International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 13 Issue: 10 | Oct 2026 www.irjet.net p-ISSN: 2395-0072

A Security and Privacy Analysis of App Terms and Conditions: Understanding the Risks of Blind User Consent

[1]Ms. Swati Kadu, [2]Tanmay Kadam, [3]Sumit Parkhe, [4]Siddharth Gavahne

[1]Assistant Professor, D. Y. Patil Deemed to be University, School of Humanities & Sciences, Belapur, Maharashtra, India.

[234]Student, Department of MCA, D. Y. Patil Deemed to be University, School of Humanities & Sciences, Belapur, Maharashtra, India.

Abstract - *Almost every mobile app asks for a tap on "I agree" before it will open, yet very few people read the Terms and Conditions (T&C) behind that button. This habit, which we call blind consent, means users routinely hand over personal data, device permissions, and legal rights without knowing what they have given up. This paper proposes a structured method for examining the security and privacy risks hidden in app T&C documents: collecting the agreements of popular apps, measuring their length and readability, sorting their clauses into risk categories, assigning each a transparent risk score, and checking whether the stated data practices match the permissions the app requests. We also outline an evaluation framework and recommendations for developers, regulators, and users. No experimental findings are claimed; the paper offers a complete methodology that can be carried out on a real collection of app agreements.*

Key Words: Terms and Conditions, Privacy Policy, Informed Consent, Mobile App Security, Data Privacy, App Permissions, Readability Analysis, Risk Assessment.

1. INTRODUCTION

Installing an app takes seconds. Between the download button and the home screen there is usually a single screen asking the user to accept the Terms and Conditions and the Privacy Policy, and most people accept without scrolling. The agreement can run to many thousands of words and may allow the company to collect contact lists, location history, and device identifiers, share this data with advertising partners, change the rules at any time, and limit the user's right to take the company to court. Because accepting is the only way to use the app, the choice is closer to "take it or leave it" than to real consent.

Researchers describe this as a failure of the notice-and-consent model, which assumes that people read, understand, and weigh the notices placed in front of them. In practice the volume of agreements and their legal language make this unrealistic [1], [2], [3]. The risk is also a security matter: every additional party that receives personal data is one more place where it can be leaked, misused, or breached. This paper therefore asks a practical question: what exactly are users agreeing to when they accept app T&C, and how can those agreements be examined in a systematic, repeatable way?

2. LITERATURE REVIEW

The difficulty of reading online agreements has been documented for more than a decade. McDonald and Cranor estimated that reading every privacy policy a person meets would cost a very large amount of working time [4]. Obar and Oeldorf-Hirsch showed that most participants in an experiment skipped the agreement of a fictitious social network and accepted deliberately unreasonable clauses [1]. Böhme and Köpsell found that the design of a consent dialog strongly affects whether people click through it [5]. Reidenberg et al. found that experts and ordinary users interpret the same policy statements differently [2], and Solove argued that privacy self-management places too much burden on individuals [3].

Other work tries to automate the reading. Wilson et al. built an annotated corpus of website privacy policies [6]; Zimmeck et al. flag places where an app's behaviour differs from its policy [7]; and Harkous et al. introduced Polisis, which labels policy segments using deep learning [8]. Felt et al. found that Android users' comprehension of permission requests was low [9]. Regulation has also moved: the GDPR requires consent to be freely given, specific, informed, and unambiguous [10], and India's Digital Personal Data Protection Act, 2023 places comparable duties on organisations [11].

3. RESEARCH GAP

Existing studies tend to examine one piece of the problem at a time: how little people read, how they misunderstand wording, or how policy text can be annotated. Fewer works combine these views into one workflow that a small team could repeat on current, widely used apps. Terms and Conditions, which carry clauses on liability, dispute resolution, and unilateral changes, receive less attention than privacy policies, and written promises are rarely put beside the permissions an app actually requests. This paper addresses that gap with a single documented procedure covering readability, clause content, risk scoring, and permission consistency.

4. PROBLEM STATEMENT

Users are expected to give informed consent to app agreements, but their length, complexity, and take-it-or-leave-it presentation make this unlikely in practice. There is a need for a clear, repeatable way of measuring how difficult app T&C documents are to read, identifying which clauses create security and privacy risk, and showing users and developers where the largest problems lie.

