MOJI: Enhancing Emoji Search System with Query Expansions and Emoji Recommendations

Yoo Jin Hong
dbwk18@kaist.ac.kr
KAIST
Daejeon, South Korea

Hye Soo Park
hyehye@kaist.ac.kr
KAIST
Daejeon, South Korea

Eunki Joung
eunki@kaist.ac.kr
KAIST
Daejeon, South Korea

Jihyeong Hong
z.hyeong@kaist.ac.kr
KAIST
Daejeon, South Korea

ABSTRACT

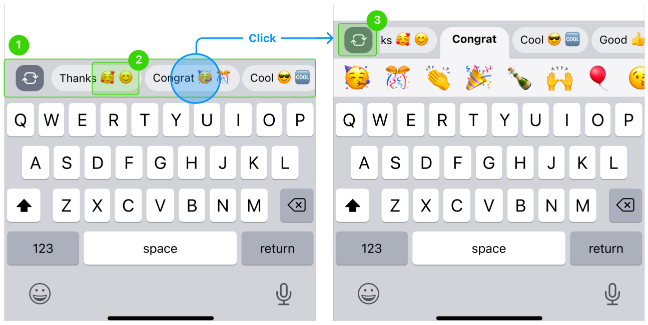


Figure 1: The MOJI system interface design. (1) The keyword-emoji picker that supports query expansion (2) The emoji preview that presents emoji prediction (3) The refresh button for new suggestions

The text-based emoji search, despite its widespread use and extensive variety of emojis, has received limited attention in terms of understanding user challenges and identifying ways to support users. In our formative study, we found the bottlenecks in text-based emoji searches, focusing on challenges in finding appropriate search keywords and user modification strategies for unsatisfying searches. Building on these findings, we introduce MOJI, an emoji entry system supporting 1) query expansion with content-relevant multi-dimensional keywords reflecting users' modification strategies and 2) emoji recommendations that belong to each search query. The comparison study demonstrated that our system reduced the time required to finalize search keywords compared to traditional text-based methods. Additionally, users achieved higher satisfaction in final emoji selections through easy attempts and modifications on search queries, without increasing the overall selection time. We also present a comparison of emoji suggestion algorithms (GPT and iOS) to support query expansion.

CHI EA '24, May 11–16, 2024, Honolulu, HI, USA

ACM ISBN 979-8-4007-0331-7/24/05
https://doi.org/10.1145/3613905.3650838

CCS CONCEPTS

• **Human-centered computing → HCI design and evaluation methods.**

KEYWORDS

User Experience, Interaction Design, Information Retrieval, Computer-mediated Communication, Usability

ACM Reference Format:
Yoo Jin Hong, Hye Soo Park, Eunki Joung, and Jihyeong Hong. 2024. MOJI: Enhancing Emoji Search System with Query Expansions and Emoji Recommendations. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '24), May 11–16, 2024, Honolulu, HI, USA.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3613905.3650838

1 INTRODUCTION

Emojis are important nonverbal cues in text-based computer-mediated communication. Emojis are used to help express one's feelings, lighten the mood, deliver the context, and reduce text ambiguity [8]. iOS started with 471 emojis in 2008 and reached 3,664 emojis this year, and the complexity of finding the "appropriate" emoji to express the text has increased accordingly. Despite the increasing scale and complexity of the emoji search process, the current emoji input interface only allows two search methods: scrolling through a tray or typing search queries. The scroll method is a "linear search task" for emojis, which gets harder and more frustrating as the number of emojis grows [18]. Also, the unpredictability of simple keyword searches creates a challenge for users attempting to force their mental model to fit with existing keyword-emoji pairs in the search system. This unpredictability contributes to the complexity of the search process, despite users expecting the emoji search task to be straightforward.

In this paper, we aim to understand the difficulties users have when using text-based emoji search and improve the experience by incorporating the concept of query expansion and keyword suggestion interface building on existing users' coping strategies. Our formative study (N=12) revealed the difficulties of text-based emoji search and different coping strategies of users when the search did not produce desirable results. Based on the findings, we developed MOJI, a keyword-based emoji search support system to curate the user's emoji search process by utilizing the concept of query expansion in the information retrieval theory [5]. Our interface design supports users to easily explore and iterate the decision of completing the search keyword and choosing the emoji. The queries were presented with four keyword categories derived from the user's coping strategy in the formative study: synonyms, hypernym, in-sentence replacement, and serendipity. We also introduce an emoji suggestion pipeline: iOS condition which retrieves emojis from conventional iOS emoji suggestion algorithm, compared with the GPT condition which uses the GPT model to retrieve emojis with semantically more relevant recommendations for expanded queries. Our user study (N=12) with three conditions (No suggestion, GPT, iOS) revealed that our system, MOJI, helped users finalize their search keyword faster by supporting more lightweight keyword revisions in a short time.

