Online User Reviews: A Treasure Trove of UX Research?

Wonil Hwang
Professor
Dept. of Industrial and
Information Systems
Engineering
Soongsil University
369 Sangdo-Ro
Dongjak-Gu
Seoul
South Korea
wonil@ssu.ac.kr

Introduction

When we go to unfamiliar places and need to choose hotels or restaurants, or when we want to purchase new products that we haven't experienced before, we often refer to online user reviews. When time is limited, we prioritize checking the average ratings, but if we have a bit more time, we read the comments from users who experienced it. In some cases, online user reviews serve valuable information for decision-making, but in other cases, they are regarded as unreliable information. So, are online user reviews a treasure trove of excellent information for future decision-making, or are they simply a collection of unreliable past user opinions? Are online user reviews only useful for individual decision-making about purchases and usage, or are they a big data of user experiences (UX) that can be utilized for academic studies, practical research, and system improvements?

Traditionally, research on user experience sought to explain and predict the causes and outcomes of user experiences by collecting quantitative or qualitative data according to research plans or to identify problems in interactive systems from a design perspective (Law, 2011). However, with the recent increasing prominence of user reviews on online platforms for interactive systems, there is growing attention toward collecting and utilizing existing, accumulated user reviews as a basis for researching user experiences and deriving problem areas for the improvement of interactive systems. Furthermore, the trend of utilizing online user reviews as user experience data is strengthening with the development of various tools for natural language processing (NLP) and big data analysis.

Online user reviews, which can be relatively easily obtained from real users, offer a wealth of UX data and are undoubtedly appealing. However, as big data containing UX content, caution must be exercised when extracting meaningful insights for scientific research. From a traditional perspective of scientific research methods, online user reviews are qualitative data composed of text, necessitating appropriate analysis methods such as content analysis. However, given that they are big data obtained from thousands or even millions of users, big data analysis methods should also be employed. Additionally, online user reviews are not pre-planned data but rather accumulations of unplanned data, which introduces additional considerations during the analysis and conclusion-drawing process. In this essay, we briefly examine the considerations for utilizing online user reviews in UX research. We account for the characteristics of data collection methods and the pros and cons of online user reviews in UX data. Finally, we discuss future challenges related to this topic.

Getting Data for UX Research

In traditional research methodologies, data collection is often classified into experimental and non-experimental methods based on the level of control over nuisance factors. However, when capturing the characteristics of online user reviews, it is better to distinguish between prospective and retrospective research based on whether the study was planned before data collection rather than relying on categorization.

- **Prospective research**: Experimentation is one of the representative research methods that involve planning and setting variables and hypotheses before data collection. In experiments, nuisance factors are controlled, and reliable data collection is achieved by setting the sample size for the desired statistical power. Much of the existing quantitative research in UX falls under this category. There is also prospective qualitative research in which research designs are planned in advance, and UX data is collected through observations and other methods.

- **Retrospective research**: Retrospective research refers to research conducted by analyzing data that has already been collected for purposes other than research; this data was collected without a pre-existing research plan. Online user reviews record user feedback on interactive systems but are not collected for research purposes. Therefore, UX research using online user reviews can be classified as retrospective research. Such research is challenging to control for the influence of nuisance factors, which results in lower reliability of research findings. However, it offers the advantage of relatively lower data collection costs.

- **Practical tips for collecting data from online user reviews**: When collecting online user reviews for retrospective research, it is important to consider the following points to derive more reliable research results. First, plan ahead by identifying the desired sources and elements, such as target web pages and links, for extracting the desired data using data scraping or crawling techniques. Second, because the data is automatically obtained from online sources, legal issues such as copyright disputes should be carefully examined. Third, collect any relevant information that can be gathered (user information, rating scores, and timestamps) to control for nuisance factors.

Benefits from Online User Reviews

Users who engage with Information and Communication Technology (ICT) in their daily lives often upload their experiences with interactive systems online, enabling communication with manufacturers, other users, and potential users. UX researchers and practitioners strive to effortlessly leverage these collected online user reviews to conduct UX research and derive usability problems. There are several advantages of using online user reviews for UX.

