

Explicit Discontinuance and Switching Signals in Generative AI Applications: A Comparative Analysis Across Nine Apps

Veysel Şahin¹

Okan Alkan²

Yunus Emre Özenoğlu³

Abstract

This study examines explicitly reported discontinuance and platform-switching signals in Turkish- and English-language Google Play reviews of nine generative artificial intelligence applications. A total of 1,145,267 reviews posted between 1 July 2025 and 30 June 2026 were screened using rule-based procedures. Of 11,960 candidate reviews, 821 that met strict contextual rules were evaluated in two blind AI-assisted coding passes. The final high-confidence analysis set comprised 692 records, including 543 signals in which discontinuance or switching was reported as having occurred and 149 signals expressing discontinuance intention. To assess the recall limitation of the strict screen, an additional application-by-language stratified sample of 500 reviews was audited from the broader pool of 11,139 records, revealing 110 missed explicit signals. The stratified weighted estimate of the proportion of missed explicit signals in the broad pool was 21.6% (95% CI: 18.2%-25.2%). Accordingly, rates reported per 10,000 reviews are interpreted as high-confidence detection indicators rather than prevalence estimates. Within the high-confidence set, the most frequent reasons were response quality/task failure, technical problems, privacy/trust,

- 1 PhD Student, Department of Management Information Systems, Graduate School of Social Sciences, Atatürk University, Erzurum, Türkiye. E-mail: veysel.sahin18@ogr.atauni.edu.tr. ORCID: 0000-0002-0827-9478
- 2 Lecturer Dr., Department of Computer Technologies, Vocational School, Erzincan Binali Yıldırım University, Erzincan, Türkiye. E-mail: oalkan@erzincan.edu.tr. ORCID: 0000-0002-6526-1645
- 3 PhD Student, Department of Management Information Systems, Graduate School of Social Sciences, Atatürk University, Erzurum, Türkiye. E-mail: yunusemre.ozenoglu12@ogr.atauni.edu.tr. ORCID: 0000-0001-5051-4622

and pricing. Among platform switches reported as having occurred, Gemini and Claude were the most frequent destinations. By separating intention from signals of reported enacted behavior and by explicitly reporting screening sensitivity, this study provides complementary evidence to the post-adoption literature on generative AI.

1. Introduction

Generative artificial intelligence (GenAI) refers to computational systems capable of generating new content in response to user inputs and based on a range of generative architectures, particularly large language models (Feuerriegel et al., 2024). Following the rapid diffusion of ChatGPT after its release in late 2022, a broader ecosystem of consumer-facing GenAI services has become increasingly visible. These systems are used for information access, content generation, problem solving, and decision support, and they are becoming increasingly prominent not only among individual users but also in educational and work settings (Feuerriegel et al., 2024). This diffusion has shifted research attention beyond initial GenAI adoption toward users' post-adoption experiences and decisions about whether to continue using these technologies.

The information systems literature treats initial technology adoption and continued use over time as distinct behavioral stages. The long-term success of an information system depends not only on initial acceptance but also on whether users continue to use it (Bhattacharjee, 2001; Xiong et al., 2026). Under the Expectation-Confirmation Model, continuance is explained by the extent to which initial expectations are confirmed after use, perceived usefulness, and user satisfaction (Bhattacharjee, 2001). Recent research in the GenAI context likewise indicates that response accuracy, interaction quality, reliability, and perceived utility or affordances are relevant to continuance intentions (Khan et al., 2026; Xiong et al., 2026). Post-adoption behavior, however, is not limited to continued use: users may reduce their frequency of use, take temporary breaks, discontinue a service altogether, or move to an alternative platform.

Against this background, interruption or discontinuance of GenAI use has become an increasingly important research topic. Existing studies identify misinformation and hallucination risks (Feuerriegel et al., 2024; Zhao & Liu, 2026), algorithmic bias and low transparency (Zhou & Wang, 2025), privacy and data-security concerns (Zhou & Zhang, 2025), and cognitive fatigue (Zhao & Liu, 2026) as negative experiences that may be associated with discontinuance intention. Response accuracy and task performance are also critical to the continuation of GenAI use (Xiong et al., 2026). Nevertheless,

a substantial part of the literature focuses on discontinuance intention, that is, a user's future inclination to stop using a service (Zhao & Liu, 2026; Zhou & Wang, 2025). Statements such as "I uninstalled the app," "I cancelled my subscription," or "I no longer use it" are conceptually different because the user explicitly reports that the behavior has already occurred. App-store reviews provide an important naturalistic source for examining such post-adoption experiences in users' unsolicited feedback (Pagano & Maalej, 2013).

Moving away from one GenAI service does not necessarily imply abandoning the technology category as a whole. Users may leave a service with which they are dissatisfied and move to another GenAI platform. Platform switching has been examined through the Push-Pull-Mooring (PPM) framework. In the GenAI context, information hallucination, privacy risk, and dissatisfaction have been treated as push factors that drive users away from the current platform; better interaction, personalization, and user experience on an alternative platform have been treated as pull factors; and switching costs have been conceptualized as a mooring factor that may constrain switching (Zhou & Li, 2025). This framework is particularly useful for explaining the conditions under which platform-switching intention may arise. Consequently, research on post-adoption behavior in GenAI should consider not only why users leave an application but also whether they subsequently move to another platform.