Our contributions are as follows:

- Empirical knowledge of user hardships and coping strategies in search-based emoji entry
- Algorithm for content-relevant multi-dimensional keywords and emojis suggestion
- Interface that supports query expansions and emoji recommendations in non-pervasive, small, and easy-to-use

2 RELATED WORKS

2.1 Difficulties of Emoji Usage

The studies of emojis have revealed that their ambiguity of interpretation makes the use of emojis challenging. Even simple emojis like "laughing with mouth open" were perceived to have different emotions by different readers [14], and presenting emojis in line with texts did not resolve ambiguity [13]. However, it is unclear how the ambiguity affects the emoji search process, and its difficulties have not been researched extensively. Previous research has focused on difficulties of search due to the extended number of emojis [1, 6, 11, 18] without observing the user's search behavior.

In this study, we observe and categorize the hardships of the emoji search process and design a system to support choosing the emoji that presents the context of text, close to our real-life emoji usage.

2.2 Query Expansion For Information Retrieval

In addressing the challenges inherent in traditional information retrieval, Azad et al. [2] identified several complexities. The main challenge lies in the search system relying solely on exact matches of index, and uncertainty of query selection until the results appear. Efforts to address these challenges have led to various attempts at QE(Query Expansion), especially automatic [4, 9] and interactive QE [22], which is a method to improve the information retrieval performance by fetching relevant terms in original query [5].

The difficulties outlined in information retrieval are similar to that users encounter in emoji search, particularly in the context of label-based searches. However, there has been no attempt to enhance users' text search experience in the emoji domain through query expansion. In response to this gap, our work aims to mitigate the identified challenges in emoji search. While the prevalent use case in emoji text search aligns with the manual QE, we seek to explore the application of automatic QE to reduce user burdens and facilitate in obtaining the desired emoji search results.

2.3 Emoji Prediction and Suggestion

There were many attempts to predict and suggest emojis based on the given text. Machine learning techniques such as support vector machine and random forest have been used to predict the accurate emoji [20]. Some tried to replicate the real-life emoji usage, such as Seq2Emoji [16] predicting multiple emojis to convey diverse emotions and stories for the main text. Several studies focused on understanding the context with richer inputs, such as referring to not only the most recent sentence but also the preceding sentences in a conversation [10], individual preferences and user gender [23], and image that is posted along the text [3].

However, most studies on emoji prediction and suggestion proposed the pipeline without considering the user's search methods and addressing how the results should be presented. Our system encompasses the technical pipeline built upon the text-based search methods that users employ, along with a non-pervasive and easy-to-use interface.

3 FORMATIVE STUDY

A formative study was done to find the hardships of text-based emoji search and people's strategies for describing ambiguous emojis in text queries. A total of 12 students from the researchers' university participated in this study.

3.1 Method

Each user was required to fill an emoji in each provided blank embedded in the sentences. Two filling tasks varied by their purposes (public event announcement and personal social media post), each containing 10-14 blanks. The order and the topic of the tasks were counterbalanced. The whole task took 10-20 minutes. After finishing the tasks, the semi-structured individual interviews were conducted.

To examine the difficulties, the participants were asked to describe the failure experiences during the text-based search. For the quantitative measures, we calculated the average number of revisions in each blank, as these revisions are presumed to signify difficulties encountered during emoji search. From the observation and the interview, we summarized three bottlenecks during the process of emoji search.

3.2 Result

3.2.1 Bottleneck 1: Finding search query candidates. The first bottleneck is in translating the mental representation into search keywords, as reported by 9 out of 12 individuals. The interviews showed emojis are not always easily expressed in verbal representation at first. This is in line with a study showing that emojis are processed not only verbally but visuo-spatially [7]. Users reported the hardships of converting the visual image of an emoji in mind to a search keyword. For instance, P6 said, *"I have an image of the emoji I wanted in my head, but it was hard to convert it into words."*. Within this context, people also struggled to find emoji keywords that suited the given context. P1 mentioned their difficulty as "I couldn't remember which emoji keywords would fit this context".

3.2.2 Bottleneck 2: Predicting search results. The uncertainty of the prediction of the search result made search query selection difficult.

6 out of 12 participants said that the keyword-emoji pair used in a search system is hard to predict because the interpretations of emojis can be diverse. One participant remarked, *"I wondered if I could find emojis I want when searching with another query."* (P5). This difficulty was particularly highlighted for the object emojis in announcement tasks. In these tasks, five blanks (e.g., "hoodies": 2.833 average revisions) led to more than or equal to 1 average revision. Users reported that the object emojis tended to have clear emoji representation in mind and it was difficult to know whether that emoji did not exist or they were just failing to find it, which led to repetitive revisions.

3.2.3 Bottleneck 3: Refining search keywords. The third bottleneck is a time-consuming query revision process for unsatisfactory results. When their search keywords do not return a proper emoji or do not show any result at all, users have to change the keywords or adapt to the result. 7 individuals found this aspect to be particularly challenging. Examining the individual revision process of finding satisfactory search results, we analyzed five different coping strategies. Users tried to search (1) hypernyms or (2) synonyms of the previous search keywords. For instance, when the search with "hoodies" did not return a result, P11 tried a new search with "clothes"(hypernym). When searching for "cheer" failed, P6 tried again with "celebrate" (synonym). They also utilized (3) other keywords appearing in the sentence, which means that users chose different words or phrases in the sentence to represent the emoji. Sometimes, users accepted the (4) unexpected search result as a serendipity. This includes cases where users enjoyed putting words with less semantic similarity but with similar sounds, such as "hamster" for "ham". Giving up on the text-based search and (5) browsing from the entire emoji tray was also a common coping strategy.

