Do Social Media Users Consent to Their Data Being Used for Research? —Ethical Reflections Based on User Acceptance Intention Postprint

Authors: Chen Yi, Li Si

Date: 2023-04-01T00:00:00+00:00

Abstract

[Purpose/Significance] This study investigates social media users' acceptance intention regarding the collection of their data for research purposes, explores the effects of familiarity, trust, and altruism on acceptance intention, and examines the necessity of informed consent procedures. It also calls for scholarly ethical reflection to provide references for establishing ethical norms in social media data research and for conducting ethical risk communication with users. [Method/Process] A psychological mechanism model of users' informed consent process was developed. Data were collected from Sina Weibo users through a questionnaire survey, yielding 320 valid responses for model validation. [Results/Conclusion] Two valid psychological pathways were identified: familiarity → altruism → acceptance intention, and familiarity → trust → altruism → acceptance intention. Familiarity and altruism are critical factors influencing users' acceptance intention. Users exhibit strong ethical needs in social media data research, underscoring the necessity of informed consent procedures. Researchers can proactively demonstrate respect and recognition to enhance users' altruism, increase their acceptance intention, and mitigate ethical controversies in the research process.

Full Text

Do Social Media Users Consent to Their Data Being Used for Research?

An Ethical Reflection Based on User Acceptance Intention

Chen Yi[1], Li Si[2]
[1]School of Information Management, Wuhan University, Wuhan 430072
[2]Department of Information Management, Peking University, Beijing 100871

Abstract:
[**Purpose/Significance**] This study investigates social media users' acceptance intention regarding the collection and use of their social media data for research purposes. It explores how familiarity, trust, and altruism influence acceptance intention, analyzes the necessity of informed consent procedures, and calls for ethical reflection among scholars. The findings provide a reference for establishing ethical norms for social media data research and for conducting ethical risk communication with users. [**Method/Process**] A psychological mechanism model was developed to track users' journey from being informed to giving consent. Data were collected from Sina Weibo users through a questionnaire survey, yielding 320 valid responses for model validation. [**Result/Conclusion**] The study verifies two effective psychological paths: familiarity → altruism → acceptance intention, and familiarity → trust → altruism → acceptance intention. Familiarity and altruism emerge as important factors affecting user acceptance intention. Users demonstrate strong ethical needs regarding social media data research, making informed consent procedures necessary. Researchers can also actively express respect and recognition to enhance users' altruism, increase acceptance intention, and avoid ethical controversies in the research process.

Keywords: ethics; social media data; informed consent; familiarity; altruism; trust

1. Introduction

Social media platforms contain massive amounts of user behavioral data, which researchers increasingly collect and utilize to explore user behavior in virtual worlds [1]. The New York Times has reported that scholars are excited about the prospects of using the vast personal data collected by Facebook, Google, and other companies, believing it could transform social science research [2]. Currently, online data research, including social media data, has become an important data source, with this new form of quantitative and statistical analysis becoming one of the most common approaches in social science research [3].

Users' social media accounts contain substantial personal information. On Twitter, 40-50% of tweets include information about the author [4], potentially including data about their contacts such as personal identity information, location data, health information, and images. Research shows that the authenticity of basic registration information for Sina Weibo users in China ranges from 17.3% to 100% [5]. Even when collected data does not include personal identity information, comprehensive personal and even private information about a user can be obtained through correlation analysis without much difficulty [6-7].

Traditional human subjects research requires procedures including informed consent to ensure that "participants understand what they are doing, what limitations exist, and what risks may arise" [8], representing a critical component of scientific research. However, the development of the internet has introduced a

series of additional complexities for scientific research, such as informed consent [9], terms of service [10-11], relationships with participants [12], and the definition of public versus private spaces [13]. In research using social media data, whether informed consent procedures are necessary remains controversial. The ambiguity of data attributes and the ease of data access often lead researchers to overlook informed consent procedures, and they may not consider potential privacy violations or ethical issues [14-15]. Consequently, many users passively participate in research without their knowledge. In practice, some studies have faced protests from social media users who believe such research is unethical and violates personal privacy [16-17].

Although new research methods and data collection approaches continue to develop, ethical guidelines and policy formulation for online research remain in their infancy [18]. While some international organizations have provided relevant guidelines for researchers [19], these guidelines rarely reference users' perspectives [20] and lack standardized informed consent procedures to provide ethical guidance for research. Some scholars note that research based on social media data is growing continuously, urgently requiring increased researcher awareness of its ethical challenges [21]. B. Brown and colleagues argue that research ethics should be built on "participant sensitivity" and "everyday practice" rather than merely legal issues [22]. Therefore, investigating social media users' acceptance intention for such research and exploring influencing factors should form the foundation for establishing ethical norms and regulating such research.

