What Makes It Findable? An Exploration on User Search Behavior and the Findability of Technical Documentation

Zhijun Gao
Peking University
University of Twente
Beijing, China
gaozhijun@pku.edu.cn

Yuxin Gao
Peking University
Beijing, China
gaoyuxin@pku.edu.cn

Jingsong Yu*
Peking University
Beijing, China
yjs@ss.pku.edu.cn

Abstract - This paper was presented at the Invited Panel session "Technical Communication in China". Findability is one of the most important qualitative factors of websites. With the rapid growth in navigation complexity and in number of technical documentations in help centers, whether users can easily locate the target document could directly determine the information retrieval task outcome. Providing users with a fine guide to target documents and then helping them find solutions to their problems is the most important function of a help center. Investigation on user search behavior data and perceived findability of documentation has to be done in order to further apply website log data to predicting user subjective assessment. In this paper we analyze the correlation between subjective document findability, subjective task complexity, and user search behavior. We found several search behavior metrics which significantly correlate with the two subjective measures above.

Index Terms - Help center evaluation, subjective document findability, subjective task complexity, user search behavior.

Introduction

A help center contains a collection of massive technical documents. Many companies use online help centers to provide convenient customer service, which facilitates users to solve problems and also helps service providers save labor costs. As the service content increases, technical documentation in a help center piles up at an alarming rate, leading to a significant increase in the structural complexity of the help center. Although the three-click rule is an unofficial website design principle, it mirrors the fact that users are always impatient in finding information and they are likely to leave within a small number of clicks. Service providers are aware of that and they have to find a way to evaluate the findability of information in their help centers. Do my customers find answers using search? Zendesk uses searches by page, search refinements, search exits by search term to answer this question [1]. Kayako [2] uses popular searches, failed searches and most popular articles for help center search and article analytics. Google Analytics [3] provides search metrics such as sessions with search, percentage of sessions that use internal search, search refinement, time after search, and search depth.

Aside from these methods to evaluate the help center search efficiency, the most intuitive and simple way is to collect the user's subjective feedback, which is also the hardest to obtain. Currently widely used help center search efficiency indicators are relatively independent, more similar to those of general websites, lacking a framework for holistic evaluation of help center search efficiency. And there are still many obstacles to quantitatively predict the subjective assessment of documents in help centers. In order to conquer this problem, we explored the correlation of subjective document findability, subjective task complexity and user search behavior as the foundation for further mining into website log data and then predicting user subjective assessment using behavior data.

This paper is structured as follows. Section 2 presents our research objectives. Section 3 reviews the literature on user search behavior. Section 4 presents our experiment design. In Section 5, results of the experiment are discussed and Section 6 presents a conclusion.

Research Objectives

Peter Morville [5] defines findability as a) the quality of being locatable or navigable, b) the degree to which a particular object is easy to discover or locate, c) the degree to which a system or environment supports navigation and retrieval. Findability is influenced by many factors, such as

2158-1002/20/$31.00 ©2020 IEEE
DOI 10.1109/ProComm48883.2020.00031

enormous growth of websites and digital documents, low level of information literacy of internet users, not adhering to the standards of the Word Wide Web Consortium and to the recommendations of information architects. Reference [6] When considering only the structure of a help center, the objective findability of certain document can be calculated using graph-based method as [7]. However, as the number of documents in a help center could be enormous, and users always come with certain questions in mind, and may lose patience quickly, objective findability of documents is not quite practical. Users with questions seeking information in help centers are like users performing searching tasks, thus we took subjective task complexity into account. In the study reported in this paper, our interest is in examining the relations between subjective document findability, post-task assessed complexity and users' information search behavior in a help center. For this study, we have posed two research questions:

RQ1: Does a user's post-task assessed complexity correlate with the user's subjective findability of the target document?

RQ2: Does user's perceived findability of a technical document in a help center correlate with user search behavior?

The study was conducted with the following goals in mind:

1) To examine relationships between subjective findability and subjective task complexity.
2) To examine relationship between subjective findability, subjective task complexity and user search behavior.
3) To examine which measures of the user search behavior are more important in predicting subjective document findability or subjective task complexity.

RELATED WORK

The majority of earlier studies concerning online searching can be categorized as focusing on either traditional information retrieval (IR) systems or online public access catalogue (OPAC) systems [8]. And user studies can be viewed as a subset within the larger area of IR system evaluation, which typically focuses on measuring the recall and precision of the system [9]. In this kind of evaluation, one takes a known document collection and executed a set of queries using a particular IR system. Based on the number of relevant and non-relevant documents retrieved, one determines recall and precision. The whole process is very systematic and user behavior has little to do with the outcome.