- **Effective way to approach real user experience**: Online user reviews, which are real users' feedback on interactive systems from real-life usage, exhibit different characteristics than UX findings collected and analyzed in a laboratory setting. The conclusions derived from analyzing online user reviews can be considered superior in terms of ecological validity. Online user reviews are an effective method to easily access real users' experiences in the field.

- **Efficient way to analyze large UX data**: Online user reviews accumulate thousands to millions of cases on the internet every year, and reviews can be collected as data for analysis without additional cost using web-crawling tools. The collected online user reviews, which constitute text-based big data, are typically analyzed for UX research using techniques from text-mining, machine learning, and NLP (Hedegaard & Simonsen, 2013). For instance, sentiment analysis using NLP techniques can be performed on online user reviews. This allows for the basic categorization of reviews into positive, negative, or neutral opinions, and further classification into multiple emotional states such as enjoyment, anger, disgust, sadness, fear, and surprise (Ho et al., 2019). Therefore, utilizing online user reviews enables the analysis of a sufficient amount of data (thousands or millions of user experiences) at a low cost, leading to theoretical or practical conclusions regarding UX.

- **Applicable contexts for online user reviews**: Examples of topic areas in which online user reviews can be academically or practically utilized as UX data are as follows. 1. By comparing and studying user expressions in text form, which are voluntarily provided in online user reviews, we can gain insights into the characteristics and limitations of user text representations in comparison to observational studies on UX. 2. In-depth exploration of the types and frequencies of UX expressions contained in large-scale online user reviews can enable new research on the essence of UX (such as forming new definitions or classifications of UX). 3. By extracting UX issues in interactive systems from online user reviews and analyzing them in conjunction with rating scores, it is possible to establish models for determining the severity, impact level, and prioritization of practical UX issues.

Possible Biases in Online User Reviews

Online user reviews serve as valuable data for UX research, allowing for the collection of large-scale data from real users at a low cost. Despite the advantages of being text data that can be analyzed using text-mining, machine learning, and NLP techniques, UX research using online user reviews inherently takes on the characteristic of retrospective research. Because online user reviews are not collected according to a pre-designed research plan, it is necessary to be cautious of systematic errors (bias) during the analysis process and carefully interpret the results. The following are some typical biases that can be found in online user reviews.

- **Selection bias**: Not all users who use interactive systems can contribute to online user reviews. Because the users who provide reviews are not randomly selected from the entire user population, we should suspect the presence of selection bias. Users voluntarily contribute online user reviews, so if only certain users with specific motivations or personal characteristics write online user reviews, selection bias can occur. For example, if there are incentives for writing online user reviews, users who are sensitive to incentives are more likely to participate, and users who are more active in online communication are more likely to contribute reviews compared to less active users. Therefore, when analyzing online user reviews and drawing conclusions, it is important to consider that they may not represent the opinions of all users due to selection bias.

- **Reporting bias**: Users tend to not record, or are unable to record, everything they experienced in online user reviews. There are often space limitations for users' comments that restrict users from writing long passages. For example, when analyzing 1,866 user reviews of Duolingo (a language learning app) collected from the Google Play Store January 1–December 31, 2022, it was found that users recorded 1–14 sentences, with an average of 5.19 sentences per user ($SD = 1.736$). Due to the nature of online user reviews, if users cannot write long comments, they must selectively choose what to include from their experiences. Depending on their motivations or situations, reporting bias can occur. Users may report only a subset of their experiences; for example, negative experiences may be reported more frequently than positive experiences, or specific event-related experiences may be prioritized over general emotional experiences. Even if not reported as a priority, there is a possibility that other, important user experiences exist. Therefore, when analyzing online user reviews and drawing conclusions, researchers must consider reporting bias.

- **Recall bias**: Online user reviews are typically based on users' memories. Users record their experiences after using interactive systems. According to ISO, user experience is defined as "a person's perceptions and responses that result from the use and/or anticipated use of a product, system, or service" (ISO 9241-11:2018, 3.2.3). Based on this definition, online user reviews do not capture user experiences related to "anticipated use," making them unanalyzable. Additionally, because user experiences are recorded based on memory, experiences that fade early from memory tend to be excluded from online user reviews. Unrecalled experiences remain unanalyzed. The presence of expressions related to emotions in online user reviews may be due to the relative difficulty of retrieving specific memories. Researchers must consider recall bias when analyzing online user reviews and drawing conclusions.