Sina Weibo is the world' s most influential Chinese-language social media platform, with 523 million monthly active users and 229 million daily active users as of November 2020. This study uses Sina Weibo users as subjects to investigate their actual views on the collection and application of their social media data for research, exploring the factors that influence user acceptance intention. The goal is to remind scholars to engage in ethical reflection when conducting similar research and to provide a reference for establishing ethical norms for social media data research and for conducting ethical risk communication with users.

2. Literature Review

Whether "social media data research" without user consent is ethical has sparked ongoing debate [23]. Some scholars argue that informed consent is unnecessary because social media data is public—users consciously choose to make their activities visible to the world, and unrestricted access to an account implies no privacy expectation [24]. However, research shows that users are not as lacking in ethical expectations as researchers assume. A scholar' s "emotional contagion" experiment using Facebook user data without permission, though involving direct experimentation rather than collecting existing data, triggered strong public protests [8,25]. C. Fiesler and N. Proferes' survey found that over half of respondents did not accept their social media data being collected for research

or indicated it would depend on circumstances [20]. M. Zimmer discovered that although researchers consider Twitter public, users agreeing to public viewing of tweets does not imply consent to collection and analysis [25]. D. Nunan and B. Yeniciglu argue that most current social media data research involves "uninformed consent," creating ethical ambiguity. They view "informed consent" not as a restriction but as enabling participants to fully engage in the research process and be protected from potential harm [26].

In practice, mature informed consent procedures have not yet been developed. B. Custers et al. note that while many current regulations address various aspects of consent, most only specify very general consent scopes without focusing on procedural details. In some cases, there is a disconnect between abstract legal regulations and the concrete implementation of social media frameworks and privacy statements [27].

Some studies have explored factors influencing user acceptance of such research, including demographic factors and user cognitive factors. M. L. Williams et al. studied the relationship between participants' demographic characteristics and their acceptance of data use for scientific research, including gender, age, race, income, and sexual orientation [28]. Research on cognitive factors primarily examines participants' understanding of social media data use for scientific research, with most studies concluding that social media users are not familiar with social media platforms and are unaware that their data may be used for research, making them feel strongly opposed when they discover their data has been collected and used. N. Proferes' research shows that users are generally unaware of APIs' existence, and some may not even realize their tweets are completely publicly visible [29]. Twitter users often have "followers" they cannot imagine, and these unexpected followers may pose certain risks when users post personal messages [30]. Additionally, S. S. J. Lee believes that addressing ethical issues in social media data research is related to trust, arguing that research is relationship-based and that a core challenge in maintaining relationships is resolving trust issues [31]. Other studies have analyzed factors influencing users' willingness to actively share information, such as increased familiarity raising willingness to participate in projects [32-33], and altruistic values increasing knowledge-sharing willingness in online communities [34-35].

In summary, while many studies have discussed ethical issues in social media data research, users—the most direct stakeholders—are mostly unaware that their data is being collected and used, and they oppose unauthorized data collection. Current explorations of users' acceptance intention and influencing factors remain insufficient. In studies of cognitive factors affecting social media users, although researchers have proposed based on experience or descriptive statistics that participants' familiarity with social media data research is an important factor, and trust has been mentioned, these have not been tested with empirical data. Therefore, this study investigates whether users accept their social media data being used for research and explores what factors influence users' consent process, hoping to provide a reference for establishing scientific data use norms

and ethical guidelines while reminding researchers to engage in ethical reflection in the new data research environment.

---

3. Research Process

3.1 Theoretical Background

M. Fishbein and I. Ajzen define behavioral intention as an individual' s action tendency to engage in a specific behavior [36]. It is a cognitive activity reflecting a person' s willingness and conscious plan to undertake a behavior [37]. Their Theory of Reasoned Action and Theory of Planned Behavior emphasize the importance of cognitive factors in determining behavioral intention, assuming that people are rational and form behavioral beliefs based on available information and information processing, thereby influencing behavioral intention formation [37-38]. From these theoretical perspectives, users consenting to their data being used for scientific research can be considered their acceptance intention for such research. Therefore, this study assumes participants are rational actors, exploring how information factors and belief factors influence their intention to accept data use for scientific research.

Based on this, the study establishes a model of factors influencing users'intention to accept data use for scientific research, using users'familiarity with social media data as an information factor and their acceptance intention for social media data research as a measure of consent level. For belief factors, existing research has demonstrated that altruistic values influence people' s willingness to share knowledge [39], and users accepting data use for scientific research can also be viewed as an information-sharing willingness. Therefore, this study hypothesizes that altruism also influences acceptance intention. Meanwhile, trust has been proven to be a precursor factor influencing behavioral intention [40]. Thus, trust and altruism are set as belief factors influencing acceptance intention. The final model is shown in Figure 1 [Figure 1: see original paper], with specific relationships between variables elaborated in Section 3.2.