App-store reviews allow users to describe their experiences without being constrained by researcher-defined survey items (Pagano & Maalej, 2013), enabling simultaneous examination of dimensions such as response quality, technical problems, usability, pricing, and service experience. Related themes are also visible in research conducted in Türkiye. Kılınç and Ayaz (2025), using a large-scale analysis of Google Play reviews, identified prominent themes in ChatGPT user experience including technical problems, feature requests, service quality, functionality, and academic use. Durmuş et al. (2025) compared user evaluations of ChatGPT, Gemini, Copilot, and DeepSeek and showed that ease of use, interface use, content satisfaction, information security, and privacy are salient dimensions of the GenAI experience. Together, these studies indicate that application performance should be evaluated not only through model capability but through the broader combination of technical functioning, trust, accessibility, and user experience.

A substantial proportion of the international literature on GenAI continuance, discontinuance, and platform switching relies on survey data and behavioral-intention measures (Khan et al., 2026; Zhou & Li, 2025; Zhou & Wang, 2025). In contrast, comparatively little research has distinguished intention statements from spontaneous app-store comments in

which users explicitly report that discontinuance or platform switching has already occurred, particularly in comparative analyses across multiple GenAI platforms. To address this gap, the present study examines Turkish- and English-language Google Play reviews of ChatGPT, Gemini, Claude, Microsoft Copilot, DeepSeek, Grok, Perplexity, Meta AI, and Kimi. It investigates the frequency with which explicit discontinuance and switching signals are detected, the reasons accompanying discontinuance or switching, whether these reasons differ across applications and review languages, and which alternative applications are named in platform switches reported as having occurred. Review language is not treated as an indicator of users' nationality, country of residence, or cultural identity. By focusing on explicit signals in which users report enacted discontinuance and switching, rather than only on continuance or discontinuance intention, the study aims to provide complementary empirical evidence on post-adoption behavior in GenAI.

2. Theoretical Framework and Research Questions

2.1. Post-Adoption Behavior and Technology Discontinuance

The information systems literature distinguishes initial technology acceptance/adoption from post-adoption continuance over time (Bhattacharjee, 2001; Limayem et al., 2007; Xiong et al., 2026). The long-term success of an information system depends not only on the initial decision to use it but also on users' continued use after experience with the system (Bhattacharjee, 2001; Khan et al., 2026). In the Expectation-Confirmation Model (ECM), continuance intention is explained by the degree to which initial expectations are confirmed after use, perceived usefulness, and user satisfaction (Bhattacharjee, 2001). Recent GenAI research similarly shows that post-use evaluations, perceived system value or affordances, and the broader user experience are important to continuance intention (Khan et al., 2026; Xiong et al., 2026).

As GenAI applications have proliferated, negative post-adoption behaviors and the interruption or discontinuance of use have also become important research topics (Zhao & Liu, 2026; Zhou & Wang, 2025). Discontinuance is not a uniform process. It may take the form of reduced use, temporary or intermittent discontinuance, complete cessation, or switching to an alternative platform (Zhao & Liu, 2026; Zhou & Li, 2025; Zhou & Wang, 2025; Zhou & Zhang, 2025). Post-adoption behavior therefore encompasses not only whether use continues but also how the user's relationship with the current service changes over time.

Many factors associated with non-continuance of GenAI use arise from negative experiences with the system. Prior research has highlighted

misinformation and hallucination (Feuerriegel et al., 2024; Zhao & Liu, 2026; Zhou & Zhang, 2025), privacy and data-security concerns (Durmuş et al., 2025; Zhao & Liu, 2026; Zhou & Zhang, 2025), algorithmic bias and low transparency (Zhou & Wang, 2025), and problems related to response accuracy and task performance (Xiong et al., 2026). Studies of user reviews have also identified technical functioning, service/functionality problems, and pricing concerns as salient negative aspects of post-use experience (Kılınc & Ayaz, 2025; Metin, 2025). In addition, cognitive fatigue associated with risk perception and the ongoing need to evaluate generated information has been linked to intention to discontinue GenAI use (Zhao & Liu, 2026).

At this point, it is important to distinguish discontinuance intention from discontinuance that users explicitly report as having occurred. Discontinuance intention refers to a user's inclination or plan to stop using a system in the future, whereas statements such as "I uninstalled the app," "I cancelled my subscription," or "I no longer use it" are explicit signals that the user reports the behavior as already enacted. Much of the existing GenAI literature relies on survey-based measures of intention (Zhao & Liu, 2026; Zhou & Wang, 2025; Zhou & Zhang, 2025). App-store reviews, by contrast, are a naturalistic source of feedback in which users describe their experiences independently of researcher-defined question formats (Pagano & Maalej, 2013). Such reviews do not verify behavior at the level of system logs; rather, they enable analysis of explicit user reports of discontinuance and other post-adoption changes.

