain from one another without any setback.

Human-to-human queries can have further complexities, as illustrated by our next example. Consider the following query made about the same time as the five queries mentioned above: "Anybody has the live streaming link? Can't find it". This query is separate than the previous ones. Though it has the very similar wordings and made by a person who lived in the same community (maybe not in the same community of practice though) as the previous five askers, he did not refer to the same game. He made this query in the context of the Australian Open. We could see that human responders understood that quickly. How can a search engine understand this difference? Distinguishing which community of practice this person belongs to (and s/he can belong to multiple), and using it to understand the context of the query is by no means a simple task. We can see that his fellow social network members could answer his query without ambiguity. It is also this community of practice that can shed light on the following three *ordinary* queries: i) "What is a chamber judge? Anyone explain please?", ii) "How many time I will see the gaps in

laws?", iii) "Any lawyer friend here? Please explain the issue of Chamber Judge?"

If we ignore the sociality and the community of practice with regard to SMQA, we might end up answering about the definition of what a chamber judge is. However, that was not the intention of those queries. It referred to a judicial situation in Bangladesh at the time of our data collection regarding the trial of a war criminal there - when a chamber judge temporarily stopped that trial for a judicial cause. Considering keywords could ignore the inherent meaning of the temporal nature of the query that humans are so adept at, as shown by the answers in those threads.

We do not grow up to be an expert member of the community in one instance. Learning, as Wenger [29] has put, is a social practice, and in learning a new form of practice, we learn a new set of ways to comprehend the world. In general, practice is about finding the world meaningful in terms of the actions that it affords. As one acquires these skills, new aspects of the environment become relevant for the activities that one performs, and the scope of potential context broadens.

VII. PRACTICE BASED DESIGN AND CONCLUDING REMARKS

If we adopt the *practice* lens to understand the context of queries in social media question asking, where does it lead us towards algorithmically providing answers for SMQA? This draws our attention from the theorist viewpoint of practice toward designers' viewpoint of it. We find inspiration and guideline from Bauer et al. [38] in this regard. Thinking from the designers' perspective, Bauer et al. did not find it feasible to discard positivist interpretation entirely. They suggested a gradual change from phenomenological perspective towards a positivist interpretation, as the software design iterates. While some might consider this approach as *impure*, their suggestions are driven by necessity and practical constraints that system designers face in real-life projects.

The phenomenological viewpoint of context suggests open-ended and adaptive design solutions; while the positivist way that designers often encoded context imposed constraints on the system. Bauer et al. discussed how the difficulty inherent in determining and responding to context forced the designers to either simplify the design space to incorporate only the most impactful contextual elements ignoring the rest or to explore previously overlooked areas of the design space. One empirical guideline they provided is to move from phenomenological understanding of context to more positivist reasoning as the design goes through different iterations of framing, encoding, unifying, and evaluating phases of the software project.

Algorithmic understanding of context and perspective is undoubtedly a challenge. Related literature in text mining has considered five approaches - sentiment analysis, emotion analysis, belief mining, intent mining, and finally context learning, thus acknowledging context learning as the most challenging part [39]. In dealing with this, a recent work by Bamman et al. [40] takes in account the immediate communicative environment between author and audience to detect

sarcasm as a contextual phenomenon on Twitter. They used features from four sources - the relevant tweet (scoped only over recent tweets), author (profile and historical data), the audience (historical interaction between author and audience), and response (interaction between immediate and predicted tweet). Their approach claimed better accuracy compared to purely linguistic ones. In another study, Tausczik et al. [41] studied the collaborative approach to solving a problem in a question answering community (MathOverflow Q&A). They identified the collaborative acts (providing information, clarifying question, critiquing/revising/extending an answer) that happen between contributors and asker in that problem-solving forum. Though collaboration in social media's people is quite different from a Q&A community forum, identifying the collaborative act while answering a query could prove helpful in learning the context. As shown earlier, both asker's and contributors metadata and historical data along with the classification of collaborative acts at that recent point of time could assist in determining the relevant community of practice and hence help us in identifying the context.

