From Complexity to Clarity: AI/NLP's Role in Regulatory Compliance

Jivitesh Jain, Nivedhitha Dhanasekaran, Mona Diab
Carnegie Mellon University
{jivitesj, ndhanase, mdiab}@cs.cmu.edu

Abstract

Regulatory data compliance is a cornerstone of trust and accountability in critical sectors like finance, healthcare, and technology, yet its complexity poses significant challenges for organizations worldwide. Recent advances in natural language processing, particularly large language models, have demonstrated remarkable capabilities in text analysis and reasoning, offering promising solutions for automating compliance processes. This survey examines the current state of automated data compliance, analyzing key challenges and approaches across problem areas. We identify critical limitations in current datasets and techniques, including issues of adaptability, completeness, and trust. Looking ahead, we propose research directions to address these challenges, emphasizing standardized evaluation frameworks and balanced human-AI collaboration.

1 Introduction

Regulatory data compliance – ensuring an organization's data handling processes, practices, and documentation adhere to applicable laws and regulations – has become increasingly critical in today's business landscape. Organizations must maintain comprehensive documentation, implement security measures, and regularly audit their practices to comply with regulations like the General Data Protection Regulation (GDPR) (European Parliament, 2016) in the European Union, the California Consumer Privacy Act (CCPA) (California State Legislature, 2018) in the U.S. state of California, among others.[1] The stakes are high, as non-compliance can result in substantial fines, legal consequences, and reputational damage (Armour et al., 2015).

The challenge of maintaining compliance has grown significantly. Organizations often need to comply with multiple regulatory frameworks across jurisdictions. The traditional approach of manual compliance checking faces significant limitations in scale, consistency, and adaptability. As regulations grow more complex and organizations handle increasing volumes of data, manual verification becomes impractical and error-prone.

Recent advances in Natural Language Processing (NLP) and Large Language Models (LLMs) have demonstrated remarkable success in complex text analysis tasks. These models have achieved near human-level performance in document summarization (Zhang et al., 2024), legal document analysis (Martin et al., 2024), and code understanding (Chen et al., 2021; Yu et al., 2024). Their ability to process complex text at scale offers promising solutions for automating compliance processes. NLP and LLMs can assist with data compliance in several ways: analyzing privacy policies and compliance documents for completeness, verifying software systems against stated policies, making regulations accessible through question-answering systems, and helping prepare for regulatory changes. These technologies can reduce manual effort while improving accuracy and comprehensiveness, let alone facilitating accountability and governance, basic tenets of Responsible AI.

This survey examines the use of NLP and machine learning techniques for regulatory data compliance. While automated compliance has received significant attention in policy research and industry practice, the application of NLP techniques to this domain remains relatively underexplored in academic literature. Despite promising advances, the field lacks standardized benchmarks and systematic comparisons of approaches, making it difficult to track progress or establish best practices. Our survey provides the first comprehensive overview of NLP applications in automated data compli-

[1] While organizations must be compliant with several kinds regulations, such as data, financial, environmental, we focus on data-protection regulations due to (1) the significant body of NLP work in this field; and (2) the sizeable impact of these regulations on organizations' everyday functioning and expenses (Chander et al., 2021; McQuinn and Castro, 2019).

Findings of the Association for Computational Linguistics: ACL 2025, pages 26629–26641
July 27 - August 1, 2025 ©2025 Association for Computational Linguistics

ance, highlighting its current state and identifying promising future directions.

This paper is organized as follows. Section 2 establishes the case for automated data compliance by examining limitations of manual approaches. Section 3 explores key problems and surveys current techniques. Section 4 reviews available datasets and their characteristics. Section 5 discusses the current state of the field and future directions.

2 The Case for Automated Regulatory Compliance

Modern organizations face increasingly complex regulatory requirements that govern how they handle data, develop and deploy software, and conduct business. Manual compliance checking – the traditional approach – faces several critical limitations that make it inadequate for today's needs.

First, regulatory frameworks have grown significantly in complexity and scope.[2] For example, the GDPR contains 99 articles with intricate requirements, and organizations often need to comply with multiple such frameworks simultaneously. This complexity makes manual interpretation time-intensive and requires scarce, expensive expertise.[3] Second, compliance checking involves analyzing large volumes of documents and software systems. Organizations maintain numerous documents and software codebases that must align with regulatory requirements. Manual verification of all these artifacts is practically infeasible. Third, manual compliance checking is prone to human error and inconsistency. This risk increases when dealing with multiple jurisdictions or when regulations are updated. Recent advances in NLP/LLMs offer promising solutions to these challenges. LLMs can process and understand complex text, while specialized tools can automate document analysis and code checking. This paper surveys these automated approaches to regulatory data compliance, examining their current capabilities and limitations.