4 SYSTEM DESIGN

In our formative research, we found that users struggled at three points when using emoji text search. Based on these findings, we derived the design goals for our proposed system.

- DG1: Reduce the cognitive load on users by supporting them in selecting their emoji search query candidates.
- DG2: Enhance the predictability of search results for each search query through emoji recommendation.
- DG3: Reduce the amount of time users spend refining their search queries by supporting search query expansion.

To support the DG1, our interface provides initial suggestions for search query candidates. Second, our interface displays keyword-emoji pairs, presenting emoji previews for each suggested search query. Lastly, to support DG3, our keyword suggestion methods were based on the four query refinement strategies identified in the formative study. We also added the refresh button to offer users a simple way to find better results when the results are unsatisfied. The interface design is described in Fig 1.

We implemented the query expansion and emoji recommendation pipeline to support our interface. The workflow of technical pipeline is illustrated in Fig 6. At first, we generated a set of query samples based on the seed query, which was the single word nearest to the input place. Based on this seed keyword, the suggestions of queries include four categories: synonym, hypernym, in-sentence replacement, and serendipity (homonym). To support the query expansion, the language model (OpenAI's GPT-3.5-turbo-1106) was prompted to suggest search queries based on each method. For synonyms and hypernyms, the model was directly asked to produce them, while serendipity was explained as a keyword that either reuses parts of the seed keyword or sounds similar to them. In the case of in-sentence replacement, the model is requested to pick the words that appeared in the sentence that will be fun and meaningful to describe. To guide the model, each prompt was augmented with 1-2 manually checked examples. The example of prompts of query expansion is provided in Appendix Table 2.

For the emoji recommendation, we built the pipeline with two design options to compare a GPT-generated emoji recommendation with the iOS default emoji recommendation. In the iOS condition, we used an Apple script and the default Apple emoji finder to retrieve the emoji recommendation results by providing our collected query sample. We deleted the queries that did not find any emoji search results. In the GPT condition, we used OpenAI API in the Python server, prompting the language model to suggest 7 emojis for each query in our query sample. An example of keywords and emojis suggested in the recommendation process for the iOS and GPT options is provided in Appendix Table 1.

In total, 20-30 queries and corresponding emojis were collected for each input. To limit the number of displayed recommendations at once, we set a rule to distribute the queries from different coping strategies. In this rule of arrangement, at most five queries are presented at once. Synonyms and hypernyms, the more popular strategies adopted by users, appear earlier than the in-sentence replacement. At the same time, 1-2 serendipity queries are always displayed. The interface that incorporates this full pipeline was built with React.js [19] and Next.js [15] for user experiments.

5 EVALUATION

We conducted a user study to evaluate MOJI's effectiveness in helping users cope with difficulties encountered during the emoji search process. We focused on measuring the impact on user difficulties identified through our formative study, which are 1) increased entry time for emoji searches due to the conversion of image-to-text for finding search keywords, and 2) an increased number of revisions caused by unpredictable search results, leading to continuous adjustments of keywords to generate the desired emoji output. We also 3) compared the difference between GPT and iOS emoji suggestion outputs that better support the search query expansion.

5.1 Participants

We recruited a total of 12 participants from the researchers' university, all of whom had experience with emoji searches and were users of iOS mobile phones. The study lasted around 40 to 60 minutes, and participants volunteered to take part. Our study was approved by the university's Institutional Review Board (IRB).

5.2 Procedure

All study sessions were conducted in person. We set up our system on the researcher's mobile phone to maintain a consistent environment for the keyboard, initializing the customized user settings in the emoji suggestion keyboard.

We conducted a within-subject study with three conditions: baseline, GPT, and iOS. For the baseline condition, we used the existing iOS default emoji search interface to compare our system. We chose the iOS emoji search interface as our baseline due to its widespread usage and familiarity among users, recognizing the significant influence of users' habitual usage patterns. The order of the task conditions was counterbalanced as described in the Appendix Table 3b, and participants were not initially aware of the distinctions among conditions. In each task, participants were presented with 12 sentences, each sentence containing an empty blank, which users were instructed to fill in with an emoji that came to mind. Three different sets of 12 sentences were created manually by researchers for each task. The sentences were shown one by one to independently measure the time for each single emoji search. The task for each condition was constructed with two types of paragraphs: one about announcement posts, and the other about personal posts on social media, with each paragraph consisting of six sentences. These scenarios were created to simulate text-writing situations where emojis are commonly used. There were no predefined answers for each blank.

To examine users' search behavior, we asked participants to utilize at least one search function for each blank in the system across all baseline, iOS, and GPT conditions. We allowed users to browse through the emoji tray after experiencing search failures for each blank. After each condition, participants completed a survey regarding their experience with the task using the system. The survey questions described in the Appendix Table 3a were consistent across all conditions. The questionnaire focused on the usability of the system interface, satisfaction with the search process and search results when using the system. The questions included the USE questionnaire [12].

