Eliciting Implicit Information Needs in E-Commerce Search by Using the Think-Aloud Method

Kosetsu Tsukuda
National Institute of Advanced Industrial Science and Technology (AIST)
Tsukuba, Japan
k.tsukuda@aist.go.jp

Atsuki Maruta
University of Tsukuba
Tsukuba, Japan
maruta.itf@gmail.com

Makoto P. Kato
University of Tsukuba
Tsukuba, Japan
mpkato@acm.org

Hideo Joho
University of Tsukuba
Tsukuba, Japan
hideo@slis.tsukuba.ac.jp

Abstract

In this paper, we elicit implicit information needs that arise during the process of deciding which products to purchase on e-commerce (EC) sites. We designed product purchase tasks to capture implicit information needs, and we conducted a user study to collect utterance data using a think-aloud method. By analyzing the utterances of participants during the tasks, we developed a taxonomy comprising five categories where people express preferences for products and 11 categories where people want to understand products. Our taxonomy includes implicit information needs that have not been captured in existing EC-related taxonomies (*e.g.*, *Preference for Subjective Attributes* and *Understanding Product Differences*). We revealed the characteristics of each category of information need in terms of timing during the tasks: *e.g.*, the information need of *Understanding Product Range* occurred very frequently in the early stage of a task. We also revealed the occurrence frequencies for different task types: *e.g.*, the information needs of *Preference for Objective Attributes*, *Understanding Product Range*, and *Understanding Terminology* had a higher occurrence when purchasing products less frequently and at a higher cost than when purchasing products frequently at a relatively low cost. Our taxonomy could be used to further improve users' purchasing processes on EC sites.

CCS Concepts

• **Information systems** → **Information retrieval**; • **Applied computing** → **Online shopping**.

Keywords

E-commerce, information needs, taxonomy, think-aloud method

CHIIR '25, March 24–28, 2025, Melbourne, VIC, Australia

ACM ISBN 979-8-4007-1290-6/25/03
https://doi.org/10.1145/3698204.3716461

ACM Reference Format:
Kosetsu Tsukuda, Atsuki Maruta, Makoto P. Kato, and Hideo Joho. 2025. Eliciting Implicit Information Needs in E-Commerce Search by Using the Think-Aloud Method. In *2025 ACM SIGIR Conference on Human Information Interaction and Retrieval (CHIIR '25), March 24–28, 2025, Melbourne, VIC, Australia.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3698204.3716461

1 Introduction

With the advancement of natural language processing technologies, web search engines like Google and Bing can now retrieve highly accurate search results for not only keyword-based queries but also natural language queries [7]. Recently, it has also become common to generate summary responses to queries via large language models (LLMs) and to display these responses on the search engine results page (SERP) [28]. While web search's flexibility has increased, it is still common on e-commerce (EC) sites such as Amazon for users to enter keyword-based queries in a search box and view a list of products in the search results [1, 12, 25]. Hence, the capability for EC sites to accept natural language queries in the search box, thereby improving the search results and generating summary responses, could enhance the user experience and further expand the EC market.

When choosing products to purchase on EC sites, users input various queries and browse through search results or compare multiple products before making a purchase decision [3, 8, 11]. In this process, users have various information needs [6, 24] such as "I want a hat with a cute design" or "I want to know the differences between these two displays." Clarification of users' information needs during product search can provide useful insights to enable more flexible search on EC sites, such as application to retrieval-augmented generation (RAG) [10, 15]. For instance, if it becomes clear that the frequency of information needs like "I want to know the differences between products" is high, it could be predetermined in RAG to refer to each product's description. By doing so, when a query like "tell me the difference between the Dell G2724D and Dell S2721QS displays" is entered, accurate comparison results can be generated on the search results page.

There has been an increase in research on conversational recommendation systems (CRSs) for purchasing assistance [13, 16, 31], and some researchers have also constructed taxonomies from user utterance intents in dialogue systems [4, 17]. Such research assumes that the purchasing process can be completed solely using the dialogue system. CRS is considered an essential future technology, and we fully recognize the significance of CRS research. However, major platforms, including Amazon, do not yet provide such services, and users still perform traditional search on EC sites. To bridge the gap between CRS and current EC site search practices, a promising approach would be for users to continue using traditional EC sites while being able to input their information needs in natural language in the search box.