3 Problems and Techniques in Automated Regulatory Compliance

Ensuring regulatory data compliance is a complex task involving multiple stakeholders (regulatory bodies, data subjects, data controllers, and data processors), processes, and documentation. Due to this complexity, there are several distinct problems that can be automated. In this section, we examine the most widely studied automation approaches in the literature. Examples of some tasks are listed in Figure 1 while the papers are organized in Figure 2.

3.1 Document Compliance Analysis

Regulations such as the GDPR require maintaining several sets of documents, each with a specific function vis-à-vis the responsibilities of the entity handling personal data (Hamdani et al., 2021). These documents include Data Processing Agreements (DPAs), Privacy Policies/Notices, Data Subject Access Requests, among others. Therefore, a substantial component of regulatory compliance checking is the verification of document compliance, which involves checking if two documents are compliant with each other. Based on whether one of those documents is a regulation, this task can be further divided into document-to-regulation compliance checking, and document-to-document compliance checking (Hamdani et al., 2021).

Document-to-Regulation Compliance Given a document D and piece of regulation R, the task is to validate the *completeness* or *compliance* of D against R. A completeness check determines whether D contains all the information mandated by R, while compliance-checking further requires that the provisions of D are permissible under R. For instance, R may mandate that D disclose what personal information is collected and stored; a comprehensive description of such information would render D *complete* with respect to R. However, if R additionally stipulates that only information strictly necessary for service provision may be collected, this compliance constraint must also be satisfied for D to be *compliant* with R.

Majority of literature surveyed uses the GDPR (European Parliament, 2016) as the regulation R, due to its complexity (11 chapters, 99 articles) and reach (covers all EU residents). Privacy policies[4] and Data Processing Agreements (DPAs),[5] as defined by the GDPR, are commonly

[2] https://www.tmf-group.com/en/news-insights/articles/global-business-complexity/global-compliance-challenges-business-complexity/

[3] https://www.thomsonreuters.com/en-us/posts/investigation-fraud-and-risk/2023-cost-of-compliance-report/

[4] In the context of the GDPR, a Privacy Policy is a publicly available statement that informs data subjects about the purposes, legal basis, and methods of processing their personal data, as well as their rights under the GDPR.

[5] In the context of the GDPR, a DPA is a contractual agreement between a controller and a processor that governs the processing of personal data, ensuring compliance with GDPR

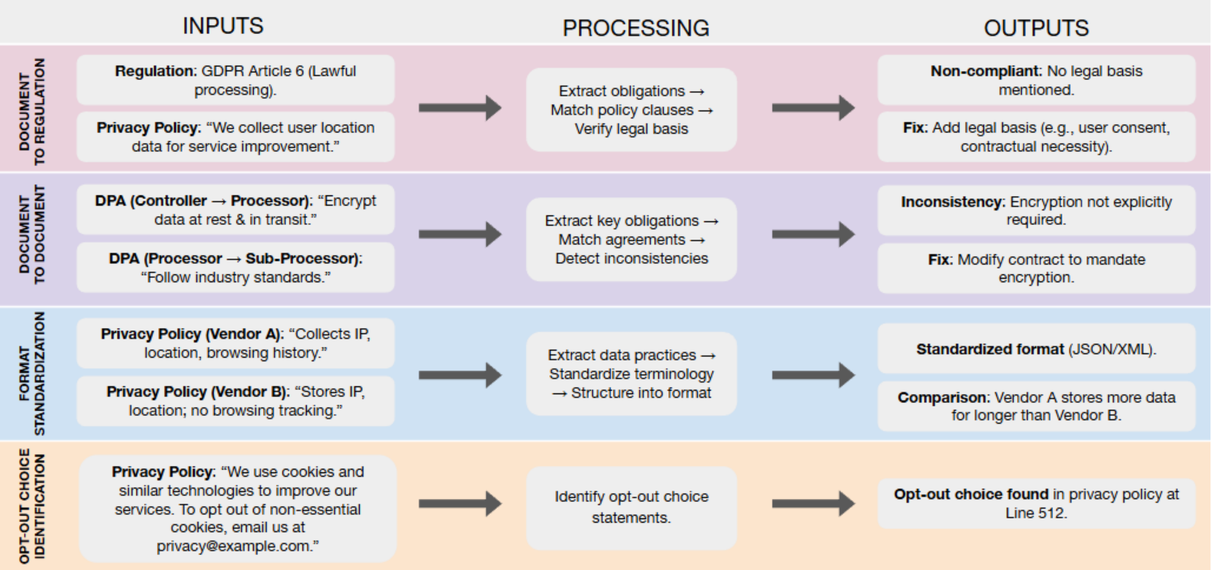
INPUTS
PROCESSING
OUTPUTS
DOCUMENT TO REGULATION
Regulation: GDPR Article 6 (Lawful processing).
Privacy Policy: "We collect user location data for service improvement."
Extract obligations → Match policy clauses → Verify legal basis
Non-compliant: No legal basis mentioned.
Fix: Add legal basis (e.g., user consent, contractual necessity).
DOCUMENT TO DOCUMENT
DPA (Controller → Processor): "Encrypt data at rest & in transit."
DPA (Processor → Sub-Processor): "Follow industry standards."
Extract key obligations → Match agreements → Detect inconsistencies
Inconsistency: Encryption not explicitly required.
Fix: Modify contract to mandate encryption.
FORMAT STANDARDIZATION
Privacy Policy (Vendor A): "Collects IP, location, browsing history."
Privacy Policy (Vendor B): "Stores IP, location; no browsing tracking."
Extract data practices → Standardize terminology → Structure into format
Standardized format (JSON/XML).
Comparison: Vendor A stores more data for longer than Vendor B.
OPT-OUT CHOICE IDENTIFICATION
Privacy Policy: "We use cookies and similar technologies to improve our services. To opt out of non-essential cookies, email us at privacy@example.com."
Identify opt-out choice statements.
Opt-out choice found in privacy policy at Line 512.

