ARTICLE

Check for updates

Key traits of top answerers on Korean Social Q&A platforms: insights into user performance and entrepreneurial potential

Moonkyoung Jang[1] & Seongcheol Kim [2✉]

Despite the proliferation of tools available to search for information online, such as search engines and generative artificial intelligence (AI), many people still actively use social question-and-answer (Q&A) platforms. This study examines the best answerers on Korean social Q&A platforms as key content producers and potential entrepreneurs. The characteristics of the best answerers, who are the core contributors to social Q&A platforms, were analysed based on the 903 best answerers from the social Q&A platform Naver Knowledge-iN, a leading social Q&A platform in South Korea. Research hypotheses were formulated based on three key dimensions: the intrinsic characteristics of the answers (length, immediacy, and similarity with the question), attributes of the answerer (self-introduction, credentials beyond the social Q&A platform), and level of engagement of the answerer on the platform (variety in the field of answers, diversity in role activities on the platform). An in-depth analysis using a multiple regression model was conducted to test the hypotheses. The findings of the analysis results indicate that enhanced user performance is positively correlated with several factors. Specifically, longer answers, high similarity between questions and answers, additional credentials beyond the social Q&A platform, a broader range of fields answered, and active participation as both asker and answerer all contribute to improved user performance. These findings highlight what sets social Q&A platforms apart from other search tools such as search engines or generative AI chatbots. This study has practical implications for social Q&A platforms seeking to identify the best contributors who are likely to produce quality content and are potential entrepreneurs.

[1] Business School, Gachon University, 1342 Seongnamdaero, Seongnam 13120, Republic of Korea. [2] School of Media and Communication, Korea University, 145 Anam-Dong Seongbuk-Gu, Seoul 02841, Republic of Korea. ✉email: hiddentrees@korea.ac.kr

Introduction

Advances in information and communication technology have made the Internet a convenient place for people to obtain information. Even as technology has increased the number of tools available to find information on the Internet, such as search engines and generative (Gen) artificial intelligence (AI) chatbots, people continue to use social question-and-answer (Q&A) platforms (Fang and Zhang, 2019). While global social Q&A platforms, such as Google Questions and Yahoo! Answers experienced severe downturns and shutdowns, social Q&A services that specialise in specific regions or languages remain popular, such as Naver Knowledge-iN (kin.naver.com) in South Korea, Baidu Zhidao (zhidao.baidu.com) in China, and Gutefrage (gutefrage.net) in Germany (Zheng et al., 2021).

While people engage with social Q&A platforms to obtain answers to their questions, other factors are known to motivate people to ask questions on social Q&A platforms as well, including a combination of social and emotional factors, user interaction, dependency on user attachment, and emotional or other social factors such as a shared language and vision (Gazan, 2010; Fang and Zhang, 2019; Wang et al., 2021; Zheng et al., 2021). Social Q&A platforms are also evolving from people participating to gain something without monetary value, such as reputation, to paid social Q&A services where askers obtain more professional answers and answerers receive financial incentives (Liu et al., 2021; Ye et al., 2021). These characteristics related to human interaction are difficult for Gen AI chatbots to replace completely, indicating that social Q&A platforms will continue for the foreseeable future.

User-generated questions and answers are the core content of social Q&A platforms. Users answer questions based on their expertise and experience and share rich, useful information. These answers are perceived to provide valuable knowledge and be helpful to other users. Answerers invest time and effort into carefully answering questions from other users and strive to maintain originality and quality (Fang and Zhang, 2019). Among the numerous users of social Q&A platforms, only a small group produce high-quality content (i.e., answers) is thus exhibit different behaviour on these platforms compared to other users (Shen et al., 2014; Zheng et al., 2021). Social Q&A platforms should manage these few answerers (i.e., the best contributors) well because they play a major role in providing the core content for social Q&A platforms, providing valuable information, and helping users (Nam et al., 2009). Their work contributes significantly to the quality and value of social Q&A platforms overall. In addition, managing the best contributors helps to maintain and support their passion and creativity (Zheng et al., 2021).

Furthermore, a small number of the best contributors on social Q&A platforms have the potential to become entrepreneurs. Therefore, they play a vital role and need to be well managed by the platform. An entrepreneur is an individual who is responsible for starting, managing, and operating a business or venture to make a profit (Portuguez et al., 2021). The best answerers on social Q&A platforms can be creative in building their reputation and recognition around a skill and an area of expertise while providing valuable information and helping users. Thus, they may have opportunities in the future to translate their expertise into commercial value and pursue independent businesses or entrepreneurship. For the small group of best answerers with the potential to become entrepreneurs, networking and building trust through social Q&A platforms can be an important starting point for the development of future business opportunities.

Thus, the best answerers on social Q&A platforms are important not only as content generators but also as potential entrepreneurs. Therefore, the characteristics that make social Q&A platform users the best answerers need to be understood. This study explored this topic based on the following research question:

- RQ: What are the characteristics of the best answerers on Korean social Q&A platforms?

Presumably, the best answerers perform well by writing quality answers that are highly rated by social Q&A users. Therefore, in this study, we consider a user to be a better answerer if their answers are selected by askers at a high rate. We used data from Naver Knowledge-iN, a popular South Korean social Q&A platform. We collected the best answerers' profile information and activity data and performed multiple regression analysis to identify their characteristics. The results indicate enhanced answerer quality under the following conditions: (1) longer answers, (2) greater similarity between questions and answers, (3) a higher number of credentials beyond social Q&A platforms, and (4) answers across diverse subject areas. Based on this study's findings, users with such characteristics can be managed separately as potential entrepreneurs to encourage them to voluntarily and continuously produce quality content (i.e., answers) on social Q&A platforms.

The remainder of this paper is organised as follows. Section 2 reviews the literature on common methods for searching for answers online and social Q&A entrepreneurs, to formulate our hypotheses. Section 3 describes the research data, variables, and methodology, and Section 4 reports the results. Finally, Section 5 discusses the significance of this study and future research directions.

Literature review

Obtaining answers online. People actively utilise various tools to obtain answers online, including search engines, social Q&A platforms, and more recently, Gen AI (Table 1). These tools differ in the language used when searching for answers. For example, search engines are keyword-based and require users to enter a refined word or phrase to search for the information, and then

Table 1 Common ways to get answers online.			
Category	**Features**	**Examples**	**Refs.**
Search engine	● Keyword-based search	Google, MS Bing, Yahoo!, Naver (Korea), Baidu (China)	Shen & Wang (2017)
Social Q&A platform	● Search using natural language ● A small number of best contributors influence the overall volume and quality of contents (i.e., answers)	Google Answers Closed 2006, Yahoo! Answers (Closed 2021), Naver Knowledge-iN, Baidu Knows, Raddit, Stachoverflow	Fang & Zhang (2019), Nam et al. (2009), Shen et al. (2014), Zheng et al. (2021)
Gen AI chatbot	● Search using natural language ● Concerned about the reliability and accuracy of information (e.g., hallucination)	OpenAI ChatGPT, MS Bing Chat, Google Bard	Koubaa et al. (2023), Rudolph et al. (2024)

search repeatedly until the desired information is found (Shen and Wang, 2017). In contrast, social Q&A platforms and Gen AI chatbots based on natural language processing (NLP) allow people to search for information as if they are having a conversation in their own language. Gen AI chatbots are AI-powered systems that generate human-like responses in conversational interactions (Nicolescu and Tudorache, 2022). While previous large language models have been able to perform various NLP tasks, recent Gen AI chatbots are optimised for conversations and particularly adept at talking in a human-like manner (De Angelis et al., 2023). Unlike rule-based chatbots that rely on predefined rules or templates, Gen AI chatbots can generate unique and contextualised responses based on input, allowing for more natural and informative interactions with users (Adamopoulou and Moussiades, 2020). Thus, people can use natural language to search for information on social Q&A platforms and with Gen AI chatbots.