In this context, we aim to construct a new taxonomy of information needs in EC search. For this purpose, we conducted a user study with eight subjects performing four product purchasing tasks. The think-aloud method [27] was applied to extract implicit information needs that do not appear in EC-related search queries [22, 25, 26], and the subjects were asked to verbalize their information needs that arose during the search process. Our contributions can be summarized as follows:

- From the utterances during the tasks, we constructed a fine-grained taxonomy comprising two top-level information needs (*Preference* and *Understanding*) and 16 low-level information needs (such as *Preference for Objective Attributes* and *Understanding Contextual Information*).
- Through comparison with existing EC-related taxonomies, we qualitatively demonstrated that our taxonomy reveals implicit information needs that have not been captured in existing taxonomies.
- Through temporal analysis of the information needs and comparison of the needs generated by different task types, we showed the differences in the characteristics of the information needs in our taxonomy.

2 Related Work

The construction of taxonomies for information needs during web search has long been a topic of interest [2, 5, 23, 29]. In contrast, the building of taxonomies for information needs in EC is a relatively new topic: the first taxonomy based on search queries entered on EC sites was proposed in 2018 [25, 26]. Rao et al. [22] collected product-related web search queries to analyze broader information needs in product search, and they constructed a taxonomy accordingly. In the cases of both search on EC sites and web search, the information needs extracted from search queries would be only a part of the information needs that users might have in the process of purchasing products. To more deeply understand users' information needs, in this study, we designed search tasks using the think-aloud method and analyzed participants' utterances during these tasks. The information needs of CRS users are also related to this study, in that they arise during the process of choosing an item. Cai and Chen [4] constructed a taxonomy by analyzing user utterances during simulated movie recommendation dialogs, and Lyu et al. [17] extended this taxonomy through restaurant recommendation tasks. While these studies also analyzed user utterances, our study aims to elicit information needs during product search on existing EC sites, rather than on dialog systems. Section 4.2 discusses the differences between the taxonomy constructed in our study and the aforementioned EC-related taxonomies.

3 User Study Design

3.1 Product Purchase Task

To avoid bias in the product purchase task, which could compromise the constructed taxonomy's generalizability, this study introduced task diversity from the perspectives of product search intents and product categories. For the search intents, we considered Target Finding (TF) and Decision Making (DM) as proposed by Su et al. [26]. Typically, in TF, both the product's brand and category are predetermined, e.g. "NETGEAR router," whereas in DM, only the category is decided, e.g. "router." For product categories, we considered Convenience Products (CP), which are purchased frequently and at a relatively low cost, and Shopping Products (SP), which are purchased less frequently and at a higher cost. In our tasks, specifically, the CP category included "salad dressings" and "mystery novels," while the SP category included "kitchen knives" and "mobile batteries for smartphones."

As described above, our study involved $2^3 = 8$ tasks based on combinations of product search intents, product categories, and specific products. Following previous studies [9, 19, 20, 26], to make the tasks realistic, scenarios were assigned to each of the eight tasks. For example, the scenario for a task with TF, SP, and mobile batteries was "Planning a domestic trip, I want to buy an Elecom mobile battery for my smartphone to ensure it does not run out of power while I am traveling (budget: 40 USD)."[1] As in this example, each scenario also specified a budget. To minimize the effects of scenario differences between TF and DM, the scenario for DM with the SP category of mobile batteries for smartphones was simply changed from "an Elecom mobile battery for my smartphone" to "a mobile battery for my smartphone," differing only by the presence or absence of the brand name.[2] Table 1 shows the scenarios of the eight tasks used in the user study.

