The University of Texas at Austin
Center for Identity

A Longitudinal Look at GDPR Compliance

Brian Kim

Trista Cao

K. Suzanne Barber

UTCID Report #25-08

August 2025

A Longitudinal Look at GDPR Compliance

Brian Kim
The University of Texas at Austin
The Center for Identity
Austin, Texas, USA
bkim994@utexas.edu

Yang Trista Cao
The University of Texas at Austin
The Center for Identity
Austin, Texas, USA
ycao55@utmail.utexas.edu

K. Suzanne Barber
The University of Texas at Austin
The Center for Identity
Austin, Texas, USA
sbarber@identity.utexas.edu

***Abstract*—This paper presents a longitudinal study investigating how the General Data Protection Regulation (GDPR) compliance of website privacy policies has evolved over a five-year period. Using an automated privacy policy evaluation tool, we assessed ten core GDPR factors across a corpus of websites originally analyzed in 2020 and re-evaluated in 2025. Our analysis reveals a mixed progression: while user-facing compliance measures such as consent, data retention notification, and data sharing transparency showed measurable improvement, technically oriented factors—such as breach notification and data encryption—experienced a decline in explicit disclosure. These findings suggest a broader trend in which privacy policies increasingly emphasize legal rights and visible consent mechanisms, while de-emphasizing backend technical safeguards. The results point to a split in compliance communication, possibly influenced by regulatory clarity, enforcement pressure, and shifts in organizational privacy strategy. This study underscores the importance of continued policy auditing and the need for complementary methods that bridge the gap between stated policy and implemented practice in the context of evolving digital governance frameworks.**

I. Introduction

In an increasingly digital world, the protection of personal data has emerged as a fundamental concern for both individuals, corporations, government agencies, and regulators. The European Union's General Data Protection Regulation (GDPR), enacted in May 2018, represents one of the most comprehensive and influential data protection laws globally. Its primary objective is to give individuals greater control over their Personally Identifiable Information (PII) while imposing clear obligations on organizations that collect, process, and store such data [1]. With its extraterritorial reach and significant penalties for non-compliance, GDPR has shaped not only the legal landscape within the EU but also influenced global privacy standards and practices.

Privacy policies are central to GDPR compliance, serving as the primary mechanism by which organizations disclose their data practices and inform users of their rights. Since its adoption, a wide range of industries and organizations have attempted to align their privacy documentation with GDPR regulatory requirements, leading to changes in the structure, content, and clarity of privacy policies [2].

Despite the GDPR's clear mandates, questions remain regarding how effectively organizations have complied with GDPR over time. Initial studies at the Center for Identity at The University of Texas (UTCID) conducted shortly after the regulation took effect revealed gaps in adherence, inconsistencies in the presentation of user rights, and a lack of standardization across policies [1]. However, as organizations have had more time to adapt and enforcement actions have increased, it is critical to assess whether and how privacy policies have improved in response.

This study presented here investigates the evolution of GDPR compliance in privacy policies. By analyzing a longitudinal dataset of privacy policies using automated tools and compliance criteria derived from GDPR text and legal interpretations, we examine trends in policy completeness, clarity, and coverage of key data protection principles. This work contributes to a growing body of literature on regulatory impact assessment [3] and provides insights into the effectiveness of GDPR as a mechanism for driving organizational transparency and accountability in data practices.

II. Related Work

Research investigating privacy policies and regulatory compliance has grown considerably in recent years, particularly in light of the GDPR. This section reviews related literature from: (1) empirical studies examining privacy policy content, user understanding, and compliance trends, and (2) automated tools designed to evaluate, analyze, or interpret privacy policies at scale.

A. Empirical Studies

A wide range of academic studies have investigated how privacy policies align with user expectations, regulatory requirements, and industry practices. Tesfay et al. [4] analyzed privacy policies for GDPR compliance and found that many lacked explicit language regarding user rights and data retention, signaling a superficial implementation of core principles. Similarly, Zimmeck et al. [5] conducted large-scale audits of mobile app privacy policies, revealing persistent gaps in transparency and vague descriptions of data practices.