3.2 Research Hypotheses

(1) Altruism influences acceptance intention. Altruism may be a direct factor influencing users' acceptance of social media data for scientific research. In studies on volunteers' intention to participate in scientific research, altruism has been proven to be one of the influencing factors [33,41]. Existing research shows that altruistic personality traits are a motivation for social media users to share knowledge on social media platforms [39]. In the context of social media data for scientific research, participants receive almost no direct return. They may be more likely to accept their data being used for scientific research out of a motivation to help others.

Therefore, Hypothesis 1 is proposed: Users' altruism positively influences their

acceptance intention for social media data research.

(2) Trust in researchers influences acceptance intention. Trust refers to openness toward people based on good knowledge of their action outcomes [42-43] and is one of the effective ways to reduce complexity [44]. With trust, people tend to take risks. T. Thornton [45] points out that in citizen science projects, trust is a prerequisite for effective cooperation between both parties. In social media data research, compared with traditional human subjects research, social media users and researchers are in weak relationships. Trust depends on users' judgment of researchers' motives and professionalism. If users believe researchers will not use social media data inappropriately in the future—that is, if they have a high level of trust in researchers—then they may accept social media data research.

Therefore, Hypothesis 2 is proposed: Users' trust in researchers positively and significantly influences acceptance intention for social media data research.

(3) Relationships among familiarity, trust, altruism, and acceptance intention. Familiarity is an understanding typically based on prior interactions and experience, as well as knowledge of what others do, why they do it, and when [44], which can reduce uncertainty [46]. Familiarity is a prerequisite for trust [44] because familiarity can increase trust in others or objects by reducing uncertainty [47]. Familiarity creates trust when experiences are favorable or destroys trust when there is distrust [46]. When people participate in clinical research, relative familiarity with the research process leads to higher trust in researchers [47]. Users' familiarity with social media platforms and social media data research is generated based on prior experience. If users understand the various functions of social media platforms, know their terms of service, and are familiar with how researchers collect and use social media data, and these experiences are favorable, they can reduce their doubts about their data being collected and used. Based on prior experience, users may trust researchers more.

Therefore, Hypothesis 3 is proposed: Familiarity positively and significantly influences trust in researchers.

Based on the above elaboration of the relationships between familiarity and trust, and trust and behavioral intention, it can be argued that familiarity may influence acceptance intention through the mediating effect of trust. Familiarity involves understanding others' current behavior, while trust involves beliefs about others' future behavior [44]. If users are familiar with how researchers collect and use social media data and understand their purposes, they can reduce their doubts about their data being collected and used, thus potentially trusting researchers more and enhancing their acceptance intention for social media data research.

Hypothesis 4 is proposed: Familiarity positively and significantly influences users' acceptance intention for social media data research through the mediating effect of trust.

Familiarity also influences users' altruistic values. Some scholars point out that familiarity is a key influencing factor of altruism [48-49], and a sense of familiarity can increase people' s altruism [50]. In this research context, the higher users' familiarity with social media data research, the more they understand the significance of this behavior for others and society. This cognition of the positive consequences of social media data research aligns with users' altruistic values, further strengthening their altruistic values and making them more accepting of ethical issues arising in research.

Therefore, this study proposes Hypothesis 5: Familiarity positively and significantly influences users' altruism.

Hypothesis 6: Familiarity positively and significantly influences users'acceptance intention for social media data research through the mediating effect of altruism.

The influence of familiarity on acceptance intention may also be mediated by both trust and altruism. First, Hypothesis 3 demonstrates that familiarity has a positive effect on trust. Additionally, research has found that people' s trust levels positively and significantly influence altruism [51]. Community trust influences knowledge-sharing willingness on social media through the mediating effect of altruism. When individuals sense a trusting atmosphere in a virtual community, those with high altruism are more likely to freely share information or discuss personal experiences in the community than those with low altruism [52].

In this research context, if users are more familiar with social media platforms and social media data research, positive experiences can enhance their trust in researchers, which in turn enhances their altruism. The motivation to help scientific research is strengthened, leading to higher acceptance intention for social media data research. Therefore, the path from familiarity to acceptance intention is reasonable.

Thus, Hypothesis 7 is proposed: Trust in researchers positively and significantly influences altruism.

Hypothesis 8: The positive influence of familiarity on acceptance intention is mediated by trust and altruism.

3.3 Questionnaire Design

To ensure measurement validity, all scales in this study were derived from existing mature scales or research findings and adapted appropriately to this research context. Table 1 presents the measurement items and sources for each variable. The questionnaire used a 5-point Likert scale. Acceptance intention strength measured users' acceptance or opposition to social media data research, with 5 indicating "very accepting," 4 "accepting," ···and 1 "very opposing." Trust and altruism variables asked respondents to indicate their degree of agreement with statements, with 5 indicating "very consistent," 4 "consistent," ···and 1 "very

inconsistent." The familiarity variable examined users' familiarity level, with 5 indicating "very familiar," 4 "familiar," ···and 1 "very unfamiliar."