3.2 Task Procedure

Here, we describe the procedure for one product purchase task. Study participants first read the task scenario and then spent 30 minutes on an EC site deciding on a product to purchase. Amazon was used as the EC site. The participants could use all available features on Amazon, including keyword search, sorting features, and navigation menus to refine search results. To collect the information needs arising during the tasks, we adopted the think-aloud method [27]. To lower the threshold for speaking, the participants were encouraged to utter any thoughts they had, even those not directly related to their information needs. It is known that consumers commonly gather information through web search when choosing a product to purchase [14, 30]. For greater realism, the participants could also use Google for web search if they wanted to investigate something during a task. Accordingly, we also asked them to utter any thoughts during the web search process. Once the participants found a candidate product, they added it to their cart on Amazon. After reviewing all the candidates, they decided what product to purchase. If a decision was made before the 30-minute

[1]Because the user study was conducted in Japanese, the example scenario given here is an English translation.

[2]The brands for salad dressings, mystery novels, kitchen knives, and mobile batteries in TF were Kewpie (a Japanese food manufacturer), Keigo Higashino (a Japanese novelist), Kai Corporation (a Japanese cutlery manufacturer), and Elecom (a Japanese computer peripheral manufacturer), respectively.

Table 1: Scenarios of the eight tasks used in the user study (TF: Target Finding; DM: Decision Making; CP: Convenience Products; SP: Shopping Products).

Intent	Category	Scenario
TF	CP	I want to buy a dressing by Kewpie, different from the one I usually use for eating salads (budget: 6 USD).
DM	CP	I want to buy a dressing different from the one I usually use for eating salads (budget: 6 USD).
TF	CP	Since I will have time to relax and read a book this weekend, I want to buy a paperback mystery novel by Keigo Higashino (budget: 13 USD).
DM	CP	Since I will have time to relax and read a book this weekend, I want to buy a paperback mystery novel (budget: 13 USD).
TF	SP	The sharpness of my current kitchen knife has deteriorated, so I want to buy a Santoku knife by Kai Corporation (budget: 33 USD).
DM	SP	The sharpness of my current kitchen knife has deteriorated, so I want to buy a Santoku knife (budget: 33 USD).
TF	SP	Planning a domestic trip, I want to buy an Elecom mobile battery for my smartphone to ensure it does not run out of power while I am traveling (budget: 40 USD).
DM	SP	Planning a domestic trip, I want to buy a mobile battery for my smartphone to ensure it does not run out of power while I am traveling (budget: 40 USD).

mark, the task ended at that point. If no decision was reached within 30 minutes, the task ended without a purchase.[3]

3.3 Participants

Eight participants took part in this user study: two females and six males. All participants were students at University of Tsukuba in Japan, and they had an average age of 23.8 years (standard deviation: 1.7 years). Each had purchased at least one item on Amazon within the last three months and performed web search more than once a week on average. Each participant was assigned four tasks (TF+CP, TF+SP, DM+CP, and DM+SP)[4] following a Latin square design to control for learning effects due to task order. Thus, four participants worked on each of the eight tasks. The participants practiced on a separate task from the eight main tasks for 10 minutes before moving on to their four assigned tasks. They took a 5-minute break between each task. The participants used their own PCs and the web browsers that they regularly use. To eliminate the influence of personalized information, if a participant used Google Chrome, for example, they were instructed to use it in Incognito mode. The computer screen and the participant's utterances were recorded during each task. The participants received a 3,000 JPY (about 20 USD) Amazon gift card as compensation. This user study was conducted with the approval by the Ethics Review Committee of Institute of Library, Information and Media Science, University of Tsukuba (Approval No. 22-144).

4 Taxonomy of Information Needs

4.1 Taxonomy Development

We constructed a taxonomy of information needs through an inductive coding approach [18]. Initially, one author transcribed all of the participants' utterances from the recorded videos. This author then identified transcript segments that contained any information needs and created draft categories (codes) for the taxonomy according to these segments. Subsequently, three authors, including the one previously mentioned, all of whom are experts in information retrieval, refined the draft taxonomy through discussion, using the extracted segments to finalize the taxonomy. By using the finalized taxonomy, the author who created the draft taxonomy labeled the extracted segments with their categories. A total of 524 segments were labeled, and as in Lyu et al. [17], multiple categories could be assigned to a single segment. Considering that Lyu et al. [17] developed a taxonomy of users' CRS information needs from 360 utterance segments by 12 participants, it can be said that our taxonomy, developed from the 524 segments by eight participants, also makes an academic contribution.