However, user behavior elements cannot be ignored in online searching studies. Once considering the user's actions, the IR system evaluation metrics are no longer good enough. A common method of studying user interactions with IR systems is using transaction logs. Transaction log analysis (TLA) uses transaction logs to discern attributes of the search process, such as user's actions, the interaction between user and the system, search strategy, the delivery of results, and user's evaluation results. TLA can provide necessary data, but it is limited [10], as TLA can only deal with searcher's actions. In this case, it lacks user profiles and context information, and thus TLA cannot be the only source of data used in search behavior studies. To complement information on users, several studies have used client-side logs to study web seeking behaviors. Since client-side logs can be collected by using a custom extension or customized tool, some unique tools were developed [11][12].

There are also many researches about understanding information seeking behavior within several kinds of contexts [13], and context remains an effable concept in information science [14]. Context is generally recognized to affect web search behavior in a variety of ways [15]. Kelly [16] discusses descriptions of the relationships between situation and context. She defines context as 8 variables from bigger variables (e.g., task, topic, usefulness). It is reported that 5 variables were related to the tasks and topics of participants' web searching behavior, including endurance, frequency, stage, persistence, and familiarity. Furthermore, there are some particularly useful studies for conceptualizing context. Vakkari [17] identifies three types of tasks used in information seeking studies and defines task as a "piece of activity to be done in order to achieve a goal".

Task complexity is recognized to be one of the most important factors that affect information-seeking strategies. Campbell [18] describes three general approaches to complexity: a) psychological (subjective), b) person-task interaction, and c) objective (defined by task characteristics). He advocated an objective understanding of task complexity and proposed four task aspects as the factors contributing to task complexity: 1) multiple possible paths to the outcome, 2) multiple outcomes, 3) conflicting interdependence among paths, 4) uncertainty linkages between paths and outcomes.

Along this line, reference [19] studies on task complexity and information seeking activities in real-life work tasks. Reference [20] finds that information activities are systematically connected to task complexity and structure of the problem at hand. [21] assesses structural complexity of website menus by path depth and menu options diversity. They found that for low-complexity tasks, menu diversity has virtually no impact upon navigational behavior. The same group assessed structural complexity of website menus by path depth and menu options diversity. They finds that for low-complexity tasks, menu diversity has virtually no impact upon navigational behavior. In [22], their study shows that the relationships between the operational measures (such as time spent, number of page visited, number of changes in search strategies) and the subjectively perceived task difficulty is of no fundamental nature and does not depend on particular search mechanisms. Reference [23] manipulates task complexity along path length and path relevance. Reference [24] discovers that when using search scope, the time lapse between searches has significant correlation with search results.

EXPERIMENT DESIGN

We conducted an information-seeking study through remote screen sharing. The experimental data is recorded using screen recording.

I. User Task

The authors of [22] claim that objective complexity is first and foremost defined by the length of the path that leads to the target information. And in [23], the authors use path relevance as one of the factors of objective complexity, which can be understood as the semantic parameters of the path. However, in order to improve search efficiency and minimize the number of clicks when searching, Alibaba Cloud's navigation of a certain product is up to four layers, which leaves us with the option to use semantic relevance as the measure of objective task complexity. According to the degree of difficulty of the information retrieval through site search, our user tasks are divided into simple and difficult groups. Altogether there are three simple tasks coming from high-traffic documents in the help center, and two difficult tasks collected from Alibaba Cloud feedback platform, which was intended to be close to the possible questions a user might actually encounter and in need of the help center to find its solution. Each participant was asked to perform one simple task and one difficult task on the help center of Alibaba Cloud.

II. Participants

Since the help center contains mostly cloud service documents, whose users are possibly programing-capable, we recruited twenty-six postgraduate students (7 females and 19 males) participants from the first-year students at School of Software and Microelectronics, Peking University. Four of the participants claim that they have no programming experience, but all of them have taken programming-related courses.

III. Procedure

The study was conducted through remote screen sharing and recording. After filling out a pre-task questionnaire, the participants were asked to find the target information according to their tasks. Then the participants would fill out a post-task questionnaire to rate the target document findability and perceived task complexity on a scale from 1(easy) to 7(difficult). To be more accurate, the participant's search behavior data would then be carefully recorded manually according to the video. Then analysis was made using SPSS Statistics.