2.2. Platform Switching Behavior

Discontinuing a GenAI application does not necessarily mean that the user abandons GenAI altogether. When dissatisfied with an existing service, users may move to another GenAI platform. This process is explained in the platform-switching literature through the Push-Pull-Mooring (PPM) framework (Zhou & Li, 2025). In the GenAI context, information hallucination, privacy risk, and dissatisfaction are treated as push factors that move users away from the incumbent platform. Better interactivity, personalization, and user experience offered by alternative platforms are among the pull factors, whereas switching costs may operate as a mooring factor that constrains transition (Zhou & Li, 2025).

In the present study, the PPM framework is not used as a structural model for testing causal relationships. Instead, it serves as a theoretical lens for interpreting why users move away from one GenAI service and toward another. Much of the existing GenAI literature examines continuance, discontinuance, or switching intentions through survey data (Khan et al., 2026; Xiong et al.,

2026; Zhou & Li, 2025). The present study complements this literature by distinguishing intention statements from spontaneous Google Play reviews in which users explicitly report that discontinuance or platform switching has already occurred.

2.3. Research Questions

Based on the theoretical framework outlined above, the following research questions were formulated to examine explicitly reported discontinuance and platform-switching signals in Turkish- and English-language Google Play reviews of nine GenAI applications:

RQ1. How frequently are explicit discontinuance and platform-switching signals detected across the nine GenAI applications?

RQ2. What are the principal reasons accompanying discontinuance or switching of GenAI applications?

RQ3. Do the reasons for discontinuance and platform switching differ across the GenAI applications examined?

RQ4. Do the reasons for discontinuance and platform switching differ between Turkish- and English-language reviews?

RQ5. Which alternative GenAI applications are named as destinations in platform switches explicitly reported as having occurred?

3. Method

3.1. Research Design and Data Source

This study is an observational secondary-data analysis of explicitly reported discontinuance and platform-switching signals in GenAI applications. The data source comprised Turkish- and English-language Google Play Store reviews of ChatGPT, Gemini, Claude, Microsoft Copilot, DeepSeek, Grok, Perplexity, Meta AI, and Kimi. The study period was pre-specified and locked as 1 July 2025 to 30 June 2026, and July 2026 data were excluded to avoid a partial-month effect. Records retrieved more than once under the same source-review identifier were deduplicated, and review language was classified from the review text. The final corpus contained 1,145,267 reviews: 88,026 in Turkish and 1,057,241 in English. Usernames and fields that could directly identify individuals were not included in the analysis dataset.

3.2. Candidate Identification and AI-Assisted Coding

A negative evaluation alone was not treated as a discontinuance signal. Only explicit statements indicating actions such as uninstalling the application, ceasing use, cancelling or not renewing a subscription, or switching to another AI platform were considered. A broad rule-based screen using Turkish and English behavioral markers identified 11,960 candidate reviews; 821 reviews satisfying stricter contextual rules were then placed in the high-priority review pool. These 821 reviews were evaluated using a locked codebook in two separate blind coding passes with OpenAI ChatGPT (GPT-5.6 Sol), with the second pass unable to see labels from the first. Raw agreement on status between the two passes was 93.06%. In each pass, status was coded as “reported enacted behavior,” “intention,” “false positive,” or “uncertain”; where applicable, behavior type, target platform, and reasons were coded in a multi-label format. Disagreements were contextually adjudicated using the same decision rules. The 93.06% value is therefore a repeatability/consistency indicator for the same AI protocol, not an estimate of independent human validation or inter-coder reliability.

3.3. Broad-Pool Sensitivity Audit

To assess the recall limitation of the strict screen, a random sensitivity sample of 500 records was drawn from the broader pool of 11,139 reviews that did not meet the strict contextual rules, stratified by application × review language. Sampling was approximately proportional, with at least two records selected from each non-empty stratum, using the fixed seed 20260905. Independently of the main labels, these records were re-evaluated in a separate AI-assisted audit pass using the same decision rules for explicit behavior. The audit identified 110 high-confidence explicit signals missed by the strict screen: 62 signals in which the behavior was reported as having occurred and 48 intention signals. After stratified weighting, the estimated proportion of missed explicit signals within the broad pool was 21.6%, with a stratified-bootstrap 95% confidence interval of 18.2%-25.2%. This result indicates that the strict screen produced a high-confidence core through conservative selection but was not exhaustive. The 692-record main set should therefore be interpreted as a high-confidence detection set rather than as the prevalence of all discontinuance/switching events. Because the audit was also conducted using the same AI system, it is not interpreted as independent human validation.