However, social Q&A platforms and Gen AI chatbots serve different purposes and have distinct characteristics regarding how people ask questions. First, Gen AI chatbots focus on generating new content based on patterns and structures learned from existing data (Ray, 2023). When someone asks a Gen AI chatbot a question, the system generates a response or output based on its learning, and these answers are generated by an algorithm rather than based on human expertise or personal experience (Koubaa et al., 2023). However, on social Q&A platforms, real people ask and answer the questions. Users of these platforms ask questions to receive answers, insights, opinions, or solutions from other users who have knowledge or experience in a related field, and the answers provided on social Q&A platforms are from people who share their expertise, insights, and personal experiences (Fang and Zhang, 2019; Zheng et al., 2021).

Second, Gen AI chatbots learn from vast amounts of data collected from various sources on the Internet. The responses Gen AI chatbots provide are based on the patterns and information present in the training data; however, the accuracy and reliability may vary because the accuracy of the information generated by the model cannot be verified (Koubaa et al., 2023; Rudolph et al., 2024). However, on social Q&A platforms, answerers provide information directly based on their knowledge, expertise, or experiences, and users can evaluate credibility and accuracy based on factors such as the contributor's reputation, expertise, and community feedback (Jin et al., 2015).

Third, when people ask Gen AI chatbots questions, responses are typically generated through a one-way process. People ask questions and the Gen AI model responds without further interaction or explanation. Gen AI models may be unable to engage in dynamic conversations or seek additional information to provide more accurate or contextualised answers (Koubaa et al., 2023). However, on social Q&A platforms, users can have interactive discussions with the answerer, such as by asking follow-up questions, seeking clarification, and engaging in back-and-forth conversations (Jin et al., 2015).

Finally, Gen AI chatbots lack the personal experience, opinions, and domain-specific expertise of human contributors (Koubaa et al., 2023). Although Gen AI chatbots can generate content based on patterns, they may not provide the same level of depth, context, or domain expertise as human-authored answers. However, on social Q&A platforms, human contributors can provide answers that reflect their knowledge, expertise, and personal insights (Fang and Zhang, 2019; Zheng et al., 2021), share real-world stories, offer nuanced perspectives, and provide contextual recommendations beyond those Gen AI chatbots.

Top contributors on social Q&A platforms. Contributors are the main content producers on social Q&A platforms because the questions and answers users voluntarily contribute are the core content. The quality of social Q&A platform services is also significantly affected by how many high-quality answers are consistently generated (Fang and Zhang, 2019); thus, social Q&A platforms want to know which users produce quality answers. Social Q&A platforms have heavy-tailed activities in which only a few answerers have many comments (Nam et al., 2009), and these answerers behave differently than other answerers or lurkers (Shen et al., 2014; Zheng et al., 2021). Although some users answer tens of thousands of questions, askers rarely ask more than a few hundred questions (Nam et al., 2009). Therefore, while it is unlikely that high-contributing askers (i.e., contributors who ask many questions) contribute significantly to site activity, high-contributing answerers (i.e., contributors who answer many questions) can help many askers. Furthermore, these few answerers differ from others in both the quantity and quality of their answers (Nam et al., 2009). Therefore, social Q&A platforms rank users and attempt to manage their best answerers as content providers (Fang and Zhang, 2019; Jin et al., 2016; Nam et al., 2009; Yan and Zhou, 2015; Zheng et al., 2021).

In addition to their ability to provide content, the best answerers on social Q&A platforms have the potential to become entrepreneurs based on their expertise, reputation, and ability to create content, similar to influencers on social media platforms (Bi and Liu, 2022). An entrepreneur is an individual who is responsible for starting, managing, and operating a business or venture to make a profit (Portuguez et al., 2021). On social Q&A platforms, the best answerers are free to create content (i.e., answers) based on their expertise and knowledge without being bound by specific guidelines or content requirements set by the platform. This independence allows them to develop unique approaches, styles, and brands as content creators. In addition, the best answerers can build strong reputations both on and off the platform through consistent, high-quality contributions. Such a reputation can attract devoted followers and build trust and credibility among audiences. Building a reputation is a crucial aspect of starting a business because it opens the door to various opportunities (Xie and Lv, 2018). Although immediate monetisation options within social Q&A platforms may be limited, the best answerers can leverage their expertise and reputation to explore different revenue streams. For example, they can expand their presence to other platforms such as by creating blogs or social media accounts, allowing them to explore revenue streams, including sponsored content, affiliate marketing, digital product sales, and consulting services. Furthermore, the best answerers can consider expanding their business beyond content creation as they gain recognition and build their brand. They can also look for opportunities for speaking engagements, workshops, training programmes, or collaborations with other businesses or organisations in their niche. Expanding into various entrepreneurial ventures can help answerers grow and diversify their businesses.

Furthermore, if the same platform offers free and paid social Q&A services, the best answerers with entrepreneurship potential are likely to be active in both. The best answerers can contribute their expertise and knowledge to a paid social Q&A service and be compensated for it, providing them the opportunity to leverage the reputation and knowledge they have built on the free social Q&A service to provide specialised services and work independently. This allows answerers to create greater value and explore business opportunities by responding to consumer needs in their areas of expertise. The best answerers who participate in both free and paid social Q&A services on the same platform can be considered platform-dependent entrepreneurs. A platform-dependent entrepreneur builds or relies heavily on a specific

platform or technology infrastructure, usually provided by another company or organisation (Cenamor et al., 2019; Cutolo and Kenney, 2021; Nambisan and Baron, 2021). Platform-dependent entrepreneurs leverage platforms to create, scale, and deliver products or services without having to build everything from scratch (Cenamor et al., 2019; Cutolo and Kenney, 2021; Nambisan and Baron, 2021). Thus, platform-dependent entrepreneurs are characterised by their dependence on a platform and cannot leave it easily. From a social Q&A platform perspective, encouraging the best answerers on a free social Q&A service to switch to a paid service can strengthen platform loyalty and attract and retain users who offer expertise through a service. Social Q&A platforms can encourage users to continue providing free services while building their expertise by earning rewards through paid services. Therefore, social Q&A platforms can support their best answerers to ensure the continued growth and development of the platform.

Hypotheses. Social Q&A platforms have their own methods of determining the best answerers among a large group of users. For example, social Q&A platforms can rank members based on certain criteria, such as the number of times they log onto the platform, write a question or answer, or upvote or downvote other answers (Fang and Zhang, 2019; Nam et al., 2009; Yan and Zhou, 2015; Zheng et al., 2021). These methods are frequently updated; however, what is important for social Q&A platforms is that users consider answers written by a particular user as reliable and useful. In this context, most social Q&A platforms have a feature that allows askers to select the answer they find the most useful among many (Jin et al., 2016), allowing the best answerer to be defined as the user whose answers are selected the most by askers.

Among many answers, askers consider various aspects of an answer and select the best. This decision-making process can be described using the dual-process theory of information processing (James, 2007), which suggests that decision-making and information processing involve two distinct cognitive processes: systematic processing and heuristic processing. Systematic processing, or the central route, involves deliberate thinking, logical reasoning, and cognitive effort to engage in systematic analysis, problem-solving, and decision-making based on available information. Heuristic processing, or the peripheral route, relies on heuristics, associations, and intuitive judgments to make decisions and is driven by emotions, biases, and past experiences. This process generates rapid responses and intuitive judgments without conscious reasoning. These dual processes work concurrently rather than separately during personal information processing. When selecting the best answers on a social Q&A platform, askers can easily see both the answer and related information (e.g., author, date, and comments) because the platform displays these together. Therefore, systematic processing allows askers to decide whether to select an answer as the best based on the response itself, while heuristic processing allows askers to decide whether to select the best answer based on related information (Jin et al., 2016). In this context, we formulated hypotheses based on four primary perspectives to comprehend the characteristics of the best answerers: the content of the answer, the characteristics of the answer's author, the level of activity exhibited by the answerer as a platform user, and social endorsement. These factors are pivotal in determining what answers are chosen as the best on social Q&A sites.