Our hypothesis regarding the effectiveness of our system focused on measuring the hardships encountered during the emoji search process, including search query entry time, emoji entry time, and the number of revisions.

- H1. Users will spend less time completing search keywords and entering the emoji with our query expansion systems (GPT and iOS conditions).
- H2. Users will revise their search queries less with our query expansion systems (GPT and iOS conditions).

To assess user search behavior and satisfaction with the emoji suggestion of the GPT condition compared to the iOS condition, we also investigated the final emoji keyword choices made within the system in both conditions.

- H3. Among keyword suggestion systems, the rate of keyword choice recommended by our system corresponding to the final selected emoji will be higher in the GPT condition compared to the iOS condition.

5.3 Metrics

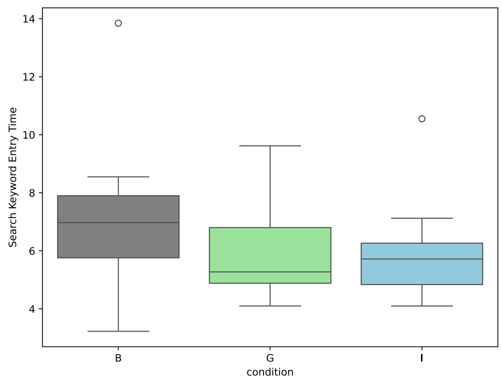
Based on our hypothesis, we identified metrics to measure during the emoji search process. To test the first hypothesis, we measured the *search keyword entry time*, which represents the time spent entering the complete search keyword. This measurement aims to quantify the cognitive load users experience when deciding on a search keyword. Additionally, we measured the *emoji entry time*, representing the time spent entering the final chosen emoji to quantify the overall search load. For testing the second hypothesis, we measured the *count of search keyword revisions* that users encountered during a single emoji entry. In the case of the third hypothesis, we counted the *final choices of search keywords within our keyword suggestions* to compare the utility of suggested emojis between the GPT and iOS conditions. To capture these metrics, we logged user clicks and typing activities within our system. Specifically, for search keyword entry time and emoji entry time, we recorded all the timestamps when users started to look at each question, typed in or clicked on each search keyword, and selected the final emoji. For final choices within the suggestions, we also logged the keywords selected during revisions, as well as the final keyword choices made by users.

5.4 Results

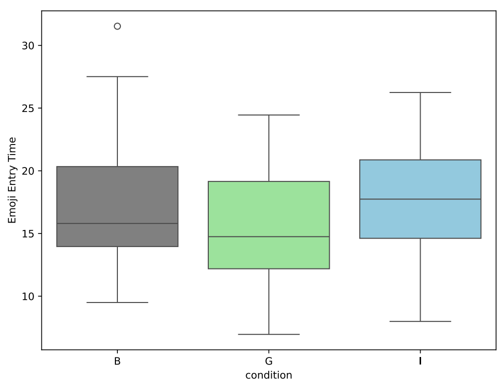
5.4.1 Entry Time. To assess users' load in searching for emojis, we incorporated measures of both search keyword entry time and emoji entry time during the search process. As the data collected from the experiment did not conform to the normal distribution, we applied Friedman's test for analyzing search keyword entry time and emoji entry time. Our analysis revealed a statistically significant result for search keyword entry time ($p < .05$) across different conditions (Fig 2a). This suggests that users experienced reduced load for completing the search keywords with our keyword suggestion system. However, there was no statistically significant difference in emoji entry time (Fig 2) indicating that there were minimal differences in total search load between the baseline and our system.

5.4.2 Search Keyword Revisions. We compared the number of search keyword revisions for single emoji entries to examine the revision behavior of emoji searches within our system. Using Welch's ANOVA test, we compared the revision counts, revealing a statistically significant difference in mean among the three conditions ($p < .05$) in Fig 3. Subsequently, we conducted pairwise Games-Howell tests, indicating a significant difference between the baseline and GPT ($p < .01$), as well as between the baseline and iOS ($p < .05$). Both GPT and iOS conditions exhibited higher revision numbers compared to the baseline. To examine the factors behind the increased search keyword revision attempts using our system, we looked into the logs of the revision search keywords. The log of the search revision flow showed repetitive keywords, showing that users looked into the same three to four search keywords over and over again such as *['Style', 'Shape', 'Image', 'Art', 'Vi(Vi-su-al)', 'Style', 'View', 'Al(Vi-su-al)', 'Shape', 'Su(Vi-su-al)', 'Style', 'Shape', 'Image', 'Art', 'Vi(Vi-su-al)']*. This behavior implies that people had attempted to explore and compare the recommended emojis corresponding to the search keyword. Additionally, in the post-experiment survey, P2 mentioned that *"It was fun to explore what emojis were included in those suggested keywords!"*. This statement also suggests a potential factor contributing to increased search revision numbers in our system.