3.4. Development of Reason Codes and Quality Control

Reasons for discontinuance/switching were initially coded using predefined categories covering response quality, hallucination, post-update degradation,

usage limits, pricing, restrictions, technical problems, feature loss, privacy/trust, payment/support, competitor superiority, and user experience. The 240 reviews initially assigned to an “other” category were then subjected to open subcoding. In addition to records reassigned to existing categories, five data-driven categories were created: ethical/political/institutional objection; environmental/resource-use concern; language/localization mismatch; perceived general societal or human harm; and ecosystem redundancy/service surplus. During quality control, 17 records were reclassified as false positives and one behavior code was corrected. The resulting high-confidence main analysis set comprised 692 reviews: 543 signals of discontinuance/switching reported as having occurred and 149 intention signals. Twelve reviews with no stated reason were flagged as K0 and excluded from reason-based comparisons.

3.5. Statistical Analysis

Analyses included frequencies and percentages, and the enacted and intention signals detected by the strict screen were descriptively standardized per 10,000 eligible reviews. Because the sensitivity audit indicated differences in capture probabilities across applications and languages, these rates were not treated as prevalence estimates, and no application- or language-level tests of behavioral rates were conducted. Comparisons of reasons were exploratory and conditional on the 692-record high-confidence set. Pearson’s chi-square test and Cramér’s V were used for categorical comparisons, and Fisher’s exact test was used for sparse 2×2 tables. Application $\times$ reason analyses were restricted to ChatGPT, Gemini, Claude, Grok, and Perplexity, each of which had at least 30 analysis-ready records, and Holm correction was applied across tests K1-K17. Platform switches were examined separately for reported enacted signals and intentions and were analyzed descriptively only; when a review contained multiple destinations, each destination was counted as a separate migration event. All analyses were observational and no causal inference was made.

4. Results

This section presents the high-confidence signals identified through strict screening and AI-assisted contextual coding in 1,145,267 Google Play reviews posted between 1 July 2025 and 30 June 2026. The main analysis set comprised 692 reviews, including 543 discontinuance/switching signals reported as having occurred and 149 intention signals. Because the broad-pool sensitivity audit showed that this set was high-confidence but incomplete, all rates reported below per 10,000 reviews should be interpreted only as descriptive indicators of detected signals.

4.1. Sensitivity Audit and Distribution of Detected Signals

In the stratified audit sample of 500 reviews drawn from the broad pool, 110 missed explicit signals were identified (22.0% of the audit sample): 62 signals reported as having occurred and 48 intention signals. The stratified weighted estimate of the proportion of missed explicit signals was 21.6% (95% CI: 18.2%-25.2%). This finding shows that the strict screen prioritized precision but had limited recall. Accordingly, application- and language-specific rates below should be read as high-confidence detection indicators generated under the same strict protocol, not as estimates of actual discontinuance prevalence or as rankings of platform performance.

The high-confidence signals detected for each application (reported enacted discontinuance/switching and intention) are presented in Table 1.

Table 1. High-confidence discontinuance/switching signals by application

Application	Corpus (n)	Detected reported enacted (n)	Detected reported enacted /10,000	Detected intention (n)	Detected intention /10,000
ChatGPT	714,116	298	4.17	74	1.04
Gemini	181,914	69	3.79	22	1.21
Claude	30,565	32	10.47	9	2.94
Copilot	35,004	19	5.43	1	0.29
DeepSeek	10,105	5	4.95	2	1.98
Grok	94,292	70	7.42	30	3.18
Perplexity	42,290	39	9.22	9	2.13
Meta AI	33,230	7	2.11	1	0.30
Kimi	3,751	4	10.66	1	2.67
Total	1,145,267	543	4.74	149	1.30

Note. Rates refer to high-confidence signals detected by the strict screen; they are not prevalence estimates. For Kimi, the rate of reported enacted discontinuance/switching signals (10.66 per 10,000 reviews) is based on only four detected signals among 3,751 reviews and should therefore be interpreted with particular caution.

As shown in Table 1, the strict-screen rate of signals reported as having occurred was 10.47 per 10,000 reviews for Claude, 9.22 for Perplexity, and 7.42 for Grok. The corresponding value for Kimi was 10.66 per 10,000; however, it was based on only four records and should therefore be interpreted with particular caution. Figure 1 visualizes this descriptive distribution across applications. Because the sensitivity audit indicated that missed signals were not distributed equally across applications, these values should not be interpreted as genuine platform differences or as a failure ranking.