First, answer content itself has various aspects, including length, immediacy, and similarity. The longer the answer, the more likely the asker is to select it as the best answer (Kim and Oh, 2009; Qi et al., 2021). More immediate answers are provided

almost as soon as the question is posted, making it more likely askers will select these answers as the best (Fu and Oh, 2019; Gazan, 2010; Jin et al., 2016). Regarding similarity, the more similar the response an answerer provides is the question, the more likely it is the asker will choose that it as the best (Jin et al., 2016; Fu and Oh, 2019). Finally, the higher the ratio at which a user-written answer is selected as the best, the better the answerer will be. Therefore, we propose the following hypotheses:

- H1-1: The longer the answers a user writes, the more likely that user is to be the best answerer.
- H1-2: The more immediate the responses of a user, the more likely that user is to be the best answerer.
- H1-3: The more similarity a user's response shares with the question, the more likely the user is to be the best answerer.

Second, regarding the information related to the answer, the asker can access information about the answerer provided by the social Q&A platform (Dong et al., 2021; Ginsca and Popescu, 2013; Zhou et al., 2012). Information about answerers can be broadly divided into information answerers write themselves and information about their activity on the platform. Because the answerer serves as the source of the answer, the asker assesses the quality of the answer based on the available information about the answerer (Jin et al., 2016; Shah and Pomerantz, 2010). Regarding self-described information, social Q&A platforms permit users to write details about themselves on their profile pages, which are visible to all users. The answerer's profile is recognised as a source characteristic; the greater the degree of openness in the user's personal information, the more trustworthy other users perceive them to be (Jin et al., 2016; Shah and Pomerantz, 2010). Thus, askers tend to trust answerers more when they provide more detailed profiles (Jin et al., 2016; Liang, 2017; Liu et al., 2021). Moreover, considering that social Q&A platforms mainly focus on answering questions, the expertise of their users is highly valued (Liu et al., 2021). Therefore, users with generally accepted credentials will presumably be more reliable (Kim, 2010). Accordingly, we propose the following hypotheses:

- H2-1: The more detailed a user's self-introduction, the more likely that user is to be the best answerer.
- H2-2: The more credentials a user has, the more likely that user is to be the best answer.

Third, regarding the activity history of an answerer on a social Q&A platform, askers can judge the quality of the answer by looking at their past posting behaviour (Gazan, 2010). Individuals who actively participate in various activities on a given platform earn a good reputation (Liu et al., 2021; Ye et al., 2021), and social Q&A platforms can be considered similar. Thus, on a social Q&A platform, the asker can determine the quality of the answer based on the breadth of the fields in which the answerer mainly provides answers. Those who contribute answers across different domains on social Q&A platforms possess diverse interests in a variety of topics and willingly share their professional expertise on the platform. This enables users in various fields to search for answers and obtain reliable information. In addition, users who are active in several fields can interact with other users and form a rich community, enabling active participation on the platform and an exchange of opinions. Furthermore, as users provide answers in various fields, platform utilisation increases. This raises awareness that the platform is suitable for various topics and users, helping attract new users and maintain existing ones.

On social Q&A platforms, users are not necessarily required to choose between being an asker or answerer; they can act in both roles as information providers and seekers (Adamic et al., 2008; Fu and Oh, 2019). Askers may believe that if answerers are users
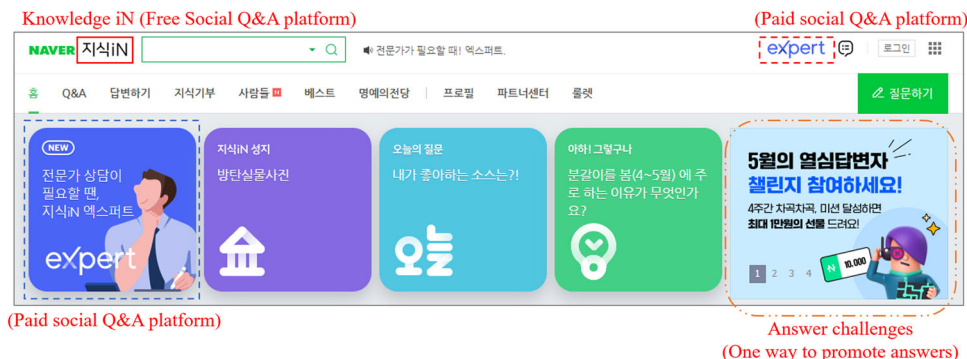
Fig. 1 Screenshot of Naver Knowledge-iN.

who have played both roles on the platform, they will be more likely to have knowledge on a wider range of activities and be more engaged with the platform. Therefore, askers may believe that these active users will provide more credible answers. Accordingly, we propose the following hypotheses:

H3-1: The more fields in which a user provides answers, the more likely that user is to be the best answerer.
H3-2: The more active a user is as both an asker and answerer, the more likely the user is to be the best answerer.

Finally, social Q&A platforms are not only important to askers and answerers but also to an overwhelming majority of lurkers. On social Q&A platforms, not only can the asker select the best answer from the answerers' post but lurkers can also react by upvoting or downvoting an answer (Fang and Zhang, 2019). If a lurker responds positively to an answer written by a user, their quality can be considered sufficient (Jin et al., 2016). Therefore, we propose the following hypothesis:

- H4: The more positive responses a user's answer receives from lurkers, the more likely the user is to receive the best answer.

Methodology
Research site: Naver Knowledge-iN. This study utilises data from Naver Knowledge-iN (kin.naver.com), a leading South Korean social Q&A platform, to identify the characteristics of the best answerers. Naver Knowledge-iN is a section of the platform launched in October 2002, where users can ask questions and obtain answers from the community. At the time Naver Knowledge-iN was launched, Google was gaining ground as a search engine for English content; however, it showed poor performance in searching for Korean content, and Korean content was also scarce online. Thus, Naver attempted to improve the performance of its search engine, while encouraging people to generate Korean content by asking and answering questions on Naver Knowledge iN. Through this service, Naver built 1.65 million knowledge DBs in a year because of the explosive response from young users, helping Naver become the leading portal site in Korea (Chae, 2003). Since its launch, 32 million cumulative users have communicated with Naver Knowledge-iN over the past two decades, and the number of questions and answers they have generated has been 300 million and 500 million, respectively, with a total of 800 million databases accumulated. In 2021, Naver Knowledge-iN added 1 million new questions and 730,000 new answers and currently averages nearly 30 million pageviews per day (Cho, 2022).

Naver Knowledge-iN offers various categories and topics that users can explore. Users can ask questions on topics such as technology, health, travel, entertainment, and education. The platform has a large user base, which people use to seek information, opinions, and advice from fellow users. Similar to other social Q&A platforms, Naver Knowledge-iN allows users to vote on answers, and the most popular and helpful answers receive higher visibility. Users can also participate in discussions, comment on answers, and build a reputation based on their contributions. In addition, as with other social Q&A platforms, user contributions are essential for the sustainability of Naver Knowledge iN; therefore, the platform has been attempting various methods to encourage more users to write better answers and participate in knowledge-sharing. For example, the platform offers monetary rewards, virtual points, badges, and rankings. In addition, the platform provides a profile page for each user and a threaded Q&A page (Figs. 2 and 3).