IV. Measures

Dependent variable used in the experiment was the participant's subjective findability of the target document and post task subjective complexity of the task collected in the post task questionnaire. Divided into four groups, the independent variables are listed in Table 1. General task metrics describe the basic information of the task. As the task consists of finding the target document in the help center and retrieving the corresponding information in it, we use find task metrics to describe the process of finding the target document to avoid using the word 'search' and causing confusion. Furthermore, more detailed indicators are grouped as search metrics and navigation metrics according to the information-seeking method. The detailed description of independent variables is in Table 1 and Table 2.

TABLE 1. VARIABLE DESCRIPTIONS.

Variable Name	**Description**
Task time	Time spent on the task
Task success	Whether the task is success
Search complexity	The cumulative number of search functions used in a task, e.g., if a participant first uses site search then navigation and then again site search, the search complexity of this task will be 3.
Search type	Search function combinations of a task with specific order, e.g., site search then navigation and in reverse order, are two different kinds of search types.
Find task time	Time spent searching for the target document.
Find click	Number of clicks made to find the target document.
Read documents	Number of total documents read.
Read target time	Time spent reading the target document.
Search numbers	Number of searches.
Search click	number of clicks made while using the search function
Time after search	Time spent browsing search results.
Search result pages	Number of result pages read after searching.
Avg time after search	Time spent browsing search results divided by the number of result pages read after searching.

Variable Name	Description
Search results	Number of search results browsed.
Navi click	Number of clicks made while using navigation.
Navi time	Time spent using navigation to find certain documents.

TABLE 2. VARIABLE GROUPS.

General task metrics	Find task metrics	Search metrics	Navigation metrics
Task time	Find task time	Search numbers	Navi click
Task success	Find click	Search click	Navi time
Search complexity	Read documents	Time after search	
Search combination type	Read target time	Search result pages	
		Avg time after search	
		Avg search results	

EXPERIMENT RESULTS

Table 3 shows correlations between findability, subjective task complexity, and general task metrics. Subjective document findability correlates positively with subjective task complexity, which is easy to understand as the more difficult a document is to find, the higher the subjective task complexity is. Findability correlates with search complexity and search type while the correlations of those and subjective task complexity are not significant.

TABLE 3. CORRELATIONS BETWEEN FINDABILITY, SUBJECTIVE TASK COMPLEXITY, AND GENERAL TASK METRICS

	Findability	Subjective task complexity
Findability	1	-.899**
Subjective task complexity	-.899**	1
Task time	-.467**	.457**
Task success	.773**	-.685**
Search complexity	-.323*	.283
Search type	-.315*	.269

NOTE: LEVELS OF SIGNIFICANCE (PEARSON CORRELATION): $^{*}P < 0.05$; $^{**}P < 0.01$

TABLE 4. CORRELATIONS BETWEEN FINDABILITY, SUBJECTIVE TASK COMPLEXITY, AND FIND TASK METRICS

	Findability	Subjective task complexity
Findability	1	-.899**
Subjective task complexity	-.899**	1
Find task time	-.535**	.568**
Find click	-.389**	.268
Read documents	-.083	-.070
Read target time	.401**	-.539**

NOTE: LEVELS OF SIGNIFICANCE (PEARSON CORRELATION): $^{*}P < 0.05$; $^{**}P < 0.01$

Table 4 presents correlations between the two dependent variables and find task metrics. Find time is more accurate in the correlations with the two dependent variables than task time because it excludes the time spent on typing in the search box, browsing the home page, and reading documents. Find click only significantly correlates with findability, thus it can be a great indicator for predicting subjective document findability. The number of documents read in the task has no significant correlation with both dependent variables.

Table 5 presents correlations between the two dependent variables and search metrics. Time spent after search, the number of search results browsed and the number of browsed search results significantly correlate with both dependent variables. Average time spent on each result page correlates only with subjective task complexity. Search numbers, search clicks have no significant correlation with both dependent variables.

TABLE 5. CORRELATIONS BETWEEN FINDABILITY, SUBJECTIVE TASK COMPLEXITY, AND SEARCH METRICS

	Findability	Subjective task complexity
Findability	1	-.899**
Subjective task complexity	-.899**	1
Search numbers	.123	.043
Search click	-.060	.030
Time after search	-.368*	.402**
Search result pages	-.320*	.337*
Avg time after search	-.224	.311*
Search results	-.462**	.489**

NOTE: LEVELS OF SIGNIFICANCE (PEARSON CORRELATION): $^{*}P < 0.05$; $^{**}P < 0.01$.