Table 2 gives the constructed taxonomy and the frequency of each category. The taxonomy comprises two top-level and 16 low-level information needs. Even categories with lower frequencies would be significant given the massive daily use of EC sites; therefore, we did not aggregate such categories into an "Others" category. As indicated by "Example" in the table, entering these sentences as natural language queries into an EC site's search box can lead to improved search results and product recommendations for *Preference*, and to the presentation of generated responses for *Understanding* through RAG.[5] To enable such support, different information sources need to be referenced depending on the category, and some categories may require referring to web pages that are not on the EC site itself. The possible primary reference sources for each category are also noted in the "Source" column.

4.2 Taxonomy Characteristics

Here, we discuss our taxonomy's characteristics in comparison with those of existing EC-related taxonomies. Typically, search queries on EC sites are either product names or product categories [1, 12, 25].

[3]In this study, there was only one instance where no purchase decision was made within 30 minutes. The participant identified several candidates to purchase but ran out of time while narrowing them down to a single product to purchase.

[4]For example, TF+CP means that the search intent is TF and the product category is CP.

[5]For readability, the example sentences provided are not direct transcriptions of the participants' utterances; however, similar utterances were made by the participants during the tasks.

Table 2: Taxonomy of information needs for product search.

Category	Description / Example	Source	Frequency
Preference			**416 (63.2%)**
Preference for Objective Attributes (PrefObjAttr)	Expression of preference for objective product attributes / "I prefer a stainless steel knife."	EC (product description)	227 (34.5%)
Preference for Product (PrefProd)	Expression of preference for the product itself / "I do not want to read this mystery novel."	EC (product description)	85 (12.9%)
Preference for Subjective Attributes (PrefSbjAttr)	Expression of preference for subjective product attributes / "I want a mobile battery with a cute design."	EC (reviews)	79 (12.0%)
Preference for Social Aspects (PrefSocial)	Expression of preference for social aspects of products / "I prefer to avoid salad dressings with few reviews."	EC (reviews)	14 (2.1%)
Preference for Personal Experience (PrefPersExp)	Expression of preference tied to personal experience / "I prefer a salad dressing flavor I have never tried before."	EC (purchase history)	11 (1.7%)
Understanding			**242 (36.8%)**
Understanding Contextual Information (UndCtxInfo)	Want to know contextual information about the product / "I want to know how long it will take to read an 800-page mystery novel."	Web	59 (9.0%)
Understanding Objective Attributes (UndObjAttr)	Want to know about objective attributes of the product / "I want to know this salad dressing's shelf life."	EC (product description)	45 (6.8%)
Understanding Product Range (UndProdRng)	Want to see a list of products / "I want to see the list of Elecom mobile batteries."	EC (product description)	36 (5.5%)
Understanding Specific Products (UndSpecProd)	Want to know if there is a product that meets specific conditions / "I want to know if a pink mobile battery is available."	EC (product description)	32 (4.9%)
Understanding Product Differences (UndProdDiff)	Want to know the differences between products "I want to know the differences between these two knives."	EC (product description)	18 (2.7%)
Understanding Terminology (UndTerm)	Want to understand the meaning of terms related to the product "I want to know what 'Damascus' means for knives."	Web	15 (2.3%)
Understanding Social Aspects (UndSocial)	Want to know about the social aspects of the product / "I want to see this salad dressing's reviews."	EC (reviews)	14 (2.1%)
Understanding E-Commerce Functions (UndECFuncs)	Want to know about EC site functions / "I want to know if I can filter the results by battery capacity."	EC (navigation menu)	11 (1.7%)
Understanding Attribute Differences (UndAttrDiff)	Want to know the differences between attribute values of products / "I want to know the differences between stainless steel and ceramic knives."	Web	5 (0.8%)
Understanding Subjective Attributes (UndSbjAttr)	Want to know about subjective attributes of the product / "I want to know this mystery novel's recommended aspects."	EC (reviews)	4 (0.6%)
Understanding Personal Relevance (UndPersRel)	Want to know information tied to personal experience / "I want to know my current mobile battery's size."	EC (purchase history)	3 (0.5%)