Other studies focused on user-centric concerns. Shvartzshnaider et al. [6] introduced the concept of Contextual Integrity annotations to compare normative expectations of privacy with actual policy language. Their findings highlighted significant mismatches between user privacy norms and how data handling is described in policies. Kang et al. [7] further emphasized these gaps by surveying user preferences and com-

paring them with policy content, demonstrating widespread dissatisfaction with data sharing and control mechanisms.

More recent studies have delved into the complexities of privacy policies, exploring their alignment with user expectations and regulatory requirements. A comprehensive scoping review by van der Schyff et al. [3] highlighted the diverse techniques employed in privacy policy analysis, ranging from manual content analysis to automated machine learning approaches.

B. Privacy Tools

To address the challenge of manual policy assessment, several automated tools have been developed to evaluate privacy policies for regulatory alignment, user-centric risks, and structural features.

Polisis [8] is one of the earliest tools that leveraged deep learning and a semantic taxonomy to segment and label privacy policy sentences. It offers a structured representation of policies and enables users to query policies using a natural language interface. PolicyLint [9] and PrivacyGuide [10] build on similar approaches, using machine learning and expert rule sets to flag incomplete or non-compliant disclosures, particularly with respect to GDPR mandates.

PolicyChecker [11] specializes in identifying GDPR completeness across mobile app privacy policies. By encoding GDPR requirements as a checklist, the tool automatically assesses whether key elements—such as data subject rights, lawful basis, and third-party sharing—are clearly disclosed. This approach is especially useful for regulators and auditors performing comparative reviews.

PrivacyCheck™ [12] provides a dual scoring system for GDPR compliance and User Control based on the Fair Information Practice Principles (FIPPs), relying on a rule-based model that interprets policy clauses across 20 key factors. The most recent iteration also allows dynamic reweighting based on user values, enabling personalized evaluations aligned with survey-informed preferences.

Together, these tools represent a significant advance in automating policy analysis, though each differs in scope, interpretability, and focus—ranging from regulatory compliance to user-centric risk scoring. Their development underlines the growing demand for scalable methods to interpret the increasingly complex landscape of digital privacy governance.

III. Approach

To assess how privacy policy compliance with the GDPR has evolved over time, this study compares past and present GDPR compliance scores using the PrivacyCheck™ tool developed at the UTCID [12]. The original PrivacyCheck™ dataset, compiled in 2020 by Zaeem et al. [1], provides baseline GDPR compliance scores for a corpus of privacy policies. By rerunning the PrivacyCheck™ browser extension on the same set of websites in 2025, this study captures a five-year window of policy evolution.

Although the dataset begins in 2020—two years after the GDPR came into effect in 2018—this buffer period offers some distinct advantages for the present study. By allowing organizations time to adapt their privacy policies to the initial regulatory requirements, the analysis can focus on more stable, long-term compliance trends rather than early-stage implementation efforts. As such, examining the evolution of GDPR compliance from 2020 to 2025 enables a clearer view of how policies have matured over time.

The original dataset consists of 392 privacy policies from companies randomly selected across the New York Stock Exchange (NYSE), Nasdaq, and American Stock Exchange (AMEX) stock exchanges [1]. PrivacyCheck™ evaluates each policy against ten specific GDPR-related compliance factors as shown in Table I. For each factor, a score of 0, 1, or 2 is assigned, where:

- **0** indicates the factor is not addressed
- **1** indicates the policy does not meet the factor's requirement
- **2** indicates full compliance with the factor

Each policy therefore receives a set of ten individual scores reflecting its degree of compliance across the evaluated dimensions. The original corpus of privacy policies was manually scored on these questions to generate training data for a machine learning model [1]. Once trained, the model was integrated into the PrivacyCheck™ browser extension, allowing users to automatically evaluate privacy policies [13]. The same scoring system is applied when analyzing the updated 2025 versions of these privacy policies using the latest iteration of PrivacyCheck™.