We understand that from the computational perspective, it might not be possible to become *omnipotent* in capturing all the contextual information; especially if we consider that some of these information are not there in the digital world and only evolved through physical-world interactions among the parties involved. However, taking the *practice* approach can enable focusing more on the ordinariness, past history, and communal organization behind the query. And through gradually refining the context (even through inquiring with the askers or their friends, as suggested by Hecht et al. [8]), such systems can improve the experience of the users. Algorithmic assistance with SMQA has a long way to go, but it is high time to distinguish between the obvious, easier, somewhat-workable solution and the difficult but promising approach to understand the complex picture of social media question asking.

REFERENCES

- [1] P. Dourish, "What we talk about when we talk about context," *Personal and Ubiquitous Computing*, vol. 8, no. 1, 2004.
- [2] M. R. Morris, J. Teevan, and K. Panovich, "What do people ask their social networks, and why?: A survey study of status message Q&A behavior," in *Proc. CHI*, 2010.
- [3] M. W. Newman, D. Lauterbach, S. A. Munson, P. Resnick, and M. E. Morris, "It's not that I don't have problems, I'm just not putting them on Facebook: Challenges and opportunities in using online social networks for health," in *Proc. CSCW*, 2011.
- [4] A. Oeldorf-Hirsch and D. Gergle, "'who knows what' audience targeting for question asking on facebook," *Proceedings of the ACM on Human-Computer Interaction*, vol. 4, no. GROUP, pp. 1–20, 2020.
- [5] N. B. Ellison, R. Gray, J. Vitak, C. Lampe, and A. T. Fiore, "Calling all facebook friends: Exploring requests for help on facebook," in *Proc. ICWSM*, 2013.
- [6] J. M. Rzeszotarski and M. R. Morris, "Estimating the social costs of friendsourcing," in *Proc. CHI*, 2014.
- [7] H. Ma, B. Yu, H. F. Cheng, and H. Zhu, "Understanding social costs in online question asking," in *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, pp. 1–6.
- [8] B. Hecht, J. Teevan, M. R. Morris, and D. Liebling, "Searchbuddies: Bringing search engines into the conversation," in *Proc. ICWSM*, 2012.
- [9] A. Oeldorf-Hirsch, B. Hecht, M. R. Morris, J. Teevan, and D. Gergle, "To search or to ask: the routing of information needs between traditional search engines and social networks," in *Proc. CSCW*, 2014.
- [10] J.-W. Jeong, J.-U. Heu, and D.-H. Lee, "Understanding the potential of social questions in the web search," in *Proc. CSCW*, 2013.
- [11] A. Singhal, "Search, plus your world, <http://googleblog.blogspot.com.au/2012/01/search-plus-your-world.html>, as of 31/08/2020."
- [12] M. A. Hearst, "'Natural' search user interfaces," *Commun. ACM*, vol. 54, no. 11, 2011.
- [13] S. Ahmed, M. T. A. Anik, M. Tasnim, and H. S. Ferdous, "Statistical analysis and implications of sns search in under-developed countries," in *Proc. OzCHI*, 2013.
- [14] Z. Liu and B. J. Jansen, "Factors influencing the response rate in social question and answering behavior," in *Proc. CSCW*, 2013.
- [15] R. W. White, M. Richardson, and Y. Liu, "Effects of community size and contact rate in synchronous social Q&A," in *Proc. CHI*, 2011.
- [16] J. Lin, D. Quan, V. Sinha, K. Bakshi, D. Huynh, B. Katz, and D. R. Karger, "What makes a good answer? The role of context in question answering," in *Proc. INTERACT*, 2003.
- [17] J. Vertesi, "Seeing like a rover: Visualization, embodiment, and interaction on the mars exploration rover mission," *Social Studies of Science*, vol. 42, no. 3, 2012.