Consequently, taxonomies representing the nature of products, such as "Minor-Item Shopping" or "Targeted Purchase," have been developed [25], but they do not clearly reveal the diverse information needs of users, as seen in Table 2. Furthermore, product-related web search often includes queries that are not related to purchasing products, such as solutions for product issues. Thus, in existing taxonomies, the categories that are relevant to purchasing information needs are limited and coarse-grained, like "Comparison" or "Informational" [22]. We subdivide these categories further into the specific categories of *UndProdDiff* and *UndAttrDiff* for "Comparison" and categories such as *UndObjAttr* and *UndTerm* for "Informational."

Taxonomies built in CRS [4, 17] contain numerous dialogue-specific categories such as "Answer" and "Acknowledgement," and they are not specialized for information needs. The desire to know about items is aggregated under the broad category of "Inquire," lacking the specificity shown in our taxonomy. In addition, the proportion of "Inquire" in the CRS-based taxonomies has been low (6.55% [4] and 9.5% [17]). In contrast, *Understanding* accounts for 36.8% in our taxonomy, which highlights the importance of supporting user understanding during product search. While CRS-based taxonomies have included categories representing "Preference," they do not distinguish between objective and subjective information. We distinguish these types of information because they have different sources, as seen in Table 2. Although *PrefObjAttr* is more prevalent, *PrefSbjAttr* is the third most frequent category, indicating the benefits of a system that can reflect these preferences. While the consideration of subjective attributes in product selection was previously noted as an important future issue [21], we contribute here by quantitatively indicating its importance.

Finally, for users deciding on a purchase, it is more frequent and important to understand surrounding information about a product, like *UndCtxInfo*, than to know about the product itself, like *UndObjAttr* or *UndProdDiff*. Typically, the information necessary to answer questions in *UndCtxInfo* is not available on an EC site and must be sourced from the web. In such cases, for example, the system might perform a web search in the background with the user's natural language query entered in the EC site's search box; then, it could use the SERP's top page as a reference in RAG.

5 Analysis

5.1 Timeline of Information Needs Occurrence

We first performed temporal analysis of information needs. Specifically, we created a list I_t^s of information needs, arranged in order of occurrence for each task t tackled by participant s. We divided I_t^s into 10 bins and counted the frequency of each category in each bin.[6] We then summed up the frequencies of each category c for each participant's task across the 10 bins.

[6]For example, if I_t^s contained 24 information needs, the first four bins would contain three needs each, while the remaining six bins would contain two needs each.

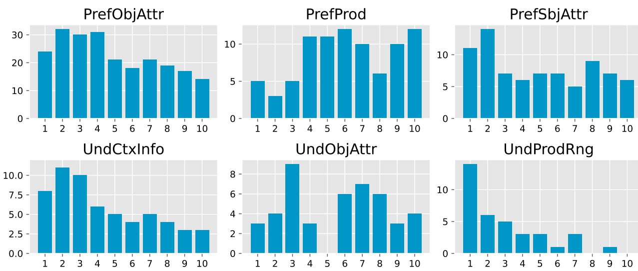

Figure 1: Timeline of information needs (x-axis: bin index; y-axis: frequency of occurrence).

The results for the top six categories in terms of occurrence frequency in Table 2 are shown in Figure 1. We can see that there are differences in the timing of information needs by category. For example, the category of *UndProdRng* occurs highly frequently at the start of a task, indicating a desire to start by getting an overview of what products are available. Therefore, for users to efficiently get such an overview, it could be useful to show diversified product search results for the *UndProdRng* information need. Additionally, the results in Figure 1 suggest potential applications for support in eliciting user information needs during product purchases. For instance, categories like *PrefSbjAttr* and *UndCtxInfo* occur highly frequently in a task's early stages. Hence, for users whose search behavior stagnates shortly after starting product search, it would be helpful to elicit information needs and advance the search phase by displaying questions on EC sites and encouraging users to enter answers in the search box. Examples of such questions include "What kind of shoes would you like to buy?" and "Is there anything you are concerned about when purchasing a digital camera?".