3.4 Data Collection

On August 13, 2019, the research team used the "Wenjuanxing" platform to distribute a pilot survey. Based on respondent feedback, some items and ambiguous concepts in the questionnaire were revised to ensure that survey objects could well understand the questionnaire items in the formal survey. A large-scale formal survey was then conducted. To ensure sample representativeness, Weibo users of different ages, genders, education levels, and occupational backgrounds were invited to complete the questionnaire, covering survey objects as broadly as possible. The survey lasted six weeks, ultimately recovering 385 questionnaires, including 320 valid questionnaires (see sample details in Table 2). The sample was distributed across 21 provinces (autonomous regions, municipalities, and special administrative regions) and 47 cities. Among respondents, 59.4% were female and 40.6% were male. Age ranged from 15 to 60, with the 21-30 age group accounting for the largest proportion at 60.3%. Nearly 60% of respondents had posted on Weibo within one month before the survey. Over 75% of respondents had registered for Weibo for four years or more. The vast majority of respondents (76%) had only one Weibo account. Weibo provides users with different privacy setting options, such as limiting "who can @ me," "who can comment," and "visible time range of posts." Among these, 85 respondents limited the "visible time range of posts," representing the largest proportion, while 207 respondents had no privacy settings on their accounts.

3.5 Reliability and Validity Tests

Table 3 shows the reliability and convergent validity test results for the scales. The Cronbach' s Alpha coefficients for all dimensions were greater than 0.7, indicating good scale reliability.

Validity includes content validity and structural validity. The measurement items for variables in this study mainly came from existing scales or literature, underwent rigorous translation and pilot testing, and were revised by experts, thus having good content validity. Structural validity includes convergent validity and discriminant validity. This study used confirmatory factor analysis to obtain indicators for testing both types of validity. During confirmatory factor analysis to obtain validity indicators, the measurement model was modified [54]. According to modification guidance, item T1 in trust and item F4 in familiarity had relatively large correlations with other items, so they were deleted before conducting confirmatory factor analysis again. After deletion, various indicators improved and approached or exceeded recommended values.

Convergent validity refers to the degree of correlation among measurement indicators of the same latent variable. Test indicators include factor loading, average variance extracted (AVE), and composite reliability (CR). Table 3 shows the val-

ues for each indicator. The composite reliability (CR) for each dimension was greater than 0.7, indicating that items in each dimension consistently explained the latent variable. Moreover, the AVE for each dimension was greater than 0.5, indicating good convergent validity for each dimension.

Discriminant validity was assessed by comparing the square root of AVE for a latent variable with correlation coefficients between that latent variable and other latent variables. In Table 4 , the square root of AVE was greater than the correlation coefficients between latent variables, indicating that the scale constructed in this study has good discriminant validity.

After reliability and validity analysis, the measurement model and sample data were considered to have good reliability and validity, providing a solid foundation for subsequent hypothesis testing.

---

4. Research Results

4.1 Users' Acceptance Intention for Social Media Data Research

Among the 320 respondents, the vast majority did not express clear opposition to "Weibo data being used for scientific research,"with only about 10% expressing opposition. However, attitudes changed significantly when asked about "if all your Weibo posting history were collected for research,"with 40% of users clearly expressing opposition or protest, and 27.5% holding a neutral or ambiguous attitude.

Acceptance levels for such research also varied across different scenarios. The proportion of respondents who "strongly opposed" "your Weibo being collected and used without notification" reached 42.8%, while this proportion dropped to 15.3% if "notification after collection and use." Compared with "computer analysis," "human reading" of collected data faced greater rejection, with the proportion choosing "strongly oppose" and "oppose" both being about three times higher. Regarding "indicating data source" and "anonymous processing," respondents clearly preferred "anonymous processing" and did not want their real identity identified. Some respondents stated: "If I agree, it's okay; collecting without permission is offensive" and "I don' t like using people' s Weibo data without informing them."

4.2 Factors Influencing Users' Acceptance Intention for Social Media Data Research

4.2.1 Correlations Among Variables Table 5 shows the correlation coefficients among measurement variables. Correlation analysis results indicate that all relationships between variables are significant, with trust, altruism, and familiarity all positively correlated with acceptance intention.

4.2.2 Direct Effects Test After testing with AMOS software, all parameters approached or exceeded recommended values (^{2}/df = 2.205; GFI = 0.949; AGFI = 0.919; NFI = 0.962; TLI = 0.979; FMIN = 0.339; RMSEA = 0.061), indicating good model fit. Table 6 and Figure 2 [Figure 2: see original paper] show hypothesis test results. Except for the non-significant effect of trust on acceptance intention, all other hypotheses were supported.