Table 6 shows correlations between the two dependent variables and navigation metrics. Only navigation time has significant correlation with findability. We expected navigation clicks could indicate possible navigation failure, but it has no significant correlation with both dependent variables.

TABLE 6. CORRELATIONS BETWEEN FINDABILITY, SUBJECTIVE TASK COMPLEXITY, AND NAVIGATION METRICS

	Findability	**Subjective task complexity**
Findability	1	-.899**
Subjective task complexity	-.899**	1
Navi click	-.267	.180
Navi time	-.303*	.109

NOTE: LEVELS OF SIGNIFICANCE (PEARSON CORRELATION): *P < 0.05; ** P < 0.01

According to Figure 1 and Figure 2, [full task time/ find task time], when participants consider the documents to be relatively easy to find–having a findability score of no less than 4–both of their task time and time needed for locating the documents show a clear sign of decrease, compared to time required for harder tasks, which conforms with our intuition that easier tasks are finished more quickly. Notice that an abnormality appears on the task with findability score of 4. This participant spends fourth times the maximum finding time of others but does not consider it a difficult task. Re-watching the video clip of his experiment, we find that there is a typo mistake when he types the product name in search box, and therefore, is misled by the search results. When the participant is aware of the mistake, he finds the target product document quickly and rates it a 4.

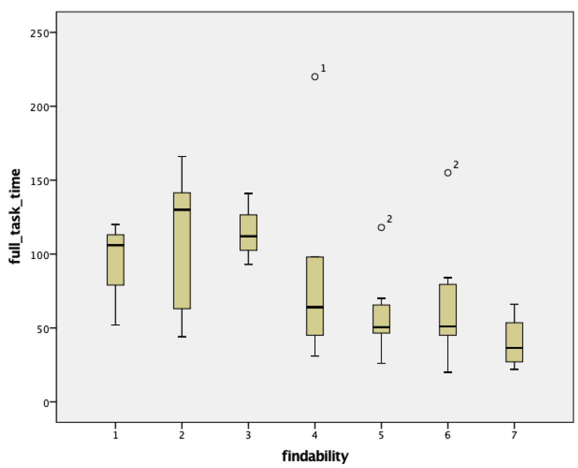

FIGURE 1. RELATION BETWEEN FULL TASK TIME AND FINDABILITY

Different patterns would be more obvious when we employ the k-means clustering, as shown in Table 7. This algorithm discovers data points that are closely to each other and group them into the same cluster. Results shows that, when we want to split our data into two clusters, all data points are centered around findability scores of 5 and 3, respectively. This indicates that, participants who rate the findability 5 behave quite differently than those who rate the findability 3; and people in the same groups perform similarly during the tasks.

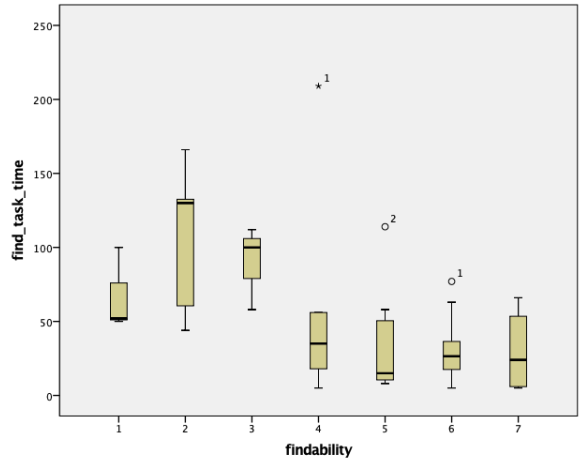

FIGURE 2. RELATION BETWEEN FIND TASK TIME AND FINDABILITY

TABLE 7. K-MEANS CLUSTERING ON FINDABILITY AND FIND TASK TIME

	Cluster	
	1	**2**
Findability	5	3
Find task time	33	133

LEVELS OF SIGNIFICANCE (PEARSON CORRELATION): *P < 0.05; ** P < 0.01

Figure 3 depicts the number of search results participants check after they have used the search function. Participants who consider their tasks as relatively easy only have to check a few search results to be able to find the information. The average number of results checked is about 5 for them. Participants who think the tasks are tough (findability score of 2 and 3) would have to look through a few dozens of search results on average. Interestingly, participants who think the task are very hard only go through a few search results, and decide a poor findability. We find that those participants simply give up on the task, after skimming through a few search results. This indicates some users with a few clicks could still lose patience, which calls for the search result optimization for the help center. The box plots in Figure 4 shows that, when participants are able to locate the information in less than 50 seconds, they tend to think the information is easy to find.