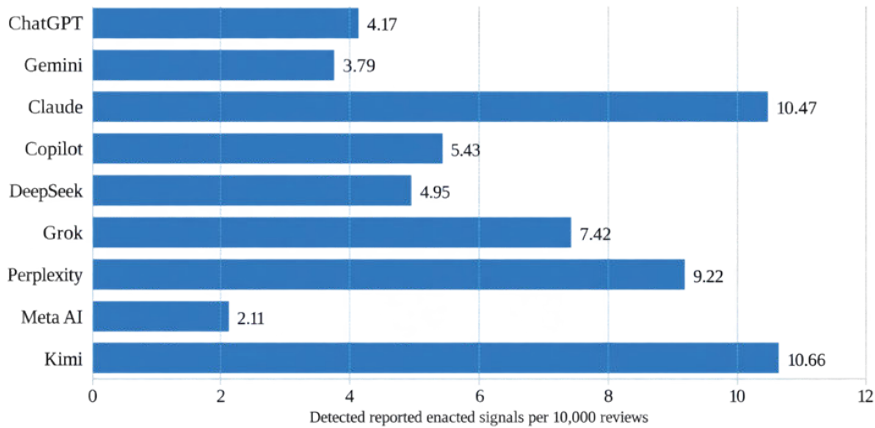


Figure 1. High-confidence rate of signals reported as having occurred by application

Note. Kimi's rate (10.66 per 10,000 reviews) is based on only four detected signals of reported enacted discontinuance/switching among 3,751 reviews and should therefore be interpreted with particular caution. Given this small event count, interpreting the value as evidence of higher discontinuance/switching prevalence for Kimi could be misleading. All rates are descriptive indicators of signals captured by the strict screen, rather than prevalence estimates or rankings of platform performance.

Rates detected by the strict screen according to review language are presented in Table 2.

Table 2. High-confidence discontinuance/switching signals by review language

Review language	Corpus (n)	Detected reported enacted (n)	Detected reported enacted /10,000	Detected intention (n)	Detected intention /10,000
Turkish	88,026	98	11.13	19	2.16
English	1,057,241	445	4.21	130	1.23
Total	1,145,267	543	4.74	149	1.30

Note. The language variable refers to the language of the review; no inference was made about users' nationality or country of residence.

As shown in Table 2, the strict-screen rate of signals reported as having occurred was 11.13 per 10,000 Turkish-language reviews and 4.21 per 10,000 English-language reviews. However, the sensitivity audit suggested that the proportion of missed signals in the broad pool was higher in Turkish-language strata than in English-language strata. The raw language difference was therefore retained only as a descriptive result of the screening protocol and was not interpreted as a behavioral, national, or cultural difference.

4.2. Reasons for Discontinuance and Platform Switching

Reasons for discontinuance/switching were coded in a multi-label format within the 692 high-confidence reviews; consequently, a single review could contain more than one reason. The final distribution of reasons is presented in Table 3, and the eight most frequent reasons are shown in Figure 2.

Table 3. Distribution of reasons for discontinuance and platform switching

Code	Reason	n	Share (%)
K1	Response quality / task failure	205	29.6
K7	Technical / access / device compatibility issue	119	17.2
K9	Privacy / security / data control / trust	94	13.6
K5	Price / subscription / paywall / perceived value	82	11.8
K3	Quality degradation after update / model change	74	10.7
K2	Hallucination / misinformation / incorrect guidance	71	10.3
K10	Payment / refund / billing / support	71	10.3
K12	User experience / interface / user control	64	9.2
K4	Usage limit / quota / capacity	52	7.5
K11	Competitor perceived as better	48	6.9
K6	Excessive restriction / refusal / content policy	40	5.8
K8	Loss of feature / model / history	28	4.0
K13	Ethical / political / institutional objection	18	2.6
K14	Environmental / resource-use concern	11	1.6
K15	Language / localization mismatch	7	1.0
K16	Perceived general societal / human harm or AI risk	6	0.9
K17	Ecosystem fit / redundancy / service surplus	3	0.4
K0*	No reason specified	12	1.7

Note. Percentages were calculated over the 692 high-confidence reviews and may sum to more than 100% because of multi-label coding. K0 is not a reason for discontinuance/switching. Additional signals detected in the sensitivity audit were not included in reason coding; therefore, this distribution is conditional on the high-confidence main set.

As shown in Table 3 and Figure 2, the most common reason in the high-confidence set was response quality/task failure (K1; $n = 205$; 29.6%). This was followed by technical/access/device-compatibility problems (K7; 17.2%), privacy and trust (K9; 13.6%), price/subscription/paywall and perceived value (K5; 11.8%), and quality degradation after an update or model change (K3; 10.7%). The distribution indicates that discontinuance/switching signals cannot be reduced to a single source of dissatisfaction; performance, access, trust, and economic conditions were all visible in the reviews.

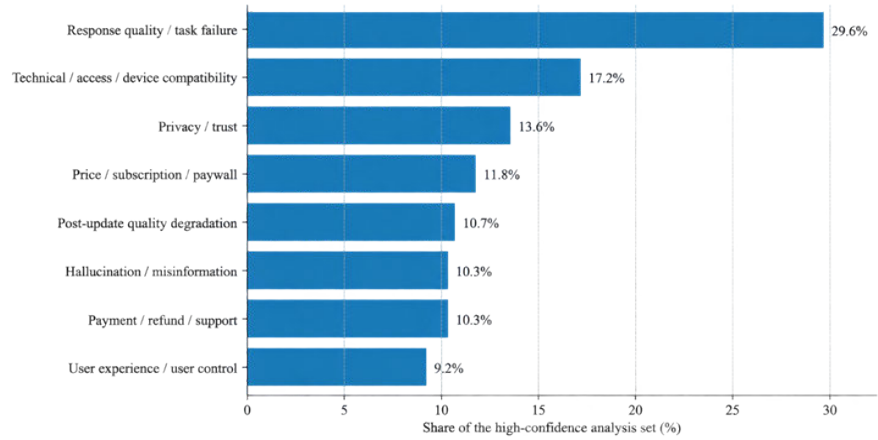


Figure 2. Most frequently detected reasons for discontinuance/platform switching in the high-confidence analysis set