5. OBJECTIVES

- To study how blind acceptance of app Terms and Conditions exposes users to security and privacy risk.
- To define a set of risk-relevant clause categories and a transparent scoring scheme for them.
- To compare the data practices stated in the text with the permissions requested by each app.
- To suggest practical measures that can help users, developers, and regulators close the gap between consent and understanding.

6. PROPOSED METHODOLOGY

The study follows a five-stage workflow (Figure 1). Each stage produces an output that feeds the next, and every decision is recorded so that another team could repeat the work.



Figure 1. Five-stage workflow of the proposed Terms and Conditions analysis

7. DATA SOURCES AND SELECTION OF APPS

The documents analysed are those publicly available on each app's store listing and website, so no private user data is needed. Apps are selected by download count and coverage of categories such as social media, messaging, shopping, finance, and gaming. The retrieval date is recorded for every document, the version shown to users in India is preferred where regions differ, and permission data is taken from the store listing or public manifest.

8. CLAUSE CATEGORIES AND RISK FRAMEWORK

Each document is read for the clause types in Table 1, chosen because each affects either the confidentiality of user data or the user's ability to seek a remedy.

Table 1. Clause categories used in the proposed risk analysis

Category	What the coder looks for	Why it matters
Data collection scope	Types of data gathered, including contacts, location, microphone, and device identifiers	Wider collection increases exposure if the data is breached
Third-party sharing	Disclosure to advertisers, analytics firms, affiliates, or unnamed partners	Each recipient is an additional point of failure outside the user's view
Purpose limitation	Whether data use is tied to stated purposes or left open-ended	Vague purposes allow later reuse the user never expected
Retention and deletion	How long data is kept and whether the user can request deletion	Data kept indefinitely stays at risk indefinitely
Unilateral changes	Right of the company to alter terms without direct notice	Consent given once can silently lose its meaning
Liability and dispute resolution	Disclaimers of responsibility, mandatory arbitration, limits on legal action	Reduces the user's remedies after a breach
Security commitments	Statements about encryption, breach notification, and safeguards	Absence of commitments leaves users with no stated protection
User rights and controls	Access, correction, withdrawal of consent, opt-out settings	Controls determine how far the user can undo a past decision

Each category receives a score s from 0 to 3: 0 means the user is protected or the practice is clearly limited; 1 means

it is disclosed with reasonable limits; 2 means the clause is broad; and 3 means it is both broad and vague, or left to the company's discretion. Each category also has a weight w (for instance 1 to 3). The Risk Index is the weighted score as a percentage of the maximum possible:

$$\text{Risk Index} = [\ \Sigma\ (w_i \times s_i) \ /\ (3 \times \Sigma\ w_i)\ ] \times 100 \qquad ...(1)$$



Figure 2. Process for calculating the Risk Index of a T&C document

The weights are a starting proposal, not a fixed standard; the team should justify them and test how app rankings change when they are varied.

Because legal text has long, nested sentences, the score is used to compare documents with one another, not as an absolute verdict on any single agreement.

9. READABILITY AND LENGTH ANALYSIS

Word count and estimated reading time describe the effort a document demands. For difficulty, the Flesch Reading Ease score is widely used; higher scores mean easier text, and scores below roughly 50 are associated with college-level material:

$$\text{Reading Ease} = 206.835 - 1.015 \times (\text{Words} / \text{Sentences}) - 84.6 \times (\text{Syllables} / \text{Words}) \qquad ...(2)$$

10. PERMISSION CROSS-CHECK

A policy can be well written and still fail the user if the app behaves in ways the policy does not describe. The cross-check lists the permissions each app requests (contacts, camera, microphone, precise location, storage) and checks whether the documents explain why each is needed. Three outcomes are recorded (Figure 3). The second, a permission requested but not explained, is the most concerning for privacy, because the user is asked for access without a stated reason.