Naver operates Naver Knowledge iN, a free social Q&A platform, and Naver eXpert, a paid social Q&A platform (expert.naver.com). On Naver eXpert, users can ask questions and obtain answers from verified experts in various fields, such as medicine, law, and finance. Naver eXpert operates on a fee-based model, meaning that users must pay a certain amount of money to ask an expert a question or access premium content provided by experts. Users are provided the opportunity to obtain professional advice and insights from experts in a specific field, and the experts are financially compensated for the knowledge and expertise they share on the platform. Because answerers of Naver eXpert can provide knowledge and monetise it, they can be considered entrepreneurs in one-person businesses. They rely on a specific platform to function; thus, they can be considered platform-dependent entrepreneurs. They rely on the platform's infrastructure, user base, and features to connect with and serve their audience and monetise their content based on the platform's revenue model. Thus, they do not rely on the platform to run their entire business, but their ability to reach customers, generate revenue, and build their reputation is highly dependent on the features and opportunities it provides.

As Fig. 1 shows, Naver eXpert and Naver Knowledge-iN are provided on the same Internet portal, Naver. Therefore, the best contributors with entrepreneurship on Naver Knowledge-iN are likely to utilise the reputation and knowledge they have built on the free social Q&A service to work on the paid social Q&A service for financial rewards, transforming them into platform-dependent entrepreneurs. In this context, Naver strives to manage the best contributors to the free social Q&A service to encourage their conversion to the paid social Q&A service, thereby strengthening users' dependence on the platform and attracting users who offer their expertise. Attracting the top contributors

Fig. 2 Example of user profile page in Naver Knowledge-iN.

from Naver Knowledge-iN to Naver eXpert has several benefits for Naver as a platform. First, it can build expertise and trust on the platform. The best contributors have built a reputation for Naver Knowledge-iN and have a high level of expertise. By bringing them to Naver eXpert, the platform can gather customers who offer specialised knowledge and services. Other users can trust the expertise of Naver eXpert, thereby increasing the value of the platform. Second, by attracting the best Naver Knowledge-iN contributors to Naver eXpert, Naver maintains the quality and consistency of the content provided on the platform. The best contributors provide authoritative answers based on their knowledge and experience and offer valuable information to platform users. This helps improve the platform's reputation and user experience. Third, answerers on Naver eXpert can provide paid services to answer questions and receive compensation for doing so; therefore, if the best contributors are converted to Naver eXpert, the platform can monetise their services which will help the platform continue to operate and grow.

Research data. For this study, we developed a Python crawler to collect the profile and activity data of the top 1000 contributors on Naver KnowledgeiN. First, as shown in Fig. 2, we crawled the profile information of these users to gather their profile information as of 15 October 2022. To investigate these users' answering activity, we randomly selected 100 posts from their replies and collected the corresponding question posts for each reply. Among the targeted users, we collected all replies written by users with fewer than 100 selections. Ultimately, we analysed 903 users after excluding posts with zero characters and those that showed an excessive difference between the question-and-answer dates.

To measure the answerers' performance as the dependent variable, we calculated the percentage of answers selected by the askers from the answers written by the answerer.

As shown in Fig. 3, the independent variables related to the content of users' answers are a key part of their activity as top contributors; therefore, we measured three aspects according to our research hypotheses. For H1-1, we assessed the level of detail and information in responses by considering response length. The level of detail or information can be examined based on both quantitative and qualitative. Quantitative aspects such as the average number of words are relatively straightforward to measure. However, qualitative aspects, such as the content, structure, and logical flow of a response, pose challenges owing to the time-consuming, expensive, and subjective nature of expert evaluations. NLP techniques offer a potential avenue for evaluating the qualitative aspects of responses; however, their effectiveness is hindered by their reliance on domain-specific analysis. In particular, the overall accuracy in the context of Korean-based NLP techniques remains relatively low (Kim et al., 2022). Despite the challenge of equating specificity solely with response length, longer answers are considered to provide more information and are more likely to be specific (Peng et al., 2020; Qi et al., 2021). Consequently, we measured the level of detail or the amount of information in a response by calculating the average number of words, providing a broad gauge of specificity. In addition, analysing the average word count across multiple responses from a user allows us to discern individual response writing patterns. A consistently high or low average word count suggests that users tend to express themselves in a specific manner.
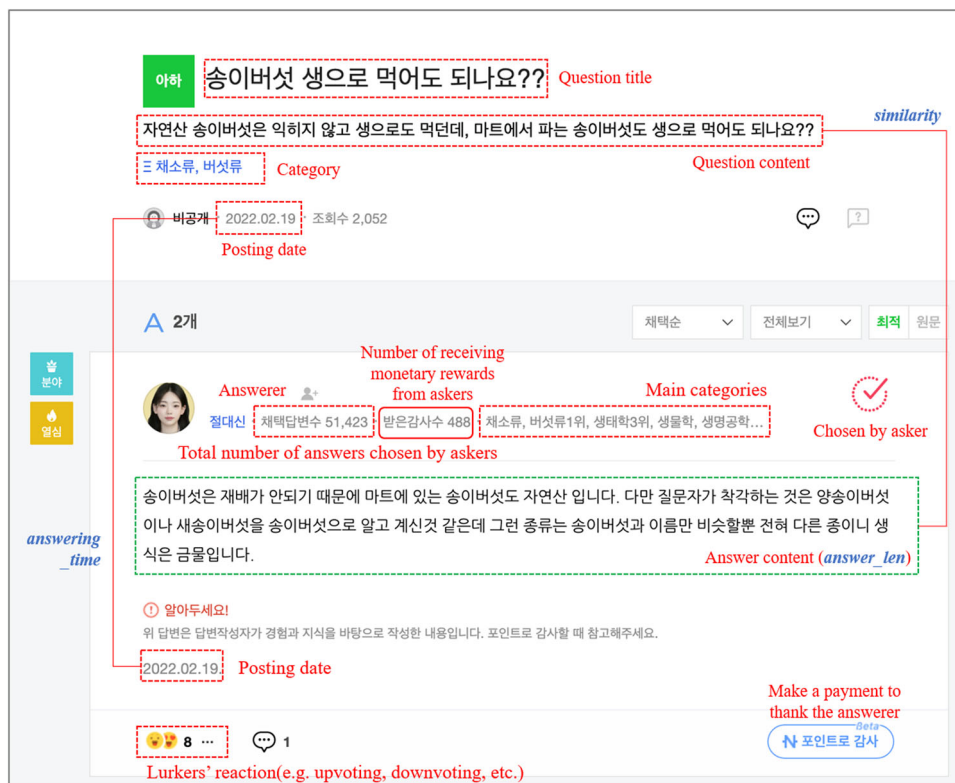
Fig. 3 Example of a question-answer exchange in Naver Knowledge-iN.

Regarding the independent variables related to the answerer's information, we first measured the degree of openness of the user's personal information, which reflects the degree of detail that the users provide to describe themselves. This was measured based on the number of words on a user's profile page, aligned with the same rationale applied to the quantitative measurement of answer specificity (H2-1). We also measured the number of credentials (H2-2) users included in their profiles, as the level of credentials is an important factor in gaining askers' trust on social Q&A platforms.

In terms of the answerer platform activity, we measured the number of activity categories in the answer threads written by the users to measure the scope of their activity on the social Q&A platform (H3-1). We measured the ratio of their selected best answers to posts written they wrote as askers to measure the extent of their contribution as an asker and answerer on the social Q&A platform (H3-2).

Regarding the level of social endorsement, we measured the number of upvotes received on a post written by that user to measure the degree of response from lurkers to a reply written by the user (H4).