To collect the current versions of these privacy policies, automated web scraping methods were employed. Starting with the URLs used in the original 2020 dataset, each website was revisited to extract its most up-to-date privacy policy. In cases where the original URLs were outdated or broken, updated links were located by scraping the website's main pages or searching for redirects and updated routes to their privacy policy sections. This ensured that the 2025 analysis used the most accurate and current documents available for each company.

Once all current policy texts were retrieved and processed, they were evaluated using the same PrivacyCheck™ scoring methodology. These results were then aggregated alongside the original scores to enable a direct comparison of GDPR compliance evolution over time.

IV. Results

This section presents the findings of the study in two parts: URL Collection and Score Comparison. The first subsection outlines the outcome of collecting updated URLs from the original 2020 PrivacyCheck™ corpus, highlighting the extent to which website privacy policy links remained stable, changed, or became inaccessible over the past five years. The second subsection presents a comparative analysis of GDPR scores generated by PrivacyCheck™ for both the original and updated privacy policy links. This comparison illustrates how GDPR compliance in privacy policies has evolved across key factors over time.

TABLE I
PRIVACYCHECK™ GDPR QUESTIONS [13]

GDPR Topic	Question for Privacy Policy
(1) Between Site Transfer	Does this website share the user's information with other websites only upon user consent?
(2) Company Location	Does this website disclose where the company is based and where the user's information will be processed and/or transferred?
(3) Right to be Forgotten	Does this website support the right to be forgotten? That is, when requested, will the website delete all of the user's information?
(4) Data Retention Notification	If this website retains information for legal purposes after the user's request to be forgotten, will they inform the user?
(5) Reject Usage of PII	Does this website allow the user the ability to reject usage of user's PII?
(6) Under 16 Protection	Does this website restrict the use of PII of children under the age of 16?
(7) Data Encryption	Does this website advise the user that their data is encrypted even while at rest?
(8) Data Processing Consent	Does this website ask for the user's informed consent to perform data processing?
(9) Data Protection Principles	Does this website implement all of the principles of data protection by design and by default?
(10) Breach Notification	Does this website notify the user of security breaches without undue delay?

A. Collecting Privacy Policies

To evaluate the current state of GDPR compliance, it was necessary to re-collect the privacy policy URLs from the original PrivacyCheck™ corpus consisting of 392 websites. Each URL was revisited to determine whether the corresponding privacy policy could still be accessed and scored using the PrivacyCheck™ tool.

Over the span of five years, a significant number of these URLs experienced structural or ownership changes. Out of the original 392 links, 377 valid and scorable URLs were successfully recovered. Among these:

- 247 URLs (65.5%) remained unchanged and still pointed to the same privacy policies as in 2020
- 130 URLs (33.2%) redirected to new domains or were associated with companies that had since been acquired or rebranded
- 15 URLs (3.8%) could not be recovered as these domains were either permanently shut down or acquired by entirely unrelated entities, rendering them unsuitable for analysis

In total, 145 out of 392 websites (37.0%) experienced major transformations, either through domain redirection, acquisition, or decommissioning.

This level of link volatility highlights the dynamic nature of the online ecosystem, where corporate mergers, rebranding, and platform shutdowns are commonplace. The fact that over one-third of the original corpus experienced substantive changes over just five years suggests that longitudinal privacy policy research must account for domain drift and organizational turnover. Furthermore, the relatively high retention of original URLs (over 60%) provides a sufficiently stable subset to support meaningful temporal comparisons of GDPR compliance.

B. Longitudinal Analysis of GDPR Compliance

This section presents a longitudinal analysis of the evolution in GDPR compliance across ten distinct factors as assessed by PrivacyCheck™. To measure changes over time, the PrivacyCheck™ tool was rerun on the updated set of privacy policy URLs, and the resulting scores for each GDPR factor were averaged. The same aggregation was applied to the original 2020 scores, allowing for direct comparison across a five-year span. Figure 1 displays the average scores for each factor in 2020 and 2025 side by side, while Figure 2 visualizes the difference between them. Each factor's direction and magnitude of score change provide insight into shifts in GDPR adherence, likely influenced by elements such as evolving regulatory expectations, legal precedents, and public awareness. To further assess whether these score differences were meaningful, independent two-sample t-tests were conducted for each factor to determine whether the observed changes are statistically significant.