5.2 Influence of Task Types

Next, we analyzed the impact of product search intents and product categories on the frequencies of information needs. For product search intents, the 32 tasks performed by the eight participants included 16 instances each for TF (Target Finding) and DM (Decision Making), and we counted the frequency of information needs in each taxonomy category for both TF and DM. The difference in frequency for each category per participant was analyzed using the Wilcoxon signed-rank test (two-tailed test, significance level of 5%). The same analysis was performed for the product categories of CP (Convenience Products) and SP (Shopping Products).

The results are given in Table 3. Although no significant difference in category frequency was observed between TF and DM, DM tended to have a higher frequency in many categories. This could be because DM, unlike TF where the brand is predetermined, involves a broader range of search targets and requires more information requests to choose potential products to purchase. On the other hand, categories such as *UndSpecProd* and *UndProdDiff* showed a higher frequency of occurrence in TF. This could be due to the need to efficiently determine if there is a product meeting specific conditions among similar branded products (*UndSpecProd*) or a desire to know the differences between similar products (*UndProdDiff*). Our future plans include conducting a larger-scale user study to more accurately analyze these trends.

Table 3: Information need frequencies for different task types: TF vs. DM, and CP vs. SP. "$*$" denotes the statistical difference at $p < 0.05$ based on the two-tailed Wilcoxon signed-rank test.

Category	TF	DM	CP	SP
PrefObjAttr	100 (33.9%)	127 (35.0%)	90 (31.6%)	137* (36.7%)
PrefProd	35 (11.9%)	50 (13.8%)	48 (16.8%)	37 (9.9%)
PrefSbjAttr	28 (9.5%)	51 (14.1%)	32 (11.2%)	47 (12.6%)
PrefSocial	6 (2.0%)	8 (2.2%)	8 (2.8%)	6 (1.6%)
PrefPersExp	6 (2.0%)	5 (1.4%)	10 (3.5%)	1 (0.3%)
UndCtxInfo	32 (10.8%)	27 (7.4%)	24 (8.4%)	35 (9.4%)
UndObjAttr	19 (6.4%)	26 (7.2%)	25 (8.8%)	20 (5.4%)
UndProdRng	16 (5.4%)	20 (5.5%)	15 (5.3%)	21 (5.6%)
UndSpecProd	21 (7.1%)	11 (3.0%)	12 (4.2%)	20 (5.4%)
UndProdDiff	14 (4.7%)	4 (1.1%)	2 (0.7%)	16* (4.3%)
UndTerm	8 (2.7%)	7 (1.9%)	4 (1.4%)	11* (2.9%)
UndSocial	4 (1.4%)	10 (2.8%)	8 (2.8%)	6 (1.6%)
UndECFuncs	3 (1.0%)	8 (2.2%)	4 (1.4%)	7 (1.9%)
UndAttrDiff	2 (0.7%)	3 (0.8%)	1 (0.4%)	4 (1.1%)
UndSbjAttr	1 (0.3%)	3 (0.8%)	2 (0.7%)	2 (0.5%)
UndPersRel	0	3 (0.8%)	0	3 (0.8%)
Total	295	363	285	373*

As for CP and SP, many categories showed a higher frequency of occurrence for SP. In fact, the average frequency of information needs per participant, when all categories were combined, was significantly higher for SP than for CP. This would be because SP-type products are typically not purchased daily and are more expensive, leading the participants to be more cautious and deliberate in their purchasing decisions. In particular, *PrefObjAttr*, *UndProdDiff*, and *UndTerm* had significantly higher frequencies for SP. These results suggest the potential usefulness of applications that generate more detailed, careful response texts when users searching for SP-type products want to know the differences between products (*UndProdDiff*) or the meanings of terms (*UndTerm*).