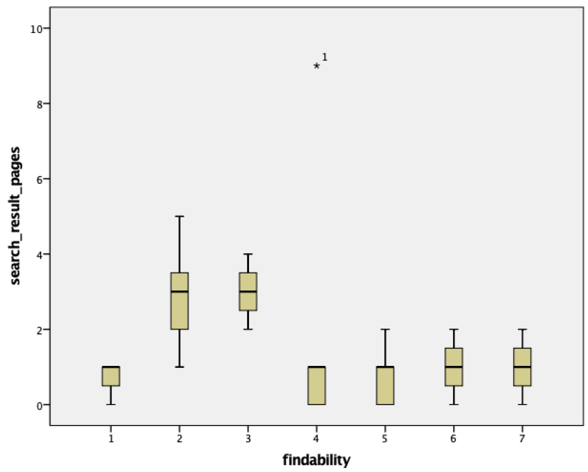

FIGURE 3. RELATION BETWEEN SEARCH RESULT PAGES AND FINDABILITY

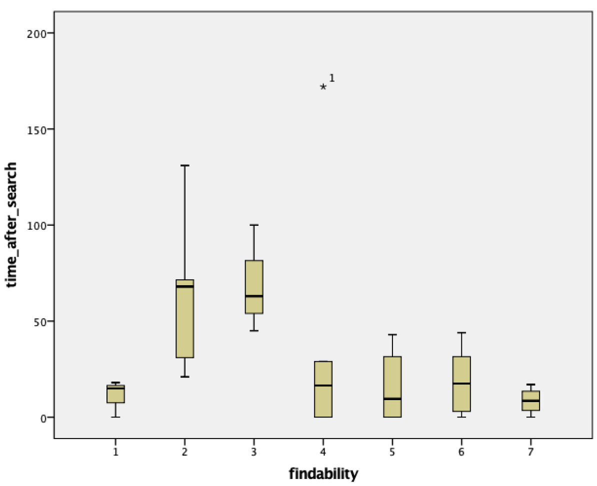

FIGURE 4. RELATION BETWEEN TIME AFTER SEARCH AND FINDABILITY

CONCLUSION AND FUTURE WORK

In our work, we find that participants' subjective ideas on document findability correlate with their behavioral data to a large extent, as stated in section 5. In such cases, user behavioral data could be used to infer the perceived findability scores of online documentations. This finding provides new perspectives on redesigning online help centers. To put it differently, if we could build a prediction model out of large amounts of user behavioral data, we would locate some of the hard-to-find documents on our help centers, and then try to improve its design by examining user paths. Thus, in future work, we aim to expand our datasets with more well-directed experiment tasks and to identify underlying patterns as well as documents with poor findability.

REFERENCES

[1] J. Maynard, Google Analytics and Help Center - Part 2: Measuring the effectiveness of search, https://support.zendesk.com/hc/enus/articles/203664146-Google-Analytics-and-Help-Center-Part-2-Measuring-the-effectiveness-of-search, 2017, last accessed Apr 12,2020.

[2] A. Ayaz, Self-service metrics for your Help Center, https://help.kayako.com/hc/en-us/articles/360006456799-Self-servicemetrics-for-your-Help-Center, 2019, last accessed Apr 13, 2020.

[3] Google, https://support.google.com/analytics/answer/1032321?hl=en, 2019, last accessed Jan 23, 2020.

[4] T. Walker, How Metrics Help Us Measure Help Center Effectiveness, https://www.k15t.com/blog/2018/01/how-metrics-help-us-measure-helpcenter-effectiveness, 2018, last accessed Feb 20, 2020

[5] P. Morville, *Ambient Findability*. Sebastopol, CA :O´Reilly and Associates, 188 p, 2005.

[6] S. Makulov´a, The Analysis and Proposal of Recommendations for Better Findability of Library Websites in Internet.