4.3. Differences in Reasons by Application and Review Language

Direct reason-level comparisons across all nine applications produced sparse cells in several categories. The exploratory inferential analysis was therefore restricted to ChatGPT, Gemini, Claude, Grok, and Perplexity, each of which had at least 30 high-confidence records, and Holm correction was applied across the 17 reason tests. These tests do not represent the full user populations; they describe conditional differences within the 692-record high-confidence detection set. Results that remained significant after Holm correction are shown in Table 4.

Table 4. Reasons showing significant differences by application and review language

Comparison	Reason	Most prominent distribution	Holm-adjusted p	Cramér's V
Application	K4–Usage limit/quota	Claude 26.8%; Grok 13.0%	<0.001	0.223
Application	K5–Price/subscription/paywall	Grok 32.0%; Perplexity 18.8%	<0.001	0.270
Application	K9–Privacy/trust	ChatGPT 18.5%; Gemini 15.4%	0.026	0.162
Application	K12–UX/user control	Perplexity 20.8%; Gemini 18.7%	<0.001	0.212
Language	K9–Privacy/trust	Turkish 25.6%; English 11.1%	<0.001	0.159

Comparison	Reason	Most prominent distribution	Holm-adjusted p	Cramér's V
Language	K14–Environmental/resource use	Turkish 6.8%; English 0.5%	<0.001	0.189

Note. Results are conditional/exploratory comparisons within the 692-record high-confidence set. Application analyses are limited to ChatGPT, Gemini, Claude, Grok, and Perplexity. Fisher's exact test was used for sparse 2 × 2 language tables, and p values were adjusted using the Holm method.

As shown in Table 4, usage-limit/quota problems (K4) were more prominent for Claude (26.8%), price/subscription/paywall reasons (K5) for Grok (32.0%), privacy/trust reasons (K9) for ChatGPT (18.5%), and user-experience/user-control problems (K12) for Perplexity (20.8%) and Gemini (18.7%) within the high-confidence set. By review language, privacy/trust (K9) appeared in 25.6% of Turkish-language reviews and 11.1% of English-language reviews, whereas environmental/resource-use concerns (K14) appeared in 6.8% and 0.5%, respectively. These differences should be interpreted only within the high-confidence detection set; review language must not be used to infer nationality, residence, or cultural identity.

4.4. Directions of Platform Switching

An explicit signal of platform switching reported as having occurred was identified in 40 reviews. Because some reviews named more than one destination, these reviews yielded 43 enacted migration events. The distribution of destination platforms is presented in Table 5.

Table 5. Destination distribution of platform migrations reported as having occurred

Destination platform	Migration event (n)	Share (%)
Gemini	15	34.9
Claude	10	23.3
ChatGPT	8	18.6
Other AI	5	11.6
Grok	3	7.0
Copilot	1	2.3
Perplexity	1	2.3
Total	43	100.0

Note. The table reports migration events rather than reviews. When a review contained more than one destination platform, each destination was counted as a separate event.

As shown in Table 5, Gemini was the most frequent destination among migrations reported as having occurred ($n = 15$; 34.9%), followed by Claude ($n = 10$; 23.3%) and ChatGPT ($n = 8$; 18.6%). The most frequent source-destination paths were ChatGPT→Gemini ($n = 12$) and ChatGPT→Claude ($n = 9$). In addition, 41 reviews contained platform-switching intentions that had not yet been enacted, producing 42 destination events. Given the small number of migration events and the unequal review volumes across applications, these findings should be interpreted as explicit directional signals in the reviewed comments rather than as indicators of market share, user-base size, or actual market-wide migration.

5. Discussion

The high-confidence analysis set showed that moving away from a GenAI application or toward another platform cannot be attributed to a single factor. Response/task quality, technical problems, privacy and trust, pricing, usage limits, and user experience were all visible. The prominence of response quality and task failure is consistent with research emphasizing the role of response accuracy, reliability, misinformation, and dissatisfaction in post-adoption evaluations of GenAI (Xiong et al., 2026; Zhou & Wang, 2025). The visibility of technical problems, service/functionality concerns, and pricing is also consistent with user-review research by Kılınç and Ayaz (2025) and Metin (2025).