6 Conclusion

We constructed a taxonomy of information needs that arise during product purchasing on EC sites by using the think-aloud method. Our study revealed implicit information needs in EC search that were not captured in previous EC-related taxonomies. We also clarified that there are differences in the timing and frequency of information needs depending on the type of task, and we discussed how to support EC search users according to these differences. Note that the provision of specific, detailed implementation methods for user support based on our taxonomy is beyond this paper's scope, as is the evaluation of such methods' effectiveness in user assistance. Nevertheless, based on the taxonomy developed in this paper, our important future work will include developing a system that can improve search results and generate responses according to user information needs, and investigating the system's impact on people's purchasing behaviors. Investigation of the impact of differences in users' ages and cultures on the taxonomy's categories and their proportions is another interesting direction for future work.

Acknowledgments

This work was supported in part by Project for University-Industry Cooperation Strengthening in Tsukuba, JST CREST Grant Number JPMJCR20D4, and JSPS KAKENHI Grant Number JP24K03048, Japan.

References

[1] Yuki Amemiya, Tomohiro Manabe, Sumio Fujita, and Tetsuya Sakai. 2021. How Do Users Revise Zero-Hit Product Search Queries?. In *Proceedings of the Advances in Information Retrieval - 43rd European Conference on IR Research (ECIR 2021)*. 185–192.

[2] Andrei Broder. 2002. A Taxonomy of Web Search. *SIGIR Forum* 36, 2 (2002), 3–10.

[3] Jozef Bucko, Lukáš Kakalejčík, and Martina Ferencová. 2018. Online Shopping: Factors That Affect Consumer Purchasing Behaviour. *Cogent Business & Management* 5, 1 (2018), 1–15.

[4] Wanling Cai and Li Chen. 2020. Predicting User Intents and Satisfaction with Dialogue-based Conversational Recommendations. In *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization (UMAP 2020)*. 33–42.

[5] B. Barla Cambazoglu, Leila Tavakoli, Falk Scholer, Mark Sanderson, and Bruce Croft. 2021. An Intent Taxonomy for Questions Asked in Web Search. In *Proceedings of the 2021 Conference on Human Information Interaction and Retrieval (CHIIR 2021)*. 85–94.

[6] Donald O Case and Lisa M Given. 2016. *Looking for Information: A Survey of Research on Information Seeking, Needs, and Behavior.* Emerald Group Publishing.

[7] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT 2019)*. 4171–4186.

[8] Paulo Duarte, Susana Costa e Silva, and Margarida Bernardo Ferreira. 2018. How Convenient is It? Delivering Online Shopping Convenience to Enhance Customer Satisfaction and Encourage e-WOM. *Journal of Retailing and Consumer Services* 44 (2018), 161–169.

[9] Yanjun Gao, Madhu Reddy, and Bernard J. Jansen. 2016. Shop Together, Search Together: Collaborative E-commerce. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA 2016)*. 2081–2087.

[10] Yunfan Gao, Yun Xiong, Xinyu Gao, Kangxiang Jia, Jinliu Pan, Yuxi Bi, Yi Dai, Jiawei Sun, Meng Wang, and Haofen Wang. 2024. Retrieval-Augmented Generation for Large Language Models: A Survey. arXiv:2312.10997

[11] Sergio Hernández, Pedro Álvarez, Javier Fabra, and Joaquín Ezpeleta. 2017. Analysis of Users' Behavior in Structured e-Commerce Websites. *IEEE Access* 5 (2017), 11941–11958.

[12] Sharon Hirsch, Ido Guy, Alexander Nus, Arnon Dagan, and Oren Kurland. 2020. Query Reformulation in E-Commerce Search. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2020)*. 1319–1328.

[13] Dietmar Jannach, Ahtsham Manzoor, Wanling Cai, and Li Chen. 2021. A Survey on Conversational Recommender Systems. *Comput. Surveys* 54, 5 (2021), 1–36.

[14] Yuka Kawada, Takehiro Yamamoto, Hiroaki Ohshima, Yuki Yanagida, Makoto P. Kato, and Sumio Fujita. 2023. Analyzing Web Search Queries of Before and After Purchase on e-Commerce Site. In *Proceedings of the 25th International Conference on Asia-Pacific Digital Libraries (ICADL 2023)*. 181–187.