Handling Online User Reviews for UX Research

Online user reviews theoretically and practically offer a cost-effective way to obtain a large amount of text data for UX research. Yet, there is a high likelihood of various biases affecting analysis and conclusions from online user reviews. How should we approach these bias issues?

- **Identify what kinds of bias have a substantial impact**: We mentioned three potential biases that may exist in online user reviews, but there could be other types of biases. To leverage online user reviews in UX research, estimate the types of biases that exist and their potential impacts. For example, there have been studies comparing usability problems derived from online user reviews with those identified through traditional usability evaluation methods such as heuristic evaluation and think-aloud (Hedegaard & Simonsen, 2014). Additionally, there are studies that map frequent keywords extracted from online user reviews to usability attributes and UX dimensions (Weichbroth & Baj-Rogowska, 2019). By comparing the results of UX analysis derived from online user reviews with those obtained through experimental UX analysis or existing UX theories from various aspects, we can empirically identify biases inherent in online user reviews and estimate their influence.

- **Find proper correction methods for biases**: Once the types of biases are identified, including those inherent in online user reviews and their respective impacts, the next step is to compensate for such biases. Compensation enables meaningful conclusions to be drawn from online user reviews in both theoretical and practical contexts. A passive approach might be to interpret the results of online user reviews in a compensatory manner, considering the presence of biases, to derive meaningful insights. A more active approach is to develop a new hybrid research method that combines the advantages of retrospective research with large-scale online user reviews and prospective research through conducting small-scale experiments alongside. Based on the results, bias correction can be attempted, creating a comprehensive research approach.

User reviews will continue accumulating online in the future, and UX research utilizing these reviews will become even more valid in terms of cost-effectiveness, feasibility of analysis techniques, and reliability resulting from a large sample size. If the limitations of retrospective research can be overcome, such as the presence of biases, online user reviews may be a treasure trove of UX research.

References

Hedegaard, S., & Simonsen, J. G. (2013, April 27–May 2). Extracting usability and user experience information from online user reviews. *Proceedings of CHI Conference on Human Factors in Computing Systems* (pp. 2089–2098). New York, NY, USA: ACM.

Hedegaard, S., & Simonsen, J. G. (2014, October 26–30). Mining until it hurts: Automatic extraction of usability issues from online reviews compared to traditional usability evaluation. *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational* (pp. 157–166). New York, NY, USA: ACM.

Ho, V. A., Nguyen, D. H. C., Nguyen, D. H., Pham, L. T. V., Nguyen, D. V., Nguyen, K. V., & Nguyen, N. L. T. (2019, October 11–13). Emotion recognition for Vietnamese social media text. *Proceedings of the 16th International Conference of the Pacific Association for Computational Linguistics, PACLING 2019* (pp. 319–333). Singapore: Springer.

International Organization for Standardization. (2018). *Ergonomics of human-system interaction-Part 11: Usability: Definitions and concepts* (ISO Standard No. 9241-11:2018).

Law, E. L.-C. (2011, June 13–16). The measurability and predictability of user experience. *Proceedings of the 3rd ACM SIGCHI Symposium on Engineering Interactive Computing Systems* (pp. 1–10). New York, NY, USA: ACM.

Weichbroth, P., & Baj-Rogowska, A. (2019, September 1–4). Do online reviews reveal mobile application usability and user experience? The case of WhatsApp. *Proceedings of the Federated Conference on Computer Science and Information Systems* (pp. 747–754). Warsaw, Poland: IEEE.

About the Author

Wonil Hwang
Wonil Hwang is a Professor in the Department of Industrial and Information Systems Engineering at Soongsil University. He received his PhD degree in Human Factors/HCI from Purdue University. His main research interests include usability evaluation, AR/VR interfaces, haptic interfaces, emotional design, and data-driven UX research.