5.4.3 Final Search Keyword Choice. To assess the utility of suggested emojis between the GPT and iOS conditions, we measured



(a) Time completing the search keyword



(b) Time completing the emoji selection

Figure 2: Entry time measures

the counts of final keyword choices made within our system corresponding to the final selected emoji. The Wilcoxon signed-rank test was conducted to compare the GPT and iOS conditions, revealing that the proportion of final choices made within the system did not significantly differ between the two suggestion conditions (Fig 4). However, looking into the search log for each question, we found out that the users in the iOS condition tried to use the text search feature directly without attempting to try out our suggested keywords. In the iOS condition, 5 participants tried to use the text search feature directly, while only 1 participant directly tried out the text search feature in the GPT condition. Related to this finding, the post-survey also shows users' greater satisfaction with GPT condition keywords compared to the iOS condition keywords. Among the seven participants who noticed differences between GPT and iOS conditions, six individuals reported being more satisfied with GPT keywords. P12 noted, *"First condition (GPT) felt like the keywords were more relevant and expanded the scope of emoji usage, whereas the last condition (iOS) felt a bit random."*. This result shows the implication of the query expansion ability in GPT and iOS conditions. Due to the emoji recommendation method relying on exact matches of keyword-emoji pairs in the iOS condition, even though we used the same sample of search keywords in both conditions, the iOS condition had more keywords that failed to support

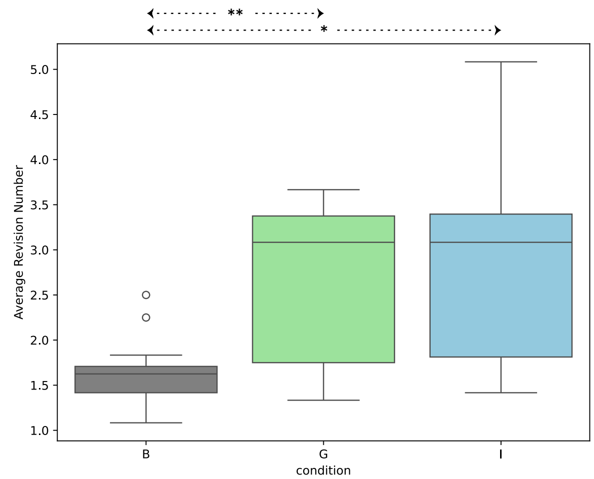


Figure 3: Search keyword revision counts per each entry

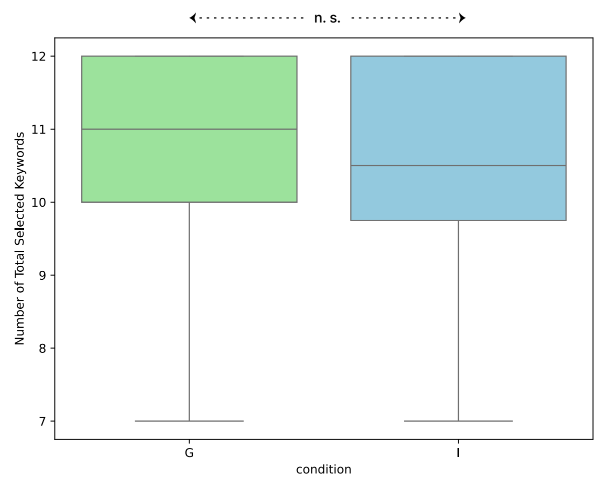


Figure 4: Number of final search keyword choices made within system suggestion

high-quality emoji recommendations for each search keyword. The finding suggests that the GPT condition is better at following the flexible query expansion flow similar to humans.

5.4.4 Satisfaction Questionnaire. We also analyzed the satisfaction questionnaire presented to users after each task. Among the six satisfaction factors (*Speed*, *Relevance*, *Humor*, *Usability*, *Learnability*, and *Enjoyment*), the results indicated statistically significant differences among *Speed*, *Usability*, *Learnability*, and *Enjoyment* (Appendix Fig 7a). These findings suggest that users were satisfied with our system in terms of its speed in searching, ease of use, learnability, and overall enjoyment compared to the existing emoji search system. Regarding questions comparing satisfaction with recommended keywords and emoji samples between GPT and iOS conditions, users expressed greater satisfaction with GPT keywords with statistical significance ($p < .05$) as described in Appendix Fig 7b.

However, there was no statistical difference in satisfaction with emoji samples.

5.5 Discussion

5.5.1 Needs of "Predictable, but Diverse" Emoji Recommendation. In previous emoji recommendation research, metrics such as accuracy and speed that assess how quickly users can find the emoji they were thinking of have been used for evaluations[10, 17, 21]. However, 5 over 12 of the participants cited various emoji recommendations as a benefit of our system such as *"Some of the recommendations were emojis that I don't use often, so I was able to use new and witty emojis."* (P10). Since one of the reasons for using emojis is to create fun and lively texts [8], users are open to emojis that they think would fit well even though it was not intended to be used at first. This result suggests that metrics about the ability to support diverse emoji use are also valuable when evaluating emoji recommendation systems.