The factor with the largest increase in score was Between Site Transfer, which rose by approximately 0.18 points—a statistically significant change (p = 0.0093). This factor evaluates whether a website shares user data with other websites only upon explicit user consent. The substantial improvement suggests that organizations have become more diligent in obtaining informed consent before engaging in cross-site data transfers. This trend may coincide with the ramifications of the Schrems II decision in 2020, which invalidated the EU–US Privacy Shield and placed stricter conditions on cross-border data sharing under GDPR Article 44 [14]. While a direct causal link cannot be confirmed, the timing suggests that some organizations may have updated their consent practices in response to increased regulatory uncertainty.

A similarly notable and statistically significant increase (p = 0.0190) of approximately 0.17 points was observed in the Data

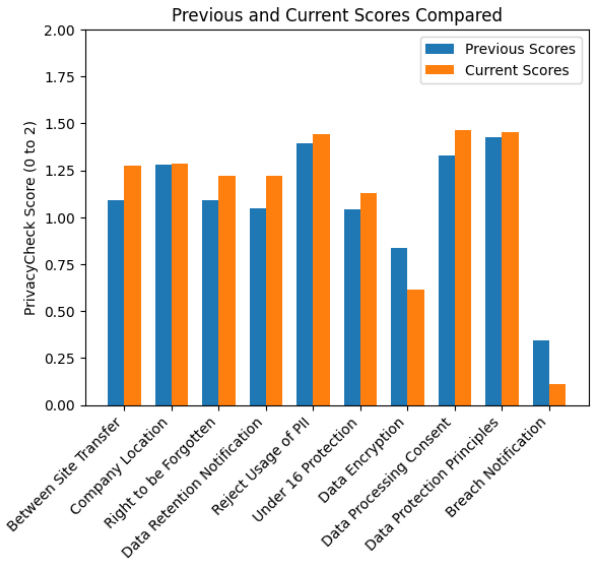

Fig. 1. Comparison of previous and current scores

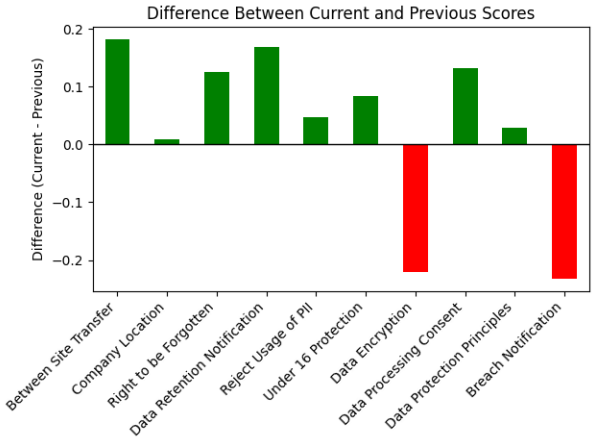

Fig. 2. Difference between previous and current scores

Retention Notification factor. This metric assesses whether websites inform users if their data will be retained for legal purposes following a deletion request. The observed increase may reflect growing transparency around data retention practices. This change occurred around the time the European Data Protection Board (EDPB) issued guidance documents in 2021–2022 [15]. These documents outlined the conditions under which continued data storage must be disclosed under Articles 13 and 17. Though a direct causal relationship cannot be established, it may indicate that the guidance could have influenced more explicit policy updates.

The Data Processing Consent factor saw an increase of approximately 0.13 points, but the change was not statistically significant (p = 0.1115). Still, this change may be indicating improved mechanisms for obtaining users' informed consent for data processing activities. This upward shift aligns with a period of increased enforcement actions against inadequate or manipulative consent practices, particularly those involving cookie banners and dark patterns. Notable regulatory actions from the French National Commission on Informatics and Liberty (CNIL) and legal challenges initiated by advocacy organizations such as None Of Your Business (NOYB) could have likely pressured companies into revising their consent flows to meet the standards of freely given, specific, informed, and unambiguous consent [16], [17]. Although this temporal alignment is very suggestive, further research would be needed to establish direct causality.