- [18] A. Spink, D. Wolfram, M. B. J. Jansen, and T. Saracevic, "Searching the web: The public and their queries," *Journal of Am. Soc. Info. Sci. & Tech.*, vol. 52, no. 3, 2001.
- [19] M. Taghavi, A. Patel, N. Schmidt, C. Wills, and Y. Tew, "An analysis of web proxy logs with query distribution pattern approach for search engines," *Computer Standards & Interfaces*, vol. 34, no. 1, 2012.
- [20] M. S. Ackerman and T. W. Malone, "Answer garden: A tool for growing organizational memory," *SIGOIS*, vol. 11, no. 2-3, 1990.
- [21] B. M. Evans and E. H. Chi, "Towards a model of understanding social search," in *Proc. CSCW*, 2008.
- [22] M. Efron and M. Winget, "Questions are content: A taxonomy of questions in a microblogging environment," in *Proc. ASIS&T*, 2010.
- [23] J. Teevan, D. Ramage, and M. R. Morris, "#TwitterSearch: A comparison of microblog search and web search," in *Proc. WSDM*, 2011.
- [24] C. Lampe, J. Vitak, R. Gray, and N. Ellison, "Perceptions of facebook's value as an information source," in *Proc. CHI*, 2012.
- [25] M. R. Morris, T. Jaime, and K. Panovich, "A Comparison of Information Seeking Using Search Engines and Social Networks," in *ICWSM*, 2010.
- [26] R. Gray, N. B. Ellison, J. Vitak, and C. Lampe, "Who wants to know?: Question-asking and answering practices among facebook users," in *Proc. CSCW*, 2013.
- [27] K. Sparck Jones, "A statistical interpretation of term specificity and its application in retrieval," *Journal of Documentation*, vol. 28, no. 1, 1972.
- [28] J.-W. Jeong, M. R. Morris, J. Teevan, and D. Liebling, *A Crowd-Powered Socially Embedded Search Engine*, ser. 2013. AAAI, 2013.
- [29] E. Wenger, *Communities of Practice: Learning, Meaning, and Identity*. Cambridge: Cambridge University Press, 1998.
- [30] J. Pierce, Y. Strengers, P. Sengers, and S. Bodker, "Introduction to the special issue on practice-oriented approaches to sustainable HCI," *ACM Trans. Comput.-Hum. Interact.*, vol. 20, no. 4, 2013.
- [31] A. Reckwitz, "Toward a theory of social practices: A development in culturalist theorizing," *European J. of Social Theory*, vol. 5, no. 2, 2002.
- [32] E. Shove, M. Pantzar, and M. Watson, *The Dynamics of Social Practice: Everyday Life and How it Changes*. SAGE Publications Ltd, 2012.
- [33] K. Schmidt, *Cooperative Work and Coordinative Practices: Contributions to the Conceptual Foundations of Computer-Supported Cooperative Work (CSCW)*. London: Springer, 2011.
- [34] K. Kuutti and L. J. Bannon, "The turn to practice in HCI: towards a research agenda," in *Proc. CHI*, 2014.
- [35] D. Nicolini, *Practice Theory, Work, and Organization. An Introduction*. UK: Oxford University Press, 2013.
- [36] T. R. Schatzki, *Social Practices: A Wittgensteinian Approach to Human Activity and the Social*. UK: Cambridge University Press, 1996.
- [37] H. S. Ferdous, D. Das, and F. M. Choudhury, "Social media question asking (smqa) whom do we tag and why?" in *Proceedings of the 30th Australian Conference on Computer-Human Interaction*, 2018.
- [38] J. S. Bauer, M. W. Newman, and J. A. Kientz, "What designers talk about when they talk about context," *Human-Computer Interaction*, vol. 29, no. 5-6, 2014.
- [39] J. Kiseleva, "Context mining and integration into predictive web analytics," in *Proc. WWW*, 2013.
- [40] D. Bamman and N. A. Smith, "Contextualized sarcasm detection on Twitter," in *Assoc. for the Advancement of Artificial Intelligence*, 2015.
- [41] Y. R. Tausczik, A. Kittur, and R. E. Kraut, "Collaborative problem solving: A study of MathOverflow," in *Proc. CSCW*, 2014.