5.5.2 Enhancing Usability by Unveiling Query Expansion Process. Through the study, we found out that our system enhances the usability of the emoji search system by revealing the query expansion process. Displaying the search query used for emoji suggestions led to increased explainability of the interface, and participants were able to get a glance at the quality of the recommendation and know whether they were satisfied with it. P3 noted that *"In the first condition(GPT), I was able to select emojis from the recommended keywords without feeling any need to search, but in the latter condition(iOS), I did manual search because I couldn't see any appropriate keywords at first glance."*. Our study extends the work to improve the explainability of the embedded search algorithm by unveiling the query expansion process, resulting in enhanced user usability.

6 CONCLUSION

In this study, a novel interface and pipeline for text-based emoji search was proposed. From the formative user study of the existing emoji search interface, we found the difficulties of textual emoji search and users' coping strategies for unsatisfactory results. Based on these findings, a system to support the user's keyword search query expansion and emoji selection was designed, which is similar to the user's existing search flow. This system suggests search keywords based on the observed coping strategies and seamlessly mixed with the existing emoji search interface. The design options with two search algorithms (a language model's semantic search and traditional dictionary-based search) were evaluated through the within-subject user study. With our system, users could try more diverse search keywords within less load for deciding each keyword compared to the baseline. Semantic emoji search results from GPT were more satisfying than the iOS results, presenting possibilities of different design options in algorithm design for future study. The initial attempt to integrate query expansion and recommendation together in the existing interface with non-pervasive design would contribute to researchers and designers who work on various search interfaces.

REFERENCES

[1] Jessalyn Alvina, Chengcheng Qu, Joanna McGrenere, and Wendy E Mackay. 2019. Mojiboard: Generating parametric emojis with gesture keyboards. In *Extended abstracts of the 2019 chi conference on human factors in computing systems*. 1–6.

[2] Hiteshwar Kumar Azad and Akshay Deepak. 2019. Query expansion techniques for information retrieval: A survey. *Information Processing & Management* 56, 5 (2019), 1698–1735. https://doi.org/10.1016/j.ipm.2019.05.009

[3] Francesco Barbieri, Miguel Ballesteros, Francesco Ronzano, and Horacio Saggion. 2018. Multimodal Emoji Prediction. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers)*, Marilyn Walker, Heng Ji, and Amanda Stent (Eds.). Association for Computational Linguistics, New Orleans, Louisiana, 679–686. https://doi.org/10.18653/v1/N18-2107

[4] Claudio Carpineto and Giovanni Romano. 2012. A Survey of Automatic Query Expansion in Information Retrieval. *ACM Comput. Surv.* 44, 1, Article 1 (jan 2012), 50 pages. https://doi.org/10.1145/2071389.2071390

[5] Efthimis N. Efthimiadis. 1996. Query Expansion. *Annual Review of Information Science and Technology* 31 (1996), 121 – 187. https://www.scopus.com/inward/record.uri?eid=2-s2.0-0030364545&partnerID=40&md5=f17fba56b23e2738ba72363cb9918fe7 Cited by: 303.

[6] Abdallah El Ali, Torben Wallbaum, Merlin Wasmann, Wilko Heuten, and Susanne CJ Boll. 2017. Face2emoji: Using facial emotional expressions to filter emojis. In *Proceedings of the 2017 chi conference extended abstracts on human factors in computing systems*. 1577–1584.

[7] Lauren A Homann, Brady RT Roberts, Sara Ahmed, and Myra A Fernandes. 2022. Are emojis processed visuo-spatially or verbally? Evidence for dual codes. *Visual Cognition* 30, 4 (2022), 267–279.

[8] Linda K Kaye, Helen J Wall, and Stephanie A Malone. 2016. "Turn that frown upside-down": A contextual account of emoticon usage on different virtual platforms. *Computers in Human Behavior* 60 (2016), 463–467.

[9] M. L. Kherfi, D. Ziou, and A. Bernardi. 2004. Image Retrieval from the World Wide Web: Issues, Techniques, and Systems. *ACM Comput. Surv.* 36, 1 (mar 2004), 35–67. https://doi.org/10.1145/1013208.1013210

[10] Joongyum Kim, Taesik Gong, Bogoan Kim, Jaeyeon Park, Woojeong Kim, Evey Huang, Kyungsik Han, Juho Kim, Jeonggil Ko, and Sung-Ju Lee. 2020. No more one liners: Bringing context into emoji recommendations. *ACM Transactions on Social Computing* 3, 2 (2020), 1–25.

[11] Jung In Koh, Josh Cherian, Paul Taele, and Tracy Hammond. 2019. Developing a hand gesture recognition system for mapping symbolic hand gestures to analogous emojis in computer-mediated communication. *ACM Transactions on Interactive Intelligent Systems (TiiS)* 9, 1 (2019), 1–35.

[12] Arnold M Lund. 2001. Measuring usability with the use questionnaire12. *Usability interface* 8, 2 (2001), 3–6.

[13] Hannah Miller, Daniel Kluver, Jacob Thebault-Spieker, Loren Terveen, and Brent Hecht. 2017. Understanding emoji ambiguity in context: The role of text in emoji-related miscommunication. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 11. 152–161.