An increase of approximately 0.13 points was also seen in the Right to be Forgotten factor, which evaluates whether users are able to request deletion of their personal data. Although this change did not reach statistical significance (p = 0.0749), the upward movement suggests that a growing number of websites now support and acknowledge this right, possibly influenced by continued public awareness and discourse following high-profile legal cases such as Google Spain v. AEPD [18]. These cases have reinforced the importance of user erasure rights and could have led organizations to incorporate clearer deletion procedures into their policies to avoid reputational or legal consequences.

The Under 16 Protection factor increased by approximately 0.08 points (p = 0.3664), reflecting modest but statistically insignificant improvements in the protection of minors' data. This shift coincides with legislative developments such as the UK's Age-Appropriate Design Code (AADC) and the EU's Digital Services Act (DSA), both of which have emphasized stronger safeguards for children online [19], [20]. It is possible that websites have increasingly implemented age-verification procedures and added language restricting the processing of data for users under the age of 16 in response to these developments.

The Reject Usage of PII factor experienced only a slight increase of approximately 0.05 points and was not statistically significant (p = 0.4642). This metric measures whether users are given the ability to reject the processing of their personally identifiable information. The relatively low change suggests that while some websites have made progress, others remain hesitant to offer granular data control options—likely due to continued reliance on user data for personalization and targeted advertising. This reflects an ongoing tension between user autonomy and data-driven business models.

The Data Protection Principles factor saw only a slight improvement, increasing by approximately 0.03 points and was not statistically significant (p = 0.7177). This factor evaluates whether a website applies the principles of data protection by design and by default. The minimal change suggests that while some companies have incrementally improved compliance with Article 25 of the GDPR, widespread and meaningful integration of these principles remains limited. This modest rise may be attributed to the influence of EDPB Guidelines 4/2019, which clarified expectations around embedding privacy into system architecture [15]. Nevertheless, the low magnitude of change implies that operationalizing these principles into practice—and reflecting them in policy language—continues

to be a challenge across industries.

The Company Location factor showed virtually no change, increasing by only 0.01 points with no statistical significance (p = 0.8991). This factor assesses whether a website discloses its geographic location and where data is processed or transferred. The lack of meaningful improvement may be because most companies had already been compliant with this requirement prior to 2020, and changes in policy wording have since plateaued.

In contrast, the Breach Notification factor experienced the largest decrease among all factors, dropping by approximately 0.23 points, and this change was statistically significant (p = 0.0197). This metric assesses whether a website commits to notifying users of data breaches without undue delay. The decline is concerning, as it suggests a reduced emphasis on breach transparency in privacy policy language. One possible explanation is that some organizations may now assume that legal obligations under Article 34 of the GDPR are already known and thus omit explicit references in their policies. Alternatively, the increased legal and reputational risk associated with breach disclosures may have led companies to intentionally obscure or generalize their notification commitments. This trend contrasts with growing public concern over cybersecurity and may warrant closer regulatory attention.

The Data Encryption factor also showed a statistically significant decrease, dropping by approximately 0.22 points (p = 0.0011). This factor evaluates whether websites inform users that their data is encrypted at rest. Although GDPR does not mandate such disclosures explicitly, the decline suggests that fewer privacy policies are making such assurances transparent to users. As encryption becomes a more common background practice—particularly when handled by third-party processors—companies may no longer highlight it. However, this lack of transparency can erode user trust and may indicate a broader regression in how organizations communicate technical safeguards to the public.

V. Discussion of Results

The results of this longitudinal analysis reveal meaningful shifts in how websites disclose and operationalize GDPR compliance in their privacy policies. While some individual factors saw notable increases or declines, the overall trend suggests a gradual but uneven maturation of GDPR-related language since 2020. The observed changes reflect a very dynamic regulatory landscape. While the timing of score shifts aligns with evolving legal interpretations, industry pressures, and public advocacy, this study does not establish definitive causal relationships between specific events and changes in disclosure patterns.