Table 2 summarises the variables used to test the hypotheses. For H1-2, we calculated the difference between the creation time of the question post and that of the answer post to measure the immediacy of the responses. Finally, for H1-3, to measure the similarity between the question-and-answer posts, we calculated the average cosine similarity (Jin et al., 2016) between them after excluding questions from the analysis if the thread was private. The detailed steps for calculating the cosine similarity between question-and-answer posts are as follows. To analyse the content of the answers, which are unstructured documents, we must first perform a structuring process. Accordingly, we performed the standard preprocessing steps of an NLP task: converting each document into a vector and tokenising the raw text. To tokenise the answers written in Korean, we used the KoNLPy library for Korean NLP in Python and the verified Python-based morphological analyser MeCab (Kang and Yang, 2018). All words were tagged as parts of speech based on their syntactic categories, and a vector space model was used to represent each document as a point in a vector space with one dimension for each term in the vocabulary. Each document d is represented by Eq. (1), where t_i denotes the term contained in d and f_i denotes the frequency of occurrence of that term in d.

$$d = \{(t_1, f_1), (t_2, f_2)\ldots(t_i, f_i)\} \tag{1}$$

The cosine similarity between question d_q and its answer d_a is provided by Eq. (2), where f_{qi} and f_{ai} represent the frequencies of term t_i in d_q and d_a, respectively; n is the number of terms contained in d_q and d_a.

$$cosine\,(d_q, d_a) = \frac{\sum_{i=1}^{n} f_{qi} \times f_{ai}}{\sqrt{\sum_{i=1}^{n} f_{qi}^2} \sqrt{\sum_{i=1}^{n} f_{ai}^2}} \tag{2}$$

Empirical analysis results

Descriptive statistics. We first conducted a descriptive analysis of the 903 best users as of 15 October 2022, and Table 3 summarises the descriptive statistics of the variables and the correlation analysis results. The mean value of the respondents' performance was 0.847 with a standard deviation of 0.150. This indicates that approximately 85% of the answers written by the best answerers were selected as the best answers. The number of words in the self-introduction of the answerer's profile ($r = 0.070$, $p < 0.05$), number of credentials of the answerer outside the social Q&A platform ($r = 0.099$, $p < 0.05$), average cosine similarity between questions and answers ($r = 0.095$, $p < 0.05$), number of fields to which the answers belongs ($r = 0.136$, $p < 0.001$), and percentage of posts and answers selected by the answer ($r = 0.114$,

Table 2 Variables.

Variables		Name	Description	Measurement	Hypothesis
Dep.	Answerer's performance	performance	Performance of user i	Of the answers made by user i, ratio of answers selected by the asker	
Indep.	Answer	answer_len	The length the answers of user i	Average word count in answers written by user i	H1-1
		answering_time	The time it takes for user i to write an answer	Average difference in creation time for questions of other users and answers of user i	H1-2
		similarity	The degree of similarity in content between the question and the answer	Average of cosine similarity between questions of other users and answers of user i	H1-3
	Answerer	profile_len	The length of self-introduction to user i	The number of words in self-introduction of user i's profile page	H2-1
		credentials	Credentials of user i outside of the social Q&A platform	The number of credentials of user i outside of the social Q&A platform	H2-2
	Platform activity	answer_diversity	Scope diversity of answers of user i in the social Q&A platform	The number of fields the answers written by user i belongs to	H3-1
		as_asker	The degree that user i acts as an asker in the social Q&A platform	The percentage of posts written by user i and answers selected by user i	H3-2
	Social endorsement	lurker_likes	The degree to which lurkers respond positively to answers of user i	Average number of recommendations (likes) for a user i's answer	H4
Control		tenure	Duration of user i in the social Q&A platform	Days between data collection date and first response date	

Table 3 Descriptive statistics and correlation matrix (903 best users).

	Mean	Std. Dev.	Min	Max	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(1) performance	0.847	0.150	0.246	1	1									
(2) profile_len[a]	120.564	105.471	0	478	0.070**	1								
(3) credentials	0.171	0.578	0	3	0.099**	0.134***	1							
(4) similarity	0.098	0.127	0	0.9397	0.095**	0.111***	−0.013	1						
(5) answering_time	0.196	0.530	0	4.83	−0.011	0.080*	−0.030	0.271***	1					
(6) answer_len[a]	176.186	502.341	0	7429.28	−0.033	0.164***	−0.029	0.319***	0.269***	1				
(7) answer_diversity	3.085	1.3560	0	4	0.136***	0.256***	0.071**	0.072**	0.048	−0.034	1			
(8) as_asker	0.805	0.314	0	1	0.114***	−0.022	0.050	0.001	−0.001	−0.067**	0.059	1		
(9) lurker_likes[a]	18473.890	54118.040	126	824,936	0.060	0.032	−0.011	0.268***	0.191***	0.206***	0.002	0.001	1	
(10) tenure	4173.513	1797.311	466	7,305	0.160***	0.039	0.015	0.116***	0.048	0.016	0.153***	0.061	0.182***	1

***p < 0.001, **p < 0.05, *p < 0.01.
[a]Before taking logarithm.

Table 4 Result of hierarchical regression analysis.

Variable		Model 1 Estimate (S.E.)	Model 2 Estimate (S.E.)	Model 3 Estimate (S.E.)	Model 4 Estimate (S.E.)
Answer	answer_len	0.014** (0.005)	0.014** (0.005)	0.016** (0.004)	0.015** (0.005)
	answering_time	−0.015 (0.010)	−0.016 (0.010)	−0.017 (0.010)	−0.018 (0.010)
	similarity	0.092 (0.040)	0.092 (0.040)	0.087** (0.040)	0.082** (0.041)
Answerer	profile_len		0.003 (0.004)	−0.000 (0.004)	−0.001 (0.008)
	credentials		0.026** (0.009)	0.024** (0.008)	0.025** (0.008)
Platform activity	answer_diversity			0.010** (0.003)	0.010** (0.004)
	as_asker			0.052** (0.015)	0.053** (0.016)
Social endorsement	lurker_likes				0.003 (0.004)
Control	tenure	0.000** (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.000** (0.000)
	Constant	0.721*** (0.025)	0.704*** (0.027)	0.641*** (0.032)	0.627*** (0.038)
R^2		0.042	0.054	0.075	0.075
Adj. R^2		0.038	0.048	0.067	0.066
R^2 change		-	0.012	0.021	0.000
F		9.910***	8.520***	9.040***	8.080***

$N = 903$; Dependent variable = performance
***$p < 0.001$, **$p < 0.05$.

$p < 0.001$) presented significant positive correlations with answerer performance.

Hypothesis testing. Hierarchical regression analysis performed with STATA (MP 14.2) was used to examine the research model. Hierarchical regression offers a distinct advantage in detecting the total variance of the outcome variables. This was demonstrated using the incremental increase in the coefficient of determination, the proportion of variance in the dependent variable that is explained by the independent variables (Petrocelli, 2003). The sequential arrangement of independent variables was methodically predetermined, guided by the research objectives and rationale. The hierarchical order of entry was strategically structured, emphasising the prioritisation of causal relationships, elimination of confounding or spurious relationships, alignment with research goals, and consideration of the structural characteristics of the factors under investigation (Cohen et al., 2003). The analytical process was divided into two blocks, each of which encapsulated a distinct set of independent variables (Falvo and Earhart, 2009). Table 4 presents the results of the hierarchical regression analysis.

In Model 1, the effect of the average answer word count on answerer performance was significant ($F = 9.910$, $p < 0.001$, $R^2 = 0.042$), indicating that a higher word count for answers ($\beta = 0.014$, $p < 0.05$) was associated with higher answerer performance.

Model 2 showed that the effect of the average answer word count and the answerer's number of credentials outside the social Q&A platform on answerer performance was significant ($F = 8.520$, $p < 0.001$, $R^2 = 0.054$). Compared with Model 1, the change in explanatory power was 1.2%, which was significant ($p < 0.001$). This indicates that the greater the word count for answers ($\beta = 0.014$, $p < 0.05$) and the greater number of credentials an answerer has outside the social Q&A platform ($\beta = 0.026$, $p < 0.05$), the higher the answerer performance.