4.2.3 Mediation Effects Test Since the direct effect test showed that trust' s direct influence on acceptance intention was not significant, Hypothesis 4 (familiarity influencing acceptance intention through trust) was not supported and was not tested in the mediation analysis.

This study used the Bootstrap procedure to test mediation effects. Using repeated random sampling, 5,000 bootstraps were drawn from the original data (N = 320). In Table 7 , the upper and lower limits of the 95% confidence interval for each path did not contain 0, indicating that all measured mediation effects were significant.

Familiarity influenced acceptance intention through the mediating effect of altruism (b = 0.121, BootSE = 0.043, Bias-corrected 95% CI = [0.047, 0.214], Percentile 95% CI = [0.044, 0.209]), with a positive indirect effect, supporting Hypothesis 6. Familiarity also influenced acceptance intention through the mediating effects of trust and altruism (b = 0.102, BootSE = 0.029, Bias-corrected 95% CI = [0.058, 0.179], Percentile 95% CI = [0.048, 0.163]), with a positive indirect effect, supporting Hypothesis 8.

---

5. Discussion and Implications

The vast amount of user behavioral data provides researchers with rich resources, but most of the time users participate in research without their knowledge. This study explores users' acceptance intention for social media data research and its influencing factors. The survey found that not all social media users agree with and accept this type of research behavior. Researchers need to consider public needs more and seek balance between technological innovation and ethical norms.

5.1 Theoretical Contributions

The survey results show that the vast majority of respondents oppose or strongly oppose researchers collecting social media data without notification. Users' clear opposition also demonstrates strong ethical needs. Researchers cannot naturally assume that social media data belongs to public space and "presume user consent." In the big data environment, traditional social, academic organizations, and regulatory agencies' distinctions between public and private may no longer be applicable to the current online environment. In fact, whether social media

data is public or private depends to some extent on the online settings themselves and whether social media users have reasonable privacy expectations [26]. This study finds that altruism is the main factor influencing user acceptance intention, and users with strong altruism show high acceptance of social media data research.

This study explores the factors influencing user acceptance of social media data research, finding that familiarity has a positive and significant impact on both trust and altruism. This again verifies that familiarity is a prerequisite for trust [52-53] and that familiarity provides support for altruism [54]. When users are familiar with various platform functions and restrictions and understand the processes and purposes of such research, they will enhance their altruistic values out of recognition of the significance of scientific research, making them more willing to accept data use for scientific research. In another path, the trust variable is required—when users' familiarity is higher, they trust researchers more, believing they will properly use and safeguard data, which enhances their motivation to help scientific research and makes them more accepting of their social media data being used for scientific research.

In the two verified paths, altruism is the influencing factor for users agreeing to social media data use for scientific research. In existing research on prosocial behaviors such as environmental protection, green product purchasing, organ donation, participation in medical research, sharing medical data, and knowledge sharing, altruism is an important variable. For example, altruism positively influences environmental protection attitudes and organ donation attitudes [55], and positively influences attitudes and willingness to share knowledge on social media [39]. However, literature on social media data for scientific research has not explored altruism' s influence on user acceptance intention. This study reveals the mechanism between altruism and user acceptance intention, confirming that altruism significantly and positively influences users' acceptance intention for social media data research. Users' recognition of the value of social media data research can reduce their resistance to it, and they are more likely to make choices beneficial to others out of a desire to express social responsibility and obtain internal satisfaction. Trust does not directly influence user acceptance intention but still influences acceptance intention through altruism, indicating that trust remains an important factor in weak-relationship research contexts.

5.2 Practical Implications

This study surveyed user acceptance intention and found that informed consent is necessary; users posting data on Weibo does not equal "consenting" to its collection and use; and familiarity and altruism are key factors influencing user acceptance intention. In practice, researchers and relevant organizations should actively take measures to avoid ethical controversies in the research process.

Implement informed consent procedures to increase user acceptance intention. The study found that familiarity influences user acceptance inten-

tion through the mediating effects of trust and altruism. "Informed consent" involves providing potential participants with clear explanations about all aspects of the research, enabling them to fully understand their role and trust researchers [56]. Making users "informed" makes their concerns about data collection and use predictable, which is a process of increasing familiarity. The informed consent procedure is also a process of building trust [57]. As previously noted, users agreeing to public viewing of tweets does not mean agreeing to collection and analysis [20]. Researchers can convey information to potential participants through this procedure to establish trust relationships, enhance trust in researchers, and increase acceptance intention. Some respondents stated: "As long as I' m informed, I' m happy to join academic research." Increasing familiarity is an effective way to enhance user trust and altruism. Therefore, through informed consent procedures that familiarize users with the research process, potential uncertainties, and relevant rights, trust can be built and altruism enhanced, creating a benign and normative research environment.