One of the clearest trends is the strengthening of user consent and rights mechanisms. Factors such as Between Site Transfer, Right to be Forgotten, and Data Processing Consent all experienced measurable improvements, suggesting that organizations are increasingly attentive to the principles of informed consent and data subject autonomy. This aligns with the broader enforcement environment in the years following the implementation of GDPR. For example, high-profile actions by the CNIL and the Irish Data Protection Commission have focused on improper consent mechanisms, especially in cookie banners and tracking technologies [21], [22]. NOYB's large-scale complaint campaigns targeting dark patterns in consent flows have also brought public and regulatory scrutiny to these practices, motivating companies to revise their language and mechanisms accordingly [17].

Complementing this, greater transparency around Data Retention Notification indicates that organizations are more frequently acknowledging users' right to know how long their data is kept and under what legal justifications. This likely stems from increased guidance by the EDPB, as mentioned in Section IV. Together, these trends suggest a positive trajectory in areas where explicit user interaction is central to compliance.

At the same time, the data reveal gaps in technical safeguard transparency and breach accountability, raising concerns about the completeness of compliance communication. Notably, Breach Notification and Data Encryption scores declined between 2020 and 2025. This is particularly striking in an era marked by highly publicized cybersecurity incidents, which have elevated user awareness and expectations around data protection. The downward trend may indicate a shift away from explicit mention of these measures, potentially due to the assumption that they are handled at the infrastructure level (e.g., via cloud service providers) or due to legal risk aversion in committing to specific breach response procedures. These omissions underscore a broader pattern: as technical practices become more standardized, organizations may increasingly deem them implicit and exclude them from policy language, even if users still value that transparency.

Taken together, these trends suggest that privacy policies are evolving along two distinct tracks. On one hand, user-facing rights and consent mechanisms are becoming more robust and better articulated—driven largely by enforcement clarity, user awareness, and advocacy pressures. On the other hand, backend practices and architectural safeguards are increasingly abstracted away from public documentation, even as they become critical to actual compliance. This split may reflect a maturing privacy pipeline where front-end and back-end compliance are handled by different teams or vendors, leading to gaps in communication.

These findings echo patterns seen in prior work. For example, Bhatia et al. [23] noted that privacy policies often improve structurally in response to legal mandates, but the readability and transparency of key terms remain limited. Similarly, Zimmeck et al. [24] highlighted that machine learning analysis of policies show lagging improvements in technical disclosures, even as user control and consent indicators improved. The divergence between policy content and technical implementation suggests a growing disconnect that future regulation or auditing mechanisms may need to bridge.

Overall, this study indicates a cautious but measurable progression in GDPR compliance communication. The areas of greatest improvement—such as consent, transfer, and

retention—reflect clearer regulatory expectations and higher visibility in public interactions. In contrast, the declines and stagnation observed in technical and architectural disclosures highlight the limitations of relying solely on policy documents to convey the full scope of privacy protections. As GDPR enforcement continues to evolve and new legislation like the DSA and AI Act come into force, it will be critical to ensure that privacy policies not only fulfill legal requirements, but also continue to serve as meaningful tools for transparency and user empowerment.

VI. Limitations

While this study offers a longitudinal perspective on how privacy policies have evolved in the years following the GDPR's adoption, several limitations must be acknowledged to contextualize its contributions and boundaries.

A. Scope of Analysis

This study assesses what organizations choose to disclose in their privacy policies, but not whether they implement those practices in reality. As such, the findings capture trends in compliance communication rather than actual technical or organizational safeguards. PrivacyCheck™ analyzes policy language and cannot verify backend practices, user experiences, or the authenticity of security measures. Consequently, conclusions about real-world data protection are necessarily limited. Additionally, while this study discusses possible associations between score changes and external events such as legal rulings or regulatory actions, these are correlational, not causal. Without more direct evidence, links between regulatory developments and observed score trends should not be interpreted definitively.

B. Dataset Constraint and Bias

This study's longitudinal dataset is constrained by the scope of the original 2020 PrivacyCheck™ corpus, which focused on publicly traded U.S. companies. These firms tend to be larger and better resourced, limiting the generalizability of our findings to smaller or non-Western organizations. And because the dataset begins two years after the adoption of GDPR, it captures more mature policy language but excludes the initial wave of compliance activity, failing to capture how organizations responded during the early implementation phase.