[14] Hannah Miller, Jacob Thebault-Spieker, Shuo Chang, Isaac Johnson, Loren Terveen, and Brent Hecht. 2016. "Blissfully happy" or "ready tofight": Varying interpretations of emoji. In *Proceedings of the international AAAI conference on web and social media*, Vol. 10. 259–268.

[15] Next.js. 2024. *Next.js by Vercel - The React Framework*. Retrieved Jan 26, 2024 from https://nextjs.org/

[16] Dunlu Peng and Huimin Zhao. 2021. Seq2Emoji: A hybrid sequence generation model for short text emoji prediction. *Knowledge-Based Systems* 214 (2021), 106727.

[17] Henning Pohl, Christian Domin, and Michael Rohs. 2017. Beyond just text: semantic emoji similarity modeling to support expressive communication👫📲😃. *ACM Transactions on Computer-Human Interaction (TOCHI)* 24, 1 (2017), 1–42.

[18] Henning Pohl, Dennis Stanke, and Michael Rohs. 2016. EmojiZoom: emoji entry via large overview maps 😄🔍. In *Proceedings of the 18th international conference on human-computer interaction with mobile devices and services*. 510–517.

[19] React. 2024. *React*. Retrieved Jan 26, 2024 from https://react.dev/

[20] Sayan Saha and Kakelli Anil Kumar. 2022. Emoji prediction using emerging machine learning classifiers for text-based communication. *J. Math. Sci. Comput* 1 (2022), 37–43.

[21] Mingrui Ray Zhang, Ruolin Wang, Xuhai Xu, Qisheng Li, Ather Sharif, and Jacob O. Wobbrock. 2021. Voicemoji: Emoji Entry Using Voice for Visually Impaired People. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 37, 18 pages. https://doi.org/10.1145/3411764.3445338

[22] Neng Zhang, Qiao Huang, Xin Xia, Ying Zou, David Lo, and Zhenchang Xing. 2022. Chatbot4QR: Interactive Query Refinement for Technical Question Retrieval. *IEEE Transactions on Software Engineering* 48, 4 (2022), 1185–1211. https://doi.org/10.1109/TSE.2020.3016006

[23] Guoshuai Zhao, Zhidan Liu, Yulu Chao, and Xueming Qian. 2020. CAPER: Context-aware personalized emoji recommendation. *IEEE Transactions on Knowledge and Data Engineering* 33, 9 (2020), 3160–3172.

A FORMATIVE STUDY

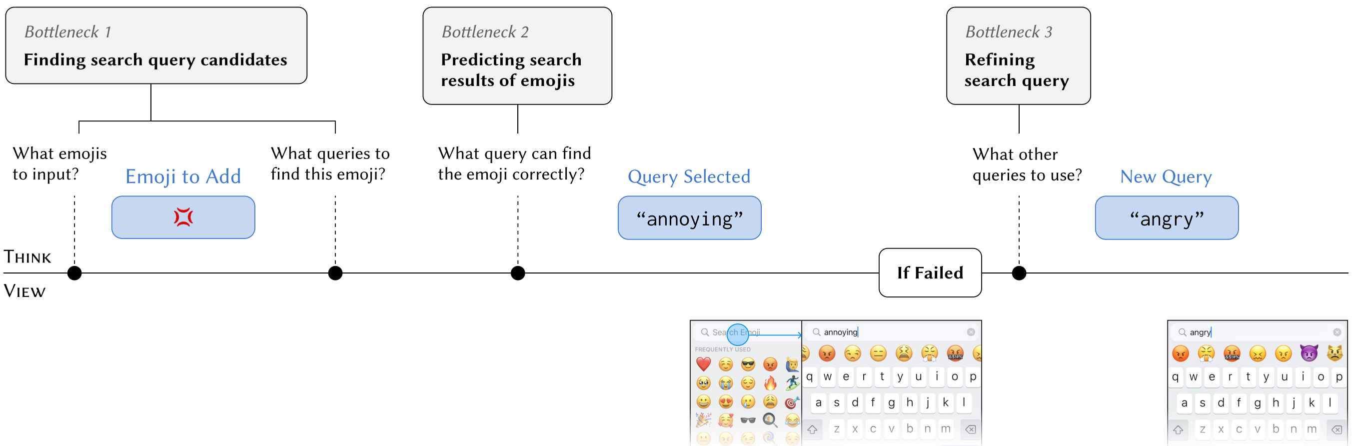


Figure 5: Formative Study Results - Bottlenecks in emoji text search process

B SYSTEM DESIGN

Table 1: An example of query expansion and emoji suggestion. The search results from iOS Emoji may differ from it because it is translated. **A mascot character of the university* ***Same letter with [col](lection) in Korean*

Task sentence		
Nike and Nub-juk-e* have a new **collaborative collection** *[emoji]*.		
Emoji Recommendation		
	iOS	GPT
Synonym	`Costume` 👯 👯‍♀	`Costume` 👚 👖, `Design` 🎨 🖌
Hypernym	`Event` □ 😶	`Event` 🎊 🎉, `Product` 🛍 📦
In-sentence replacement	`New` 🆕	`New` 🆕 🌟 , `Collaborative` 👥 🤝
Serendipity	`Curl**` 💪 □	`Curl**` 💇‍♀ 💅