The prominence of privacy and trust among high-confidence discontinuance/switching signals is also consistent with prior research. Zhou and Zhang (2025) found that privacy concerns and information hallucination can increase cognitive dissonance, which in turn promotes intermittent discontinuance, while Zhao and Liu (2026) linked risk perception to discontinuance intention through cognitive fatigue. Zhou and Wang (2026) further showed that algorithmic bias, misinformation, privacy concerns, and unexplainability can increase usage exhaustion, which promotes intermittent discontinuance. The greater visibility of privacy/trust reasons among Turkish-language reviews in the present high-confidence set is broadly consistent with the low evaluations of information security and privacy reported by Durmuş et al. (2025). Nevertheless, because the sensitivity audit indicated different capture rates across language strata, this result cannot be interpreted as a cultural or national effect.

Differences in the reason profiles across applications suggest that platform-specific service conditions may create different points of friction within the high-confidence detection set. Usage-limit/quota problems were more visible for Claude, price and subscription problems for Grok, privacy/trust

concerns for ChatGPT, and user-experience/control problems for Gemini and Perplexity. Zhou and Li (2025) conceptualize hallucination, privacy risk, and dissatisfaction as push factors and better user experience as a pull factor in GenAI switching. Demir and Zaimoğlu (2025) likewise reported subject-specific accuracy limitations and the high cost of premium features as notable disadvantages in ChatGPT use. The present application comparisons, however, are conditional on the high-confidence detection set and should not be generalized to all users of the platforms.

The prominence of Gemini and Claude among platform switches reported as having occurred, particularly in explicit reports of movement away from ChatGPT toward these platforms, provides concrete user-stated directional evidence within the app-store comments examined. Zhou and Li (2025) found dissatisfaction to be a common core condition triggering GenAI platform-switching intention. The present study complements that intention-based literature by separately identifying signals in which a switch was reported as already having occurred. However, only 43 enacted migration events were detected, and review volumes differed substantially across applications; these results therefore cannot be interpreted as evidence of general market migration or platform failure.

The emergence of environmental and resource-use concerns as a distinct reason through open coding is noteworthy. Although numerically uncommon, this theme suggests that awareness of the energy and resource demands of GenAI can enter users' post-adoption evaluations. Klesel and Messer (2026) found that users were aware of GenAI-related energy consumption and proposed a range of pro-environmental strategies, including reducing use and optimizing prompts. Environmental concern may therefore represent a useful dimension for future post-adoption research to test explicitly.

The sensitivity audit also makes the study's interpretive boundary explicit. Finding 110 additional explicit signals in the 500-review broad-pool audit demonstrates that the 692-record main set is high-confidence but incomplete. Accordingly, the study does not estimate the prevalence of discontinuance or switching; it examines the reason and direction patterns among signals captured by the strict screen. The use of the same AI system for both main coding and the sensitivity audit does not provide independent validation. Generalizability is further limited by the self-selected nature of app-store reviews, the exclusive use of Google Play and Turkish/English reviews, and the possibility that screening sensitivity varies by application and language. At the same time, the explicit separation of reported enacted signals from intentions, the large

corpus, the multi-platform scope, and transparent reporting of the precision-recall trade-off constitute the study's principal methodological contributions.

6. Conclusion

This study examined explicitly reported discontinuance and platform-switching signals in Turkish- and English-language Google Play reviews of nine GenAI applications. Strict screening followed by two blind AI-assisted coding passes produced a high-confidence analysis set of 692 records, comprising 543 discontinuance/switching signals reported as having occurred and 149 intention signals. In addition, the broad-pool sensitivity audit of 500 reviews identified 110 missed explicit signals, and the stratified weighted estimate of the proportion of missed explicit signals was 21.6%. These findings demonstrate that the main set should be treated as a high-confidence detection set rather than as an estimate of the prevalence of all discontinuance/switching events.

Within the high-confidence set, the leading reasons for discontinuance/switching were response quality and task failure, technical problems, privacy/trust, pricing, and post-update quality degradation. Reason profiles differed across applications, while privacy/trust and environmental/resource-use concerns were more visible in Turkish-language reviews. However, because the sensitivity audit indicated different capture probabilities across application and language strata, these differences must not be interpreted as behavioral prevalence or as differences attributable to nationality, residence, or culture.

Platform-switching reports showed that some users explicitly described moving to alternative GenAI services rather than terminating GenAI use altogether after a negative experience. Gemini and Claude were the most frequent destinations among switches reported as having occurred, and the most frequent explicit paths were ChatGPT→Gemini and ChatGPT→Claude. Because the number of events was limited and review volumes varied markedly across applications, these patterns should be interpreted as directional signals explicitly reported in the reviewed comments, not as market share or general user migration.

The study's main contribution lies in examining, separately from discontinuance and switching intentions, signals in which users spontaneously report that the behavior has already occurred, while also making the recall limitation of the strict screen visible through a stratified sensitivity audit. Because the same AI system was used for both main coding and the audit, this procedure does not constitute independent human validation, and that limitation is explicitly retained. Future studies should consider full coding of the broader candidate pool using human-assisted or multi-model validation,

comparisons across additional app stores and languages, and longitudinal tracking of platform switching.