Additionally, 37% of the original URLs became inaccessible or were redirected due to organizational changes. While the remaining 247 stable URLs provide a sufficient base for longitudinal comparison, this attrition may introduce potential selection bias. Firms that disappeared from the dataset may differ systematically in compliance behaviors, possibly raising questions about survivorship effects.

C. Fixed Evaluation Criteria

The ten GDPR factors evaluated in this study are also derived from the original PrivacyCheck™ model. While these factors cover many important aspects of GDPR's scope, they may not fully capture its entire breadth of regulatory standards. Furthermore, the original scores from 2020 do not reflect any amendments to the GDPR since then and updating PrivacyCheck™'s scoring critera was outside the scope of this longitudinal study.

VII. Conclusion

This study provides a five-year longitudinal analysis of GDPR compliance as reflected in website privacy policies, offering a comparative view of how key compliance factors have changed since 2020. The findings reveal a mixed trajectory: on one hand, there has been meaningful progress in articulating user rights, consent mechanisms, and data transfer transparency—possibly driven by increased enforcement clarity and advocacy efforts. On the other hand, disclosures related to technical safeguards and systemic principles have seen stagnation or even regression, indicating a shift toward implicit assumptions about backend compliance.

These trends suggest that while privacy policies are evolving in response to regulatory pressure and public expectation, they may not fully capture the breadth or depth of organizational data practices. A growing gap between front-facing legal disclosures and backend operational safeguards may hinder user trust and challenge the notion of informed consent.

Ultimately, as the regulatory landscape continues to evolve with newer digital governance frameworks, there remains a critical need for tools and methodologies that not only assess policy text but also verify substantive compliance. Bridging the gap between what is promised in policy and what is implemented in practice will be essential in advancing transparency, accountability, and user empowerment in the digital age.

References

[1] R. N. Zaeem and K. S. Barber, "The effect of the GDPR on privacy policies: Recent progress and future promise," *ACM Transactions on Management Information Systems (TMIS)*, vol. 12, no. 1, pp. 1–20, 2020.

[2] T. Linden, R. Khandelwal, H. Harkous, and K. Fawaz, "The privacy policy landscape after the gdpr," *arXiv preprint arXiv:1809.08396*, 2018.

[3] K. Van Der Schyff, S. Prior, and K. Renaud, "Privacy policy analysis: A scoping review and research agenda," *Computers & Security*, p. 104065, 2024.

[4] W. M. Tesfay, P. Hofmann, S. Kiltz, S. Wagner, T. Mahler, H. Pohl, and H. de Meer, "Privacy-Aware Data Disclosure for Personalized Online Services," in *2018 IEEE European Symposium on Security and Privacy (EuroS&P)*. IEEE, 2018, pp. 112–126.

[5] S. Zimmeck, P. Wang, L. Zou, P. Story, N. Sadeh, J. R. Reidenberg, N. C. Russell, T. B. Norton *et al.*, "Automated analysis of privacy requirements for mobile apps," in *NDSS*, 2019.

[6] Y. Shvartzshnaider, N. Apthorpe, N. Feamster, and H. Nissenbaum, "Identifying the provision of contextual integrity in privacy policies," in *IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*. IEEE, 2019, pp. 77–83.

[7] R. Kang, Y. Zhang, L. Dabbish, and Y. Wang, "Coarse privacy norms versus detailed policy descriptions: Comparing user expectations and legal disclosure," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, 2021, pp. 1–14.

[8] H. Harkous, K. F. Rahman, K. Aberer, K. G. Shin, and J.-P. Hubaux, "Polisis: Automated analysis and presentation of privacy policies using deep learning," in *27th USENIX Security Symposium (USENIX Security 18)*, 2018, pp. 531–548.

[9] B. Andow, S. Y. Mahmud, W. Wang, J. Whitaker, W. Enck, B. Reaves, K. Singh, and T. Xie, "PolicyLint: Investigating Internal Privacy Policy Contradictions on Google Play," in *28th USENIX security symposium (USENIX security 19)*, 2019, pp. 585–602.