Model 3 demonstrated that the effects of the average answer word count, average cosine similarity between questions and answers, answerer's number of credentials outside the social Q&A platform, number of fields to which the answers belong, and percentage of posts and answers selected by the user on answer performance were significant ($F = 9.040$, $p < 0.001$, $R^2 = 0.075$). Compared with Model 2, the change in explanatory power was 2.1%, which was significant ($p < 0.001$). This indicates that, higher answer word count ($\beta = 0.016$, $p < 0.05$), average cosine similarity

between questions and answers ($\beta = 0.087$, $p < 0.05$), number of answerer credentials outside the platform ($\beta = 0.024$, $p < 0.05$) number of fields to which the answers belongs ($\beta = 0.010$, $p < 0.05$), and percentage of posts and answers selected by the user ($\beta = 0.052$, $p < 0.05$) are associated with higher answerer performance.

Model 4 also demonstrated that the average answer word count, average cosine similarity between questions and answers, answerer's number of credentials outside the social Q&A platform, number of fields to which the answers belong, and percentage of posts and answers selected by the user had significant effects on answerer performance ($F = 8.080$, $p < 0.001$, $R^2 = 0.075$). However, compared to Model 3, the change in explanatory power was found to be 0%. This indicates that higher answer word count ($\beta = 0.015$, $p < 0.05$), average cosine similarity between questions and answers ($\beta = 0.082$, $p < 0.05$), number of credentials of the answerer outside the platform ($\beta = 0.025$, $p < 0.05$), number of fields to which the answers belongs ($\beta = 0.010$, $p < 0.05$), and percentage of posts and answers selected by the user ($\beta = 0.053$, $p < 0.05$) are associated with higher answerer performance. Table 5 summarises the verification results for the hypotheses derived from the analysis of Model 4.

Discussion and conclusion
Discussion. This study analysed the data of 903 of the best answers from an actual social Q&A platform, Naver Knowledge iN, to identify the characteristics of contributors who generate high-quality answers, which are the core content of social Q&A platforms. The results of our hierarchical regression model show that longer answers, more similarity between the question and answer, more answerer credentials outside social Q&A platforms, greater number of categories to answer, and more active participation as an asker are associated with better performance among the top answerers.

Our analysis results can be explained as follows. First, regarding content characteristics, the predictive power of answer length for answerer performance was validated (H1-1), which is in line with our expectations and previous literature regarding the length of answers on social Q&A platforms (Fu and Oh, 2019; Jin et al., 2016). Longer answers are more likely to be comprehensive and informative, likely because they provide more context and elaboration, which can help users understand the answer better and apply it to their own situation. In addition, longer answers

Table 5 Hypothesis testing results.

Hypothesis		Supported or not
Contents	H1-1: The longer the answers a user writes, the more likely that user is to be the best answerer.	Supported
	H1-2: The more immediate the responses of a user, the more likely that user is to be the best answerer.	Not supported
	H1-3: The more similarity a user's response shares with the question, the more likely the user is to be the best answerer.	Supported
Answerer	H2-1: The more detailed a user's self-introduction, the more likely that user is to be the best answerer.	Not supported
	H2-2: The more credentials a user has, the more likely that user is to be the best answer.	Supported
Platform activity	H3-1: The more fields in which a user provides answers, the more likely that user is to be the best answerer.	Supported
	H3-2: The more active a user is as both an asker and answerer, the more likely the user is to be the best answer.	Supported
Social endorsement	H4: The more positive responses a user's answer receives from lurkers, the more likely the user is to receive the best answer.	Not supported

may convey a sense of sincerity and effort on the part of the answerer, which can positively influence the asker's perception of the answer quality.

However, answer immediacy was not found to have a predictive effect on answerer performance (H1-2), contrary to our expectations. Based on the existing literature (Fu and Oh, 2019; Gazan, 2010; Jin et al., 2016), we expected answer immediacy to have a positive effect on answerer performance; however, the results in our analysis were not significant. This suggests that the value placed on answer immediacy varies depending on the context and user expectations. What people value when asking a question differs between social Q&A platforms and Gen AI chatbots. When people ask a Gen AI chatbot a question, they want an immediate answer. However, when they ask a question on a social Q&A platform, they are looking for something other than immediacy. Social Q&A platforms cater to a more thoughtful and nuanced approach to information-seeking, whereas chatbots prioritise immediacy for quick on-the-go queries. Understanding these differences can help in designing and optimising these platforms to meet their users' specific needs.

The predictive effect of similarity between answers and questions on answerer performance was supported (H1-3). This suggests that contributors who understand the nuances of a question are more likely to provide helpful answers. This finding is in line with our expectations and previous literature regarding the similarity between questions and answers on social Q&A platforms (Fu and Oh, 2019; Jin et al., 2016). The ability to grasp the nuances of a question reflects a contributor's deep understanding of the subject matter and aptitude for effective communication. These qualities are essential for providing high-quality answers that are both informative and engaging, which ultimately contributes to the platform's overall success.

Second, regarding the profile information provided by the best answerers, the detail in a user's self-introduction did not have a significant effect (H2-1); however, the number of credentials did (H2-2), which is consistent with previous research (Kim, 2010). Thus, askers judge the answerer's credibility based on the number of credentials they have rather than details in their self-introduction. This indicates that askers tend to emphasise objective indicators such as credentials. This aligns with the concept of 'expertise signals' in social settings. People often rely on external cues to assess others' credibility, particularly in situations where direct experience or knowledge is limited. On social Q&A platforms, credentials serve as visible markers of expertise, providing a quick and easy way for askers to gauge the trustworthiness of potential answerers.

Third, regarding the scope of platform activity based on the best answerers' posting behaviours, the predictive power of the number of answer categories for answerer performance was

verified (H3-1). The results showed that those who answered questions in more categories performed better than those who answered questions in fewer categories. The positive relationship between the number of answer categories and answerer performance highlights the value of versatility on social Q&A platforms. Contributors who demonstrate proficiency in multiple areas are more likely to provide valuable insights, gain recognition, and establish themselves as trusted sources of information within the community. This can be explained by the heightened performance of answerers engaged across diverse categories, supporting H3-2, which posits that enhanced performance is associated with users who actively participate as both askers and answerers. This dual engagement allows contributors to gain a deeper understanding of community needs, build stronger connections, identify emerging trends, and contribute to knowledge-sharing, ultimately leading to enhanced performance and a more valuable platform for users.

Finally, regarding social endorsement, the degree to which lurkers responded positively to answers did not have a significant effect (H4). As an information-seeker, a lurker's response may be important on social Q&A platforms for engaging a high number of lurkers; however, it does not have a significant effect on the answer an asker selects as the best. Askers on social Q&A platforms prioritise their own assessment of answer quality and may not be swayed solely by lurker endorsements. This highlights the importance of answerers consistently providing high-quality, relevant, and informative responses to establish themselves as trusted sources of information within a community.

Implications. Our study examined the characteristics of social Q&A platforms, focusing on Naver Knowledge-iN, a representative social Q&A platform in Korea. We expect the results of this study to provide practical implications for social Q&A platforms, not only in Korea but also in other countries, for identifying top contributors who are likely to produce quality content. The best answerers on social Q&A platforms can produce quality answers as content providers, which is a crucial capability that social Q&A platforms must manage to continue providing services. Accordingly, understanding the distinguishing traits of exemplary contributors, or 'best answerers', offers universal insights that can be applied by social Q&A platforms worldwide. In addition, the best answerers have the potential to become entrepreneurs based on their expertise, reputation, and content-creation skills; thus, social Q&A platforms should manage them well. Social Q&A platforms must keep their best users with entrepreneurial skills active on their platforms and not on others. To achieve this, social Q&A platforms must systematise how their best users can consistently receive tangible and intangible rewards for the content they create. Based on our findings, social Q&A platforms worldwide can proactively manage users who have the

potential to become the best answerers by screening them in advance. For example, to encourage detailed and informative answers, a platform can implement features that promote longer and more comprehensive responses, such as word count indicators or badges for detailed answers. Furthermore, to highlight answer–question similarity, the platform can develop algorithms that identify answers that closely align with the original question's intent and context and prioritise those answers in search results or recommendations. Platforms can also encourage contributors to expand their areas of expertise by providing incentives for answering questions in multiple categories. This can diversify the expert pool and increase the platform's coverage of diverse topics.