Social media platforms should also play an active role in informed consent procedures. Currently, relevant terms issued by social media platforms, such as "Personal Information Protection Policy" and "Developer Agreement," do not have clear specifications for academic research data collection, leaving considerable uncertainty in actual operations. However, foreign social media platforms have explicit provisions for such usage purposes. For example, Twitter' s privacy policy clearly informs users that researchers can use their public data to analyze trends or user insights, advising users to think carefully when posting public information [58]. Social media should become a useful information platform between researchers and users, providing positive guidance to both parties, formulating data collection and use terms for research purposes, and actively pushing them to users. This can regulate researchers' data use behavior, reduce ethical risks involved in the research process and post-hoc ethical controversies, and also reduce user concerns, enhance user trust and familiarity with social media platforms and social media data research, and guide users toward positive attitudes toward scientific research.

Actively expressing respect and recognition can enhance users' altruism. Some respondents explicitly stated: "If it helps academic research, providing Weibo data is feasible." Researchers can prevent ethical controversies in the research process by enhancing users' altruistic values. R. R. Cottone and R. E. Claus [59] point out that people' s ethical attitudes are influenced not only by personal moral principles but also by "external factors" —their interactions with society, such as stakeholder requirements, national laws, social expectations and consensus, and personal or group gains and losses. Researchers can adjust these "external factors" to meet users' ethical expectations. During data collection, researchers should clearly express respect for users, affirm their dedication, enhance users' altruistic values, and thereby increase their acceptance intention for social media data research.

Informed consent means that personal data should be collected, processed, used,

and transmitted only after fully informing data subjects and obtaining their consent. In the big data environment, the huge data volume and ambiguous data attributes raise questions about the validity, necessity, authenticity, and economy of informed consent, but this does not mean researchers can ignore research ethics and arbitrarily collect and use users' online data. This study does not recommend solving this dilemma by prohibiting this research method—completely disallowing the use of online platform data in research may be as problematic as using online platform data without any ethical consideration [20]. Ethical decision-making is a deliberative process in which researchers should consider as many participants and resources as possible, including research collaborators, participants, research review committees, ethical guidelines, and applicable legal precedents [19]. As an exploratory study, this research mainly explores cognitive factors influencing user acceptance intention for social media data research from the user perspective. To improve the theory, future research will introduce more cognitive factors and further consider the influence of external environmental factors, providing a reference for exploring scientific informed consent procedures and formulating relevant ethical norms.

References

[1] JAMES N, BUSHER H. Ethical issues in online research[J]. Education and information technologies, 2016, 21(3): 573-580.

[2] GOEL V. As data overflows online, researchers grapple with ethics[N]. The New York Times, 2014-08-12(1).

[3] KOENE A, PEREZ E, CARTER C J, et al. Research ethics and public trust, preconditions for continued growth of internet mediated research[J]. Social science electronic publishing, 2015, 21(2): 89-94.

[4] NAAMAN M, BOASE J, LAI C H. Is it really about me? Message content in social awareness streams[C]//Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work. New York: ACM, 2010.

[5] CHANG W Y, LIU B. An empirical study on personal information leakage of Weibo users based on credibility analysis—taking Sina Weibo as an example[J]. Journal of intelligence, 2015(8): 169-175.

[6] SCHWARTZ P M, SOLOVE D J. The PII problem: privacy and a new concept of personally identifiable information[J]. Social science electronic publishing, 2011, 86(6): 1814-1894.

[7] British Psychological Society[EB/OL]. [2021-04-08]. https://www.bps.org.uk/news-and-policy/ethics-guidelines-internet-mediated-research-2017.

[8] DOUGLAS M D. Should internet researchers use ill-gotten information?[J]. Science & engineering ethics, 2018, 24: 1221-1240.

[9] BAROCAS S, NISSENBAUM H. Big data' s end run around anonymity and consent[M]//LANE J, STODDEN V, BENDERS S, et al., Eds. Privacy, big data, and the public good: frameworks for engagement. Cambridge: Cambridge University Press, 2014: 44-75.

[10] FIESLER C, LAMPE C, BRUCKMAN A S. Reality and perception of copyright terms of service for online content creation[C]//Proceedings of the ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW). New York: Association for Computing Machinery, 2016: 1450-1461.

[11] VACCARO K, KARAHALIOS K, SANDVIG C, et al. Agree or cancel? Research and terms of service compliance[EB/OL]. [2020-01-20]. https://pdfs.semanticscholar.org/3773/15ac09cd7fba7f1f154b0b1481d7c24ad5eb.pdf?_{ga}=2.80698775.10125228839498.1562862749.