[10] W. B. Tesfay, P. Hofmann, T. Nakamura, S. Kiyomoto, and J. Serna, "PrivacyGuide: towards an implementation of the EU GDPR on internet privacy policy evaluation," in *Proceedings of the fourth ACM international workshop on security and privacy analytics*, 2018, pp. 15–21.

[11] R. Pandita, S. Zimmeck, P. Story, and N. Sadeh, "PolicyChecker: Analyzing the GDPR Completeness of Mobile App Privacy Policies," in *Proceedings on Privacy Enhancing Technologies*, vol. 2021, no. 4. SAGE, 2021, pp. 449–467.

[12] R. N. Zaeem, R. L. German, and K. S. Barber, "PrivacyCheck: Automatic summarization of privacy policies using data mining," *ACM Transactions on Internet Technology (TOIT)*, vol. 18, no. 4, pp. 1–18, 2018.

[13] R. Nokhbeh Zaeem, S. Anya, A. Issa, J. Nimergood, I. Rogers, V. Shah, A. Srivastava, and K. S. Barber, "PrivacyCheck v2: A tool that recaps privacy policies for you," in *Proceedings of the 29th ACM international conference on information & knowledge management*, 2020, pp. 3441–3444.

[14] Court of Justice of the European Union, "Case C-311/18: Data Protection Commissioner v. Facebook Ireland Ltd and Maximillian Schrems (Schrems II)," 2020, judgment of 16 July 2020. [Online]. Available: https://curia.europa.eu/juris/document/document.jsf?docid=228677

[15] European Data Protection Board, "Guidelines 4/2019 on Article 25 Data Protection by Design and by Default, Version 2.0," 2020, adopted 20 October 2020. [Online]. Available: https://edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-42019-article-25-data-protection-design-and_en

[16] Commission Nationale de l'Informatique et des Libertés (CNIL), "The sanctions issued by the CNIL," 2025. [Online]. Available: https://www.cnil.fr/en/investigating-and-issuing-sanctions/sanctions-issued-cnil

[17] NOYB – European Center for Digital Rights, "More than 500 GDPR complaints on deceptive cookie banners," 2021. [Online]. Available: https://noyb.eu/en/projects/cookie-banner-task-force

[18] Court of Justice of the European Union, "Case C-131/12: Google Spain SL and Google Inc. v. AEPD and Mario Costeja González," 2014, judgment of 13 May 2014. [Online]. Available: https://curia.europa.eu/juris/document/document.jsf?docid=152065

[19] Information Commissioner's Office, "Age Appropriate Design Code: A Code of Practice for Online Services," Sep. 2020. [Online]. Available: https://ico.org.uk/for-organisations/childrens-code/

[20] European Union, "Digital Services Act (EU) 2022/2065," https://eur-lex.europa.eu/eli/reg/2022/2065/oj, 2022, official Journal of the European Union, L277.

[21] Commission Nationale de l'Informatique et des Libertés (CNIL), "Dark Patterns in Cookie Banners: CNIL issues formal notice to website publishers," 2024. [Online]. Available: https://www.cnil.fr/en/dark-patterns-cookie-banners-cnil-issues-formal-notice-website-publishers

[22] The Data Protection Commission Ireland, "Guidance on Cookies and Other Tracking Technologies," 2020. [Online]. Available: https://www.dataprotection.ie/en/dpc-guidance/guidance-cookies-and-other-tracking-technologies

[23] J. Bhatia, T. D. Breaux, and J. R. Reidenberg, "Designing privacy policy text to improve readability and comprehension," in *Proceedings of the IEEE International Conference on Requirements Engineering (RE)*, 2020, pp. 13–23.

[24] S. Zimmeck, P. Story, K. A. Bamberger, R. Balebako, and N. Sadeh, "PrivacyCheck: Automatic Compliance Evaluation of Privacy Policies Using AI," *Proceedings on Privacy Enhancing Technologies (PoPETs)*, vol. 2022, no. 2, pp. 22–42, 2022.

WWW.IDENTITY.UTEXAS.EDU