Table 2: An example of prompts using in pipeline (synonym). The prompt is augmented with 1-2 manually checked samples of [role, content]

	Prompt

Synonym	`Based on the input keyword,`
	`suggest five answers that could be used to search for emoji to depict the similar concept.`
	`for each answer, suggest 7 emojis.`
	`Format response as a json dict: the key should be each answer, and the value should be list of corresponding emojis.`
	`"role": "user", "content": "Input is '" + keyword + "'"`
	`"role": "system", "content": [example of emojis]`

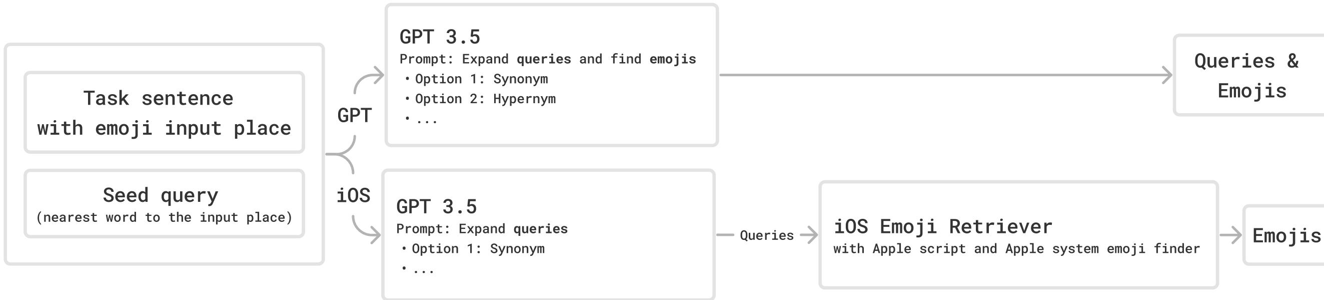


Figure 6: System pipeline for query expansion and emoji recommendation

C EVALUATION

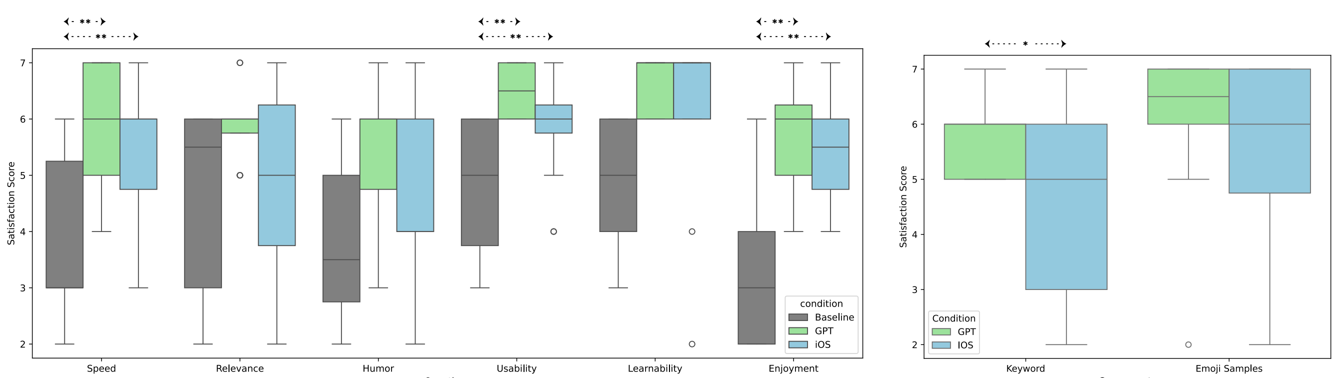
(a) The survey questions of the 7-Likert scale. ***Questions only for evaluation on our interface (iOS, GPT)***

	Questions
Speed	This allowed me to "quickly" find the emoji I wanted.
Relevance	This made it easy to include emojis that "fit the context".
Humor	This made it easy to include "witty and funny" emojis.
Usability	This emoji input method is easy to use.
Learnability	This emoji input method is easy to learn.
Enjoyment	This emoji input method is pleasant to use.
Keyword*	The keywords suggested for emoji search were satisfactory.
Emoji Preview*	The sample emojis presented next to each keyword bubble were helpful in determining keywords

(b) Description of counterbalanced task order

Task Order	**1st**	**2nd**	**3rd**
Sentences	Text 1	Text 2	Text 3
	Condition 1	Condition 2	Condition 3
P01, P07	Baseline	iOS	GPT
P02, P08	Baseline	GPT	iOS
P03, P09	iOS	Baseline	GPT
P04, P10	iOS	GPT	Baseline
P05, P11	GPT	iOS	Baseline
P06, P12	GPT	Baseline	iOS

C.1 Results



(a) Satisfaction toward *Speed*, *Relevance*, *Humor*, *Usability*, *Learnability*, and *Enjoyment* factors among all conditions

(b) Satisfaction toward keywords and emoji list among GPT and iOS conditions)

Figure 7: Responses of post-survey satisfaction questionnaire