Our results provide actionable recommendations for social Q&A platforms seeking to cultivate supportive and engaging community environments. By implementing features that encourage detailed and informative answers and prioritising answer–question similarity, platforms can foster an atmosphere conducive to quality contributions. These strategies apply not only to Korean social Q&A platforms but also to foreign social Q&A platforms aiming to enhance user satisfaction and platform relevance.

Our results can also be used when social Q&A platforms worldwide want to use other tools (e.g., Gen AI chatbots) rather than humans to generate content. Despite the variety of tools (e.g., search engines, social Q&A platforms, and Gen AI chatbots) available to search for answers online, we found that people often use the same criteria to determine a quality answer regardless of the search tool. Our results show that when people seek answers on social Q&A platforms, they place greater importance on knowing that the answers are sincere and empathetic than factors such as the immediacy of the answer or specificity of an answerer's self-introduction. These qualities, which are often overlooked in other information-seeking tools, can significantly affect the quality and value of answers. Since for the evolution of social Q&A platforms alongside AI-powered tools is a global trend, recognising and nurturing these human qualities will be essential for their continued success and relevance.

Limitations. This study suggests several avenues for future research. First, the cross-sectional nature of the data introduced certain constraints. A more nuanced understanding could be achieved by utilising panel data, allowing for an in-depth examination of dynamics such as whether the asker's selection of the best answer is influenced by the behaviours of distinguished contributors or lurkers.

Second, we did not perform a comparative analysis of the characteristics of the top and bottom answerers. Several previous investigations have explored distinctions between active users and lurkers using survey data (e.g., Zheng et al., 2021); however, studies that have systematically analysed the actual behavioural data of social Q&A platform users are scarce. While the limited number of posts from the bottom answerers presents a challenge, augmenting our dataset with additional user information, including demographics and other platform activities, could provide richer insights into the factors that influence answer quality.

Third, to unravel the qualities that contribute to the performance of top answerers, we adopted a comprehensive perspective to scrutinise the various dimensions of best-answerer participation. However, despite our efforts to incorporate additional variables, including the depth of the response, readability, objectivity, keyword density, and topic similarity, as recommended by Fu and Oh (2019) and Mousavi et al. (2020), these could not be incorporated because of the limited size of the dataset. Although we attempted to include these variables, we

failed to obtain significant results. This limitation is attributed to inherent constraints in the dataset, intricate relationships, and potential interactions among the variables. Accordingly, future studies would require access to larger datasets or the application of strategic sampling techniques to bolster the statistical power, enabling a more comprehensive exploration of the multifaceted nature of responses from top answerers.

Future studies should also consider alternative variables to explore the scope of answerers' platform activities in more depth. Utilising metrics, such as the Gini coefficient or entropy, could offer a viable approach for quantifying diversity across domain categories. This study could not use Gini or entropy metrics because of the dynamic nature of Naver Knowledge-iN's answering fields. These measures may provide a more nuanced and comprehensive understanding of the various activities on a platform, thereby contributing to a more sophisticated analysis of user engagement and expertise across domains. By using these measures, future studies may quantify the diversity of domain categories more effectively. Although our study emphasises user engagement, we acknowledge the potential value of incorporating alternative metrics and encourage future researchers to explore these avenues for a more nuanced understanding of the varied characteristics that influence response quality.

Furthermore, future studies should explore alternative approaches for categorising answer fields to better capture the nuances associated with different certifications. The current problems in classifying answer fields, as highlighted by the dynamic nature of Naver Knowledge-iN's platform and the diverse certifications users may hold, necessitate a more sophisticated analysis. Researchers could investigate alternative methods that consider not only the number of certifications but also the specific domains and expertise levels associated with each certification. This approach could provide a more nuanced understanding of how user qualifications and certifications relate to their answering behaviour, thereby contributing to a more comprehensive interpretation of response quality.

Additionally, our analysis was focused on data from Naver Knowledge-iN, a prominent social Q&A platform in Korea. While the results yielded statistical significance, the scope of the study was confined to a single country and exclusively focused on social Q&A platforms operating in Korea. Future research should consider extending the analysis to include data from social Q&A platforms in diverse countries. This approach would enable cross-cultural and cross-linguistic comparisons, allowing the variations attributed to cultural and linguistic factors to be explored.

In summary, addressing these aspects in future research will contribute to a more comprehensive understanding of the dynamics and factors influencing user engagement and answer quality on social Q&A platforms.

Data availability

All data generated or analyzed during this study were uploaded in the open repository. They can be found at: https://osf.io/ua896/?view_only=80e8e61c44cc4eb5b4b4614175034fb6.

References

Adamic LA, Zhang J, Bakshy E, Ackerman, MS (2008) Knowledge sharing and yahoo answers: everyone knows something. In Proc. 17th international conference on World Wide Web, ACM Press, 665. https://doi.org/10.1145/1367497.1367587

Adamopoulou E, Moussiades L (2020) Chatbots: History, technology, and applications. Mach Learn Appl 2:100006. https://doi.org/10.1016/j.mlwa.2020.100006

Bi CQ, Liu L (2022) From fame to fortune: Influencer entrepreneurship in a digital age. In Proceedings of PACIS 2022, 5–9 July 2022, Taipei-Sydney (p. 1358). https://aisel.aisnet.org/pacis2022/208/

Cenamor J, Parida V, Wincent J (2019) How entrepreneurial SMEs compete through digital platforms: the roles of digital platform capability, network capability and ambidexterity. J Bus Res 100:196–206. https://doi.org/10.1016/j.jbusres.2019.03.035

Chae (2003) If you use Naver, 'Knowledge-iN', Digital times (in Korean), https://n.news.naver.com/mnews/article/029/0000044359?sid=103, Accessed 25 June 2023

Cho (2022) 20th anniversary of NAVER Knowledge-iN, accumulating 800 million questions and answers, Tech World News (in Korean), https://www.epnc.co.kr/news/articleView.html?idxno=228285, Accessed 25 June 2023

Cohen J, Cohen P, West SG, Aiken LS (2003) Applied multiple regression/ correlation analysis for the behavioral sciences (3rd ed.). Mahwah, NJ: LEA. https://doi.org/10.4324/9780203774441

Cutolo D, Kenney M (2021) Platform-dependent entrepreneurs: power asymmetries, risks, and strategies in the platform economy. Acad Manag Perspect 35(4):584–605. https://doi.org/10.5465/amp.2019.0103

De Angelis L, Baglivo F, Arzilli G, Privitera GP, Ferragina P, Tozzi AE, Rizzo C (2023) ChatGPT and the rise of large language models: the new AI-driven infodemic threat in public health. Front Public Health 11:1166120. https://doi.org/10.3389/fpubh.2023.1166120

Dong L, Zhang J, Huang L, Liu Y (2021) Social influence on endorsement in social Q&A community: moderating effects of temporal and spatial factors. Int J Inf Manag 61:102396. https://doi.org/10.1016/j.ijinfomgt.2021.102396

Falvo MJ, Earhart GM (2009) Six-minute walk distance in persons with Parkinson disease: a hierarchical regression model. Arch Phys Med Rehabil 90(6):1004–1008. https://doi.org/10.1016/j.apmr.2008.12.018