[12] BEAULIEU A, ESTALELLA A. Rethinking research ethics for mediated settings[J]. Information, communication & society, 2012, 15(1): 23-42.

[13] ZIMMER M. "But the data is already public" : on the ethics of research in Facebook[J]. Ethics and information technology, 2010, 12(4): 313-325.

[14] VITAK J, PROFERES N, SHILTON K, et al. Ethics regulation in social computing research: examining the role of institutional review boards[J]. Journal of empirical research on human research ethics, 2017, 12(5): 372-382.

[15] ZIMMER M, PROFERES N J. A topology of Twitter research: disciplines, methods, and ethics[J]. Aslib journal of information management, 2014, 66(3): 250-261.

[16] MCNEAL G S. Facebook manipulated user news feeds to create emotional responses[EB/OL]. [2020-01-20]. https://www.forbes.com/sites/gregorymcneal/2014/06/28/facebook-manipulated-user-news-feeds-to-create-emotional-contagion/#bca47b639dc7.

[17] OkCupid plays with love in user experiments[N]. The New York Times, 2014-07-29.

[18] WARRELL J G, JACOBSEN M. Internet research ethics and the policy gap for ethical practice in online research settings[J]. Canadian journal of higher education, 2014, 44(1): 22-37.

[19] AoIR Ethics working committee. Ethical decision-making and internet research: recommendations from the AoIR Ethics Working Committee (Version 2.0)[EB/OL]. [2020-11-20]. http://www.aoir.org/reports/ethics.pdf.

[20] FIESLER C, PROFERES N. "Participant" perceptions of Twitter research ethics[J]. Social media + society, 2018, 4(1): 1-14.

[21] TAYLOR J, PAGLIARI C. Mining social media data: how are research sponsors and researchers addressing the ethical challenges?[J]. Research ethics, 2017, 14(2): 1-39.

[22] BROWN B, WEILEMANN A, MCMILLAN D, et al. Five provocations for ethical HCI research[C]//Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. New York: Association for Computing Machinery, 2016: 852-863.

[23] CLARK K, DUCKHAM M, GUILLEMIN M, et al. Advancing the ethical use of digital data in human research: challenges and strategies to promote ethical practice[J]. Ethics & information technology, 2019, 21: 59-73.

[24] ZIMMER M. Is it ethical to harvest public Twitter accounts without consent?[EB/OL]. [2020-11-10]. https://www.michaelzimmer.org/2010/02/12/is-it-ethical-to-harvest-public-twitter-accounts-without-consent/.

[25] KRAMER A D I, GUILLORY J E, HANCOCK J T. Experimental evidence of massive-scale emotional contagion through social networks[J]. Proceedings of the national academy of sciences, 2014, 111(24): 8788-8790.

[26] NUNAN D, YENICIOGLU B. Informed, uninformed and participatory consent in social media use: the gap between user expectations and EU Personal Data Protection Law[J]. Social science electronic publishing, 2013, 63(3): 543-557.

[27] CUSTERS B, SIMONE V D H, SCHERMER B, et al. Informed consent in social media use: the gap between user expectations and EU Personal Data Protection Law[J]. Social science electronic publishing, 2013, 63(3): 543-557.

[28] WILLIAMS M L, BURNAP P, SLOAN L. Towards an ethical framework for publishing Twitter data in social research: taking into account users' views, online context and algorithmic estimation[J]. Sociology, 2017, 51(6): 1149-1168.

[29] PROFERES N. Information flows and exploitation in an exploratory study of Twitter accounts. Social media + society[EB/OL]. [2020-11-10]. https://doi.org/10.1177/2056305117698493.

[30] LITT E. Knock, knock. Who' s there? The imagined audience[J]. Journal of broadcasting & electronic media, 2012, 56(3): 330-345.

[31] LEE S S J. Studying "Friends" : the ethics of using social media as research platforms[J]. The American journal of bioethics, 2017, 17(3): 1-2.

[32] VANESSA W, JAMES S A S H O K R, et al. Can visual familiarity influence attitudes toward brands? An exploratory study of advergame design and cross-cultural consumer behaviour[J]. Entertainment computing, 2018, 27: 194-208.

[33] BRANDT D S. "Factors associated with young adults' reported intention of willingness to participate in clinical research" [EB/OL]. [2020-04-08]. https://ir.uiowa.edu/cgi/viewcontent.cgi?referer=&httpsredir=1&article=5604&context=etd.

[34] ZHANG X, LIU S, DENG Z, et al. Knowledge sharing motivations in online health communities: a comparative study of health professionals and normal users[J]. Computers in human behavior, 2017, 75(1): 797-810.

[35] ACAR O A. Motivations and solution appropriateness in crowdsourcing challenges for innovation[J]. Research policy, 2018, 48(8): 103716.