[7] J. Pawel, Kalczynski, S. Senecal, and J. Nantel. "Predicting On-Line Task Completion with Clickstream Complexity Measures: A Graph-Based Approach," *International Journal of Electronic Commerce*, vol. 10, no. 3, pp. 121–141, 2006.

[8] B. J. Jansen, U. Pooch, "A review of Web searching studies and a framework for future re- search," *Journal of the American Society for Information Science and Technology*, vol. 52, no. 3, pp. 235–246, 2001.

[9] K. Sparck-Jones Ed., *Information Retrieval Experiments.* London: Butterworth, 1981.

[10] T. Peters, "The history & development of transaction log analysis," *Library Hi Tech*, vol. 42 no. 11, pp. 41–66, 1993.

[11] L. D. Catledge, J. E. Pitkow, "Characterizing browsing strategies in the World-Wide Web," *Computer Networks and ISDN System*, vol. 27, no. 6, pp. 1065–1073, 1995.

[12] R. B. Patrick Dubroy, A Study of Tabbed Browsing Among Mozilla Firefox Users. In: *CHI 2010: Browsing, The research on the actual condition of the internet usage in Korea. KISA, Seoul*, pp. 673–682, 2009.

[13] P. Ingwersen, K. Jarvelin, Information retrieval in context. In: Ingwersen, P., Belkin, N.J. (Eds.) *Proceedings of the SIGIR 2004 Workshop on Information Retrieval in Context*, pp. 6–9. ACM Press, New York, 2004.

[14] C. Cool, "The concept of situation in information science," *Annual Review of Information Science and Technology,* vol. 35, pp. 5–42, 2001.

[15] T. Saracevic, The stratified model of information retrieval interaction: Extension and ap- plications. In: Schwartz, C.S. (ed.) *Proceedings of the Annual Meeting of the American Society for Information Science (ASIS 1997), American Society for Information Science and Technology, MD*, 1997, pp. 3–9..

[16] D. Kelly, C. Cool, "The effects of topic familiarity on information search behavior," In: Hersh, W., Marchionini, G. (eds.) *Proceedings of the Second ACM/IEEE Joint Confe- rence on Digital Libraries (JCDL 2002),* pp. 74–75. ACM Press, New York, 2002.

[17] P. Vakkari, "Task-based information searching," *Annual Review of Information Science and Technology*, vol. 37, 413–464, 2003.

[18] D. J. Campbell, "Task complexity: A review and analysis," *Academy of Management Review*, vol. 13, pp. 40-52, 1988.

[19] K. Bystr¨om, "Information and information sources in tasks of varying complexity," *Journal of the American Society for Information Science and Technology*, vol. 53, pp. 581-591, 2002.

[20] P. Vakkari, "Task complexity, problem structure and information actions," *Integrating studies on information seeking and retrieval Information Processing and Management*, vol. 35, pp. 819-837. 1999

[21] S. S. Bodrunova, A. Yakunin, "Impact of Menu Complexity upon User Behavior and Satisfaction," in *Information Search[C]// International Conference on Human Interface the Management of Information.* Springer, Cham, 2018.

[22] J. Gwizdka and I. Spence, "What Can Searching Behavior Tell Us About the Difficulty of Information Tasks? A Study of Web Navigation," *Proceedings of the American Society for Information Science and Technology*, vol. 43, no. 1, pp. 1-22, 2006.

[23] H. Van Oostendorp, R. I. Madridb, and M. Carmen Puerta Melguizo, "The effect of menu type and task complexity on information retrieval performance," *Ergon. Open J.,* vol. 2, pp. 64–71, 2009.

[24] P. Huntington, D. Nicholas, R. Hamid, and M. Jamali, *Employing log metrics to evaluate search behavior and success: Case study BBC search engine.* New York, NY: Sage Publications, Inc. 2007.

About the Authors

Zhijun Gao is a lecturer and researcher at the Department of Digital Art and Technical Communication at Peking University (Beijing, China). He is also a PhD candidate at the University of Twente (Enschede, the Netherlands). He is also the director for China Technical Communication Alliance. His research interests include technical communication, computer-aided translation, and Global UX.

Yuxin Gao is a master student of the Department of Digital Art and Technical Communication at Peking University (Beijing, China). She is interested in technical communication and user experience.

Jingsong Yu is an associate professor of natural language processing at the Department of Digital Art and Technical Communication at Peking University (Beijing, China). He is also vice director of the department. He has founded the first master's program in computer-aided translation in mainland China.