7. Ethics and Data Use Statement

The study is based solely on secondary analysis of publicly available Google Play reviews. Usernames and directly identifying information were not included in the analysis dataset, and no verbatim user quotations are reproduced in this chapter. Findings are reported only in aggregate form, and review language was not used as an indicator of nationality, residence, or cultural identity.

8. Artificial Intelligence Use Statement

OpenAI ChatGPT (GPT-5.6 Sol) was used as an assistive tool for two blind AI coding passes under a locked codebook, contextual adjudication, the broad-pool sensitivity audit, and academic-language/formatting refinement. AI outputs are not presented as independent human-coder validation. The authors retain scientific responsibility for analytical decisions, statistical reporting, and the final manuscript.

References

- Bhattacharjee, A. (2001). Understanding information systems continuance: An expectation-confirmation model. *MIS Quarterly*, 25(3), 351–370. <https://doi.org/10.2307/3250921>
- Demir, A. N., & Zaimoğlu, S. (2025). Enhancing or deteriorating? The influence of ChatGPT on university students' academic writing skills in Türkiye. *Bayburt Eğitim Fakültesi Dergisi*, 20(46), 778–800. <https://doi.org/10.35675/befdergi.1647707>
- Durmuş, A., Tozkoparan, S. B., & Çekiç, S. (2025). Üretken yapay zekâ kullanımına ilişkin öğrenci görüşleri - chatgpt, gemini, copilot ve deepseek örneği. *Uluslararası Bilişim Sistemleri ve Uygulamaları Dergisi*, 1(2), 105–118. <https://izlik.org/JA82MY75WY>
- Feuerriegel, S., Hartmann, J., Janiesch, C., & Zschech, P. (2024). Generative AI. *Business & Information Systems Engineering*, 66(1), 111–126. <https://doi.org/10.1007/s12599-023-00834-7>
- Khan, A. R., Mahmood, S. M. F., Akbar, M., & Hossain, S. (2026). Understanding generative AI continuance intention: A dual-factor perspective on facilitators and barriers. *The Electronic Library*, 44(4), 856–875. <https://doi.org/10.1108/EL-07-2025-0328>
- Klesel, M., & Messer, U. (2026). Public perceptions of AI energy consumption: Awareness, responsibility attribution, and strategies. *Electronic Markets*, 36, Article 59. <https://doi.org/10.1007/s12525-026-00908-w>
- Kılınç, M., & Ayaz, A. (2025). Exploring user perceptions of ChatGPT: A large-scale topic modeling analysis using LDA. *Academic Platform Journal of Engineering and Smart Systems*, 13(3), 120–130. <https://doi.org/10.21541/apjess.1677061>
- Limayem, M., Hirt, S. G., & Cheung, C. M. K. (2007). How habit limits the predictive power of intention: The case of information systems continuance. *MIS Quarterly*, 31(4), 705–737. <https://doi.org/10.2307/25148817>
- Metin, S. (2025). Üretken yapay zeka modellerinin teknik özellikleri ve kullanıcı geri bildirimlerinin analizi: chatgpt örneği. *İnönü Üniversitesi Uluslararası Sosyal Bilimler Dergisi*, 14(2), 744–770. <https://doi.org/10.54282/inijoss.1763596>
- Pagano, D., & Maalej, W. (2013). User feedback in the AppStore: An empirical study. *2013 21st IEEE international requirements engineering conference (RE)* (pp. 125–134). IEEE. <https://doi.org/10.1109/RE.2013.6636712>
- Xiong, L., Zhou, Y., & Hua, Y. (2026). Why do users continue to use GCAI? A mixed-methods study. *Technology in Society*, 86, 103309. <https://doi.org/10.1016/j.techsoc.2026.103309>
- Zhao, W., & Liu, J. (2026). Risk perception and generative AI discontinuance intention: Mediating effect of cognitive fatigue and moderating effect of AI

- dependence. *Frontiers in Psychology*, 17, 1879080. <https://doi.org/10.3389/fpsyg.2026.1879080>
- Zhou, T., & Wang, M. (2026). Understanding AI-generated content user intermittent discontinuance: A COR theory perspective. *The Electronic Library*, 44(4), 820–837. <https://doi.org/10.1108/EL-09-2025-0366>
- Zhou, T., & Li, S. (2025). Examining user switching intention between generative AI platforms: A push-pull-mooring perspective. *Information Development*, 41(3), 692–704. <https://doi.org/10.1177/026666669241306735>
- Zhou, T., & Wang, M. (2025). Examining generative AI user discontinuance from a dual perspective of enablers and inhibitors. *International Journal of Human–Computer Interaction*, 41(20), 13140–13150. <https://doi.org/10.1080/10447318.2025.2470280>
- Zhou, T., & Zhang, C. (2025). Examining generative AI user intermittent discontinuance from a C-A-C perspective. *International Journal of Human–Computer Interaction*, 41(10), 6377–6387. <https://doi.org/10.1080/10447318.2024.2376370>