[15] Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, Sebastian Riedel, and Douwe Kiela. 2020. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. In *Proceedings of the 34th International Conference on Neural Information Processing Systems (NeurIPS 2020)*. 9459–9474.

[16] Yuanxing Liu, Weinan Zhang, Baohua Dong, Yan Fan, Hang Wang, Fan Feng, Yifan Chen, Ziyu Zhuang, Hengbin Cui, Yongbin Li, and Wanxiang Che. 2023. U-NEED: A Fine-grained Dataset for User Needs-Centric E-commerce Conversational Recommendation. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2023)*. 2723–2732.

[17] Shengnan Lyu, Arpit Rana, Scott Sanner, and Mohamed Reda Bouadjenek. 2021. A Workflow Analysis of Context-driven Conversational Recommendation. In *Proceedings of the Web Conference 2021 (WWW 2021)*. 866–877.

[18] Hendrik Müller, Aaron Sedley, and Elizabeth Ferrall-Nunge. 2014. *Survey Research in HCI.* Springer New York, 229–266.

[19] Xi Niu, Xiangyu Fan, and Tao Zhang. 2019. Understanding Faceted Search from Data Science and Human Factor Perspectives. *ACM Transactions on Information Systems* 37, 2 (2019), 1–27.

[20] Gustavo Penha, Eyal Krikon, Vanessa Murdock, and Sandeep Avula. 2022. Helping Voice Shoppers Make Purchase Decisions. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems (CHI EA 2022)*. 1–8.

[21] Filip Radlinski, Craig Boutilier, Deepak Ramachandran, and Ivan Vendrov. 2022. Subjective Attributes in Conversational Recommendation Systems: Challenges and Opportunities. *Proceedings of the AAAI Conference on Artificial Intelligence* 36, 11 (2022), 12287–12293.

[22] Nikitha Rao, Chetan Bansal, Subhabrata Mukherjee, and Chandra Maddila. 2020. Product Insights: Analyzing Product Intents in Web Search. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management (CIKM 2020)*. 2189–2192.

[23] Daniel E. Rose and Danny Levinson. 2004. Understanding User Goals in Web Search. In *Proceedings of the 13th International Conference on World Wide Web (WWW 2004)*. 13–19.

[24] Sosuke Shiga, Hideo Joho, Roi Blanco, Johanne R. Trippas, and Mark Sanderson. 2017. Modelling Information Needs in Collaborative Search Conversations. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2017)*. 715–724.

[25] Parikshit Sondhi, Mohit Sharma, Pranam Kolari, and ChengXiang Zhai. 2018. A Taxonomy of Queries for E-commerce Search. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval (SIGIR 2018)*. 1245–1248.

[26] Ning Su, Jiyin He, Yiqun Liu, Min Zhang, and Shaoping Ma. 2018. User Intent, Behaviour, and Perceived Satisfaction in Product Search. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining (WSDM 2018)*. 547–555.

[27] Maarten Van Someren, Yvonne F Barnard, and J Sandberg. 1994. The Think Aloud Method: A Practical Approach to Modelling Cognitive. *London: AcademicPress* (1994).

[28] Ryen W White. 2024. Advancing the Search Frontier with AI Agents. *Commun. ACM* (2024).

[29] Ryen W. White, Matthew Richardson, and Wen-tau Yih. 2015. Questions vs. Queries in Informational Search Tasks. In *Proceedings of the 24th International Conference on World Wide Web (WWW 2015 Companion)*. 135–136.

[30] Yuki Yanagida, Makoto P. Kato, Yuka Kawada, Takehiro Yamamoto, Hiroaki Ohshima, and Sumio Fujita. 2023. What Web Search Behaviors Lead to Online Purchase Satisfaction?. In *Proceedings of the 15th ACM Web Science Conference 2023 (WebSci 2023)*. 324–334.

[31] Yongfeng Zhang, Xu Chen, Qingyao Ai, Liu Yang, and W. Bruce Croft. 2018. Towards Conversational Search and Recommendation: System Ask, User Respond. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management (CIKM 2018)*. 177–186.