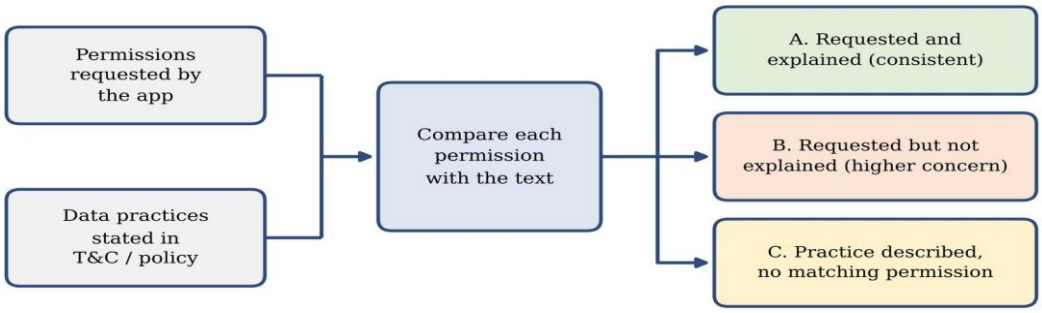


Figure 3. Permission cross-check and its three possible outcomes

11. SYSTEM ARCHITECTURE

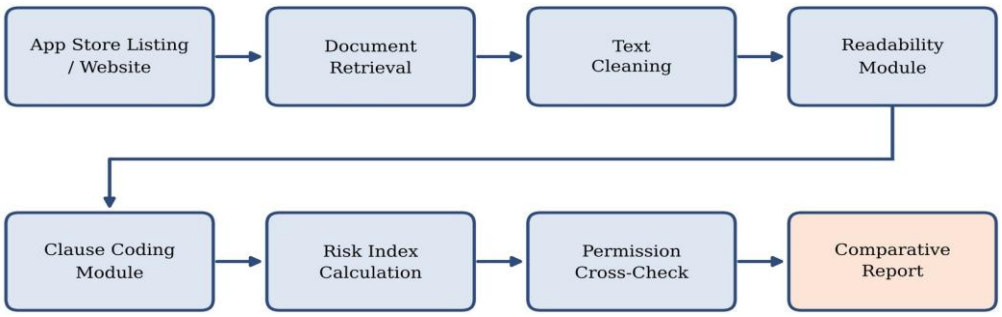


Figure 4. System architecture of the proposed analysis pipeline

The stages are kept separate on purpose. Text cleaning removes menus, repeated headers, and formatting symbols so that only the agreement is analysed. The clause coding module can be run by human coders, an automated classifier, or both, without changing the rest of the pipeline; a semi-automated approach, where a model proposes labels and humans confirm them, keeps a person responsible for the final judgement.

12. EVALUATION FRAMEWORK

This paper proposes a method and reports no measured results. Once carried out, its quality should be judged with the measures in Table 2.

Table 2. Measures for evaluating the proposed analysis

Measure	Purpose
Inter-coder agreement (Cohen's kappa)	Shows whether two coders label the same clause in the same way, which tests whether the categories are clear enough
Weight sensitivity of the Risk Index	Shows whether app rankings stay stable when category weights are changed
Readability and length statistics	Allow documents and app categories to be compared on effort required of the reader
Permission mismatch count	Counts permissions requested without an explanation in the documents
Precision and recall of automated labels	Applies only if a classifier is used; compares its labels with human-coded ones

Agreement between two coders is measured with Cohen's kappa, which corrects for chance agreement:

$$\kappa = (P_o - P_e) / (1 - P_e) \quad \text{...(3)}$$

13. EXPECTED FINDINGS AND DISCUSSION

Based on earlier studies, we would expect most documents to be long and hard to read [1], [2], [4]; third-party sharing, open-ended purposes, and unilateral change clauses to be common sources of high risk scores; and arbitration and liability clauses to appear mainly in the Terms and Conditions. These are expectations from the literature, not findings of this paper.

Results would depend on app selection, clause interpretation, and weights, so these choices and retrieval dates should be reported openly. A high risk score describes what a document permits, not what a company actually does, but the user cannot know the difference from the text, and that uncertainty is itself part of the problem.

14. RECOMMENDATIONS

Whatever the measured results, the literature points to several practical steps.