[36] AJZEN I. From intentions to actions: a theory of planned behavior[M]//KUHL J, BECKMANN J. Action control: from cognition to behavior. Berlin: Springer, 1985: 11-39.

[37] FISHBEIN M, AJZEN I. Belief, attitude, intention, and behavior: an introduction to theory and research[M]. Reading, MA: Addison-Wesley, 1975.

[38] ICEK A, DRIVER B L. Application of the theory of planned behavior to leisure choice[J]. Journal of leisure research, 1992, 24(3): 207-238.

[39] OH S. The characteristics and motivations of health answerers for sharing information, knowledge, and experiences in online environments[J]. Journal of the Association for Information Science and Technology, 2013, 55(6): 791-808.

[40] MULLER C E. Examining the inter-relationships between procedural fairness, trust in actors, risk expectations, perceived benefits, and attitudes toward power grid expansion projects[J]. Energy policy, 2020, 141: 111465.

[41] SENGUPTA S, STRAUSS R P, DEVELLIS R, et al. Factors affecting African-American participation in AIDS research[J]. Jaids journal of acquired immune deficiency syndromes, 2000, 24(3): 275-284.

[42] RIEGELSBERGER J, SASSE M A, MCCARTHY J D. Trust at first sight? A test of users' ability to identify trustworthy e-commerce sites[EB/OL]. [2020-04-13]. http://sec.cs.ucl.ac.uk/fileadmin/sec/publications/Riegelsberger_{{{Sasse}}{{McCarthy}}}A{{{Test commerce_{{{Sites}}_{{HCI2003}}}.pdf.

[43] OKYEREKE E, NORK M. Individual factors and knowledge sharing[J]. American journal of economics and business administration, 2011, 3(1): 66-72.

[44] GEFEN D. E-commerce: the role of familiarity and trust[J]. Omega, 2000, 28(6): 725-737.

[45] THORNTON T, LEAHY J. Trust in citizen science research: a case study of the groundwater education through water evaluation & testing program[J]. Jawra journal of the American Water Resources Association, 2012, 48(5): 1032-1040.

[46] LUHMANN N. Trust and power[M]. Chichester, UK: Wiley, 1979.

[47] YOON C, ROLLAND E. Knowledge-sharing in virtual communities: familiarity, anonymity and self-determination theory[J]. Behaviour & information technology, 2012, 31(11): 1133-1143.

[48] BODEN W F, MILLER J, O' GORMAN R, et al. Increased costs reduce reciprocal helping behaviour of humans in a virtual evacuation experiment[EB/OL]. [2020-05-13]. https://www.nature.com/articles/srep15896#citeas.

[49] KHALIL E L. Adam Smith and three theories of altruism[J]. Louvain economic review, 2001, 67(4): 421-435.

[50] KALCHEV D. The impact of positive emotions on altruism in the presence of familiarity[EB/OL]. [2020-05-13]. https://trepo.tuni.fi/bitstream/handle/10024/101633/GRADU-1498464800.pdf?sequence=1.

[51] ZHU B. A study on the influence of individual trust level on altruistic behavior[D]. Jinhua: Zhejiang Normal University, 2016.

[52] CHEN H, FAN H, TSAI C. The role of community trust and altruism in knowledge sharing: an investigation of a virtual community of teacher professionals[J]. Educational technology & society, 2014, 17(3): 168-179.

[53] GENG R L, SHEN J. Research on motivations for knowledge sharing behavior of social network users from different cultural perspectives[J]. Journal of library science in China, 2019, 45(1): 60-81.

[54] QIAN L L. Public service demand for e-government and its impact path on system success[D]. Shanghai: Fudan University, 2010.

[55] MOSTAFA M M. Altruistic, cognitive and attitudinal determinants of organ donation intention in Egypt: a social marketing perspective[J]. Health marketing quarterly, 2010, 27(1): 97-115.

[56] WILES R, HEATH S, CROW G, et al. Informed consent in social research: a literature review[EB/OL]. [2020-05-05]. http://eprints.ncrm.ac.uk/85/1/MethodsReviewPaperNCRM-001.pdf.

[57] SALMONS J. Getting to yes: informed consent in qualitative social research[M]//WOODFIELD K, ed. The ethics of online research. Bingley: Emerald Publishing Limited, 2017: 109-134.

[58] Twitter privacy policy[EB/OL]. [2020-09-20]. https://twitter.com/en/privacy.

[59] COTTONE R R, CLAUS R E. Ethical decision-making models: a review of the literature[J]. Journal of counseling & development, 2000, 78(3): 275-283.

Author Contributions:
Chen Yi: Conceptualization, research framework design, writing.
Li Si: Data analysis and processing, paper revision.

Note: Figure translations are in progress. See original paper for figures.