Fang C, Zhang J (2019) Users' continued participation behavior in social Q&A communities: a motivation perspective. Comput Hum Behav 92:87–109. https://doi.org/10.1016/j.chb.2018.10.036

Fu H, Oh S (2019) Quality assessment of answers with user-identified criteria and data-driven features in social Q&A. Inf Process Manag 56(1):14–28. https://doi.org/10.1016/j.ipm.2018.08.007

Gazan R (2010) Microcollaborations in a social Q&A community. Inf Process Manag 46(6):693–702. https://doi.org/10.1016/j.ipm.2009.10.007

Ginsca AL, Popescu A (2013) User profiling for answer quality assessment in Q&A communities. In Proc. 2013 workshop on Data-driven user behavioral modelling and mining from social media (pp. 25–28). https://doi.org/10.1145/2513577.2513579

Koubaa, A., Qureshi, B., Ammar, A., Khan, Z., Boulila, W., & Ghouti, L. (2023). Humans are still better thanChatGPT: Case of the IEEEXtreme competition. Heliyon, 9(11). https://doi.org/10.1016/j.heliyon.2023.e21624

James, W (2007). The principles of psychology (Vol. 1). Cosimo, Inc, (pp.10113–0416). New york, NY

Jin J, Li Y, Zhong X, Zhai L (2015) Why users contribute knowledge to online communities: an empirical study of an online social Q&A community. Inf Manag 52(7):840–849. https://doi.org/10.1016/j.im.2015.07.005

Jin J, Yan X, Li Y, Li Y (2016) How users adopt healthcare information: an empirical study of an online Q&A community. Int J Med Inform 86:91–103. https://doi.org/10.1016/j.ijmedinf.2015.11.002

Kang H, Yang J (2018) Selection of the Optimal Morphological Analyzer for a Korean Word2vec Model. In Proc. Korea Information Processing Society Conference (pp. 376–379). Korea Information Processing Society. https://doi.org/10.3745/PKIPS.y2018m10a.376

Kim S (2010) Answerers' strategies to provide credible information in question answering community. J Korean Soc Inf Manag 27(2):21–35. https://doi.org/10.3743/KOSIM.2010.27.2.021

Kim S, Oh S (2009) Users' relevance criteria for evaluating answers in a social Q&A site. J Am Soc Inf Sci Technol 60(4):716–727. https://doi.org/10.1002/asi.21026

Kim Y, Kim JH, Lee JM, Jang MJ, Yum YJ, Kim S, Song S (2022) A pre-trained BERT for Korean medical natural language processing. Sci Rep 12(1):13847. https://doi.org/10.1038/s41598-023-36519-0

Liang Y (2017) Knowledge sharing in online discussion threads: what predicts the ratings? In Proc. 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, 146–154. https://doi.org/10.1145/2998181.2998217

Liu Z, Zhao YC, Chen S, Song S, Hansen P, Zhu Q (2021) Exploring askers' switching from free to paid social Q&A services: a perspective on the push-pull-mooring framework Inf Process Manag 58(1):102396. https://doi.org/10.1016/j.ipm.2020.102396

Mousavi R, Raghu TS, Frey K (2020) Harnessing artificial intelligence to improve the quality of answers in online question-answering health forums. J Manag Inf Syst 37(4):1073–1098. https://doi.org/10.1080/07421222.2020.1831775

Nam KK, Ackerman MS, Adamic LA (2009) Questions in, knowledge in?: A study of naver's question answering community. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 779–788. https://doi.org/10.1145/1518701.1518821

Nambisan S, Baron RA (2021) On the costs of digital entrepreneurship: role conflict, stress, and venture performance in digital platform-based ecosystems. J Bus Res 125:520–532. https://doi.org/10.1016/j.jbusres.2019.06.037

Nicolescu L, Tudorache MT (2022) Human-computer interaction in customer service: the experience with ai chatbots—a systematic literature review. Electronics 11(10):1579. https://doi.org/10.3390/electronics11101579

Peng CH, Yin D, Zhang H (2020) More than words in medical question-and-answer sites: a content-context congruence perspective. Inf Syst Res 31(3):913–928. https://doi.org/10.1287/isre.2020.0923

Petrocelli JV (2003) Hierarchical multiple regression in counseling research: common problems and possible remedies. Meas. Eval Counseling Dev 36(1):9–22. https://doi.org/10.1080/07481756.2003.12069076

Portuguez Castro M, Gómez Zermeño MG (2021) Being an entrepreneur post-COVID-19 – resilience in times of crisis: a systematic literature review. J. Entrep Emerg Econ 13(4):721–746. https://doi.org/10.1108/JEEE-07-2020-0246

Qi T, Wang T, Yan J (2021) The spillover effects of different monetary incentive levels on health experts' free knowledge contribution behavior. Internet Res 31(6):2143–2166. https://doi.org/10.1108/INTR-08-2020-0445

Ray PP (2023) ChatGPT: a comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. Internet Things Cyber-Phys Syst 3:121–154. https://doi.org/10.1016/j.iotcps.2023.04.003

Rudolph J, Ismail FMM, Popenici S (2024) Higher education's generative artificial intelligence paradox: the meaning of chatbot mania. J Univ Teaching Learn Pract 21(06). https://doi.org/10.53761/54fs5e77

Shah C, Pomerantz J (2010) Evaluating and predicting answer quality in community QA. In Proc. 33rd international ACM SIGIR conference on Research and development in information retrieval (pp. 411–418). https://doi.org/10.1145/1835449.1835518

Shen H, Wang G (2017) Can dynamic knowledge-sharing activities be mirrored from the static online social network in yahoo! answers and how to improve its quality of service? IEEE Trans Syst Man Cybern Syst 47(12):3363–3376. https://doi.org/10.1109/TSMC.2016.2580606

Shen H, Ze L, Jinwei L, Joseph EG (2014) Knowledge sharing in the online social network of Yahoo! answers and its implications. IEEE Transac Comput 1–1. https://doi.org/10.1109/TC.2014.2322598

Wang X, Lu J, Ow TT, Feng Y, Liu L (2021) Understanding the emotional and informational influence on customer knowledge contribution through quantitative content analysis. Inf Manag 58(2):103426. https://doi.org/10.1016/j.im.2020.103426

Xie X, Lv J (2018) Female technology entrepreneurs: Resource shortages and reputation challenges – a view of institutional support. Int Entrep Manag J 14(2):379–403. https://doi.org/10.1007/s11365-017-0450-y

Yan Z, Zhou J (2015) Optimal answerer ranking for new questions in community question answering. Inf Process Manag 51(1):163–178. https://doi.org/10.1016/j.ipm.2014.07.009

Ye H, Yang X, Wang X, Stratopoulos TC (2021) Monetization of digital content: drivers of revenue On Q&A Platforms J Manag Inf Syst 38(2):457–483. https://doi.org/10.1080/07421222.2021.1912934

Zheng X, Shi X, Yang F (2021) Media system dependency and user attachment in social Q&A communities: do active users and lurkers differ? Inf Technol People 34(7):1863–1889. https://doi.org/10.1108/ITP-12-2019-0628

Zhou ZM, Lan M, Niu ZY, Lu Y (2012) Exploiting user profile information for answer ranking in cqa. In Proc. 21st international conference on World Wide Web (pp. 767–774). https://doi.org/10.1145/2187980.2188199

Acknowledgements

Author contributions
Corresponding author contributes to conceptualisation, review and editing, supervision, and funding acquisition. The first author contributes to methodology, formal analysis, and writing—original draft preparation.

Competing interests
The authors declare no competing interests.

Ethical Approval

Ethical approval was not required as the study did not involve human participants.

Informed Consent

Informed consent was not necessary as this research did not include human subjects.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1057/s41599-024-03239-3.

Correspondence and requests for materials should be addressed to Seongcheol Kim.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.