- **For users:** check permissions before installing, decline those that do not fit the app's purpose, and prefer apps that explain their practices in plain language.
- **For developers:** present a short layered summary ahead of the full text, request permissions when needed, and state retention periods and sharing partners specifically.
- **For regulators and institutions:** enforce specific and informed consent [10], [11], and encourage standard plain-language disclosure formats.
- **For educators:** include agreement literacy in digital safety and cybersecurity teaching.

15. CHALLENGES AND LIMITATIONS

Coding legal texts requires judgement that two people may exercise differently, so coder agreement must be checked. Agreements are revised frequently, and region-specific versions can make it unclear which text applies. The risk scheme simplifies a complicated reality and its weights are subjective. Finally, the analysis covers only written documents and listed permissions; it does not observe real network behaviour, so it cannot confirm whether an app follows its own policy.

16. FUTURE SCOPE

Language models could label clauses automatically and produce plain-language summaries, with human review. Monitoring documents over time would show how agreements change. Dynamic analysis of network traffic could test whether actual data flows match declarations. A browser extension that scores an agreement before the user accepts it, and studies of regional-language agreements, would bring the work to the people it concerns.

17. CONCLUSION

Accepting app Terms and Conditions without reading them is a routine part of digital life, but it is not harmless: behind the "I agree" button there may be broad data collection, wide third-party sharing, rights to change terms at will, and limits on legal remedies. This paper has set out a reproducible method for examining these risks, combining document collection, readability measurement, clause coding with a transparent risk index, and a comparison between written promises and requested

permissions. No experimental findings are claimed, but the framework offers a clear path from a familiar worry to measurable evidence. Closing the gap between legal consent and informed consent will require effort from users, developers, and regulators alike.

18. ACKNOWLEDGEMENT

The authors would like to thank their guide, Ms. Swati Kadu, and the Department of MCA, D. Y. Patil Deemed to be University, School of Humanities & Sciences, Belapur, for their continued guidance and support throughout this work.

19. REFERENCES

[1] J. A. Obar and A. Oeldorf-Hirsch, "The Biggest Lie on the Internet: Ignoring the Privacy Policies and Terms of Service Policies of Social Networking Services," Information, Communication & Society, vol. 23, no. 1, pp. 128–147, 2020.

[2] J. R. Reidenberg, T. Breaux, L. F. Cranor, B. French, A. Grannis, J. T. Graves, F. Liu, A. McDonald, T. B. Norton, R. Ramanath, N. C. Russell, N. Sadeh, and F. Schaub, "Disagreeable Privacy Policies: Mismatches between Meaning and Users' Understanding," Berkeley Technology Law Journal, vol. 30, no. 1, 2015.

[3] D. J. Solove, "Introduction: Privacy Self-Management and the Consent Dilemma," Harvard Law Review, vol. 126, no. 7, pp. 1880–1903, 2013.

[4] A. M. McDonald and L. F. Cranor, "The Cost of Reading Privacy Policies," I/S: A Journal of Law and Policy for the Information Society, vol. 4, no. 3, pp. 543–568, 2008.

[5] R. Böhme and S. Köpsell, "Trained to Accept? A Field Experiment on Consent Dialogs," Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI), 2010.

[6] S. Wilson et al., "The Creation and Analysis of a Website Privacy Policy Corpus," Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (ACL), 2016.

[7] S. Zimmeck, Z. Wang, L. Zou, R. Iyengar, B. Liu, F. Schaub, S. Wilson, N. Sadeh, S. M. Bellovin, and J. Reidenberg, "Automated Analysis of Privacy Requirements for Mobile Apps," Network and Distributed System Security Symposium (NDSS), 2017.

[8] H. Harkous, K. Fawaz, R. Lebret, F. Schaub, K. G. Shin, and K. Aberer, "Polisis: Automated Analysis and Presentation of Privacy Policies Using Deep Learning," 27th USENIX Security Symposium, 2018.

[9] A. P. Felt, E. Ha, S. Egelman, A. Haney, E. Chin, and D. Wagner, "Android Permissions: User Attention, Comprehension, and Behavior," Proceedings of the Symposium on Usable Privacy and Security (SOUPS), 2012.

[10] European Parliament and Council of the European Union, "Regulation (EU) 2016/679 (General Data Protection Regulation)," Official Journal of the European Union, 2016.

[11] Government of India, "The Digital Personal Data Protection Act, 2023," 2023.