Figure 1: Illustrative examples of inputs, summary of processing steps, and outputs involved in automated data compliance tasks.

considered as the document D in question.

Many approaches pre-process the relevant provisions of the GDPR, employing techniques such as semantic role labeling and verb-predicate matching to extract actions, their actors, and beneficiaries (Cejas et al., 2023; Xiang et al., 2023; Liu et al., 2021; Zhang et al., 2023; Azeem and Abualhaija, 2024). Alternatively, researchers construct taxonomies of information types mandatorily, optionally, or conditionally required to be disclosed by data controllers and processors, as delineated in GDPR articles (Amaral et al., 2021, 2023, 2022; Hamdani et al., 2021). These structured sets are then compared against documents using rule-based (Cejas et al., 2023; Xiang et al., 2023) or ML-based methods. The latter can be broadly divided into two categories: the first involves the use of textual similarity scoring techniques (Zhang et al., 2023; Amaral et al., 2022) while the second trains classifiers for predicting whether a clause in D satisfies a specific condition extracted from R during pre-processing (Amaral et al., 2023; Hamdani et al., 2021; Liu et al., 2021; Amaral et al., 2022; Azeem and Abualhaija, 2024; Fan et al., 2020).

Although regulations need to only be preprocessed once and their computation cost is amortized across multiple document analyses, these approaches necessitate substantial manual re-engineering when regulations are revised and prove challenging to generalize across different jurisdiction frameworks. Additionally, their limitations, as identified by Hamdani et al. (2021) and others, include the lack of explainability in ML-based techniques, the inability of textual similarity scores to capture legal nuance, and their inability to correctly reason about complex cases such as the partial satisfaction of a requirement.

Conversely, approaches such as Hassani et al. (2024); Rodríguez Torrado et al. (2024) directly feed text chunks from documents and regulations into LLMs such as Llama (Touvron et al., 2023) and GPT (Brown et al., 2020), circumventing manual processing entirely. These methods offer greater adaptability across regulation revisions and jurisdictions and provide some explainability owing to LLMs' ability to generate plausible justifications (Agarwal et al., 2024). However, they sacrifice on extracting high-fidelity, structured information from R, thereby needing more computationally expensive LLMs during inference.

Document-to-Document Compliance By requiring documentation at every step of data flow, the GDPR establishes a hierarchical chain of documents that need to be consistent with each other. For instance, the DPA between a processor and sub-processor must provide data protection guarantees equivalent to the DPA between the processor and the controller. As these documents are not static, only end-to-end automated approaches that do not require any manual processing can be applied here. Although none of the literature we surveyed addresses this task directly, Hassani et al. (2024);

obligations as outlined in Article 28.

Rodríguez Torrado et al. (2024) and other similar methods will apply here with little modification.

3.2 Operational Compliance Analysis

Operational data compliance ensures that processes, guidelines, and systems managing data adhere to the compliance documentation governing their processing. This is a multifaceted problem, encompassing aspects such as the organization's information security practices, employee protocols and training around handling sensitive data, mitigation strategies in case of a data breach, and the data-handling software and infrastructure. This section focuses on the latter – verifying the compliance of a software system with a regulatory document.

Such assessment of a software system can either be carried out with complete access to the software's code base, configuration, and documentation, or only with opaque access to a running instance of the software or its decompiled machine-level code. The former setting is typical of open-source software or organizations auditing their own systems, often in partnership with an external auditor, regulatory, or certification authority. Due to the complexity of modern software systems, this task often involves several layers of checks, including manual code reviews, static code analysis checks, as well as a compliance-focused test suite (sometimes known as "Compliance as Code"). Several proprietary and open-source[6] software platforms exist for this purpose. Based on the recent success of LLMs in code understanding and generation (Chen et al., 2021; Yu et al., 2024), we hypothesize that LLM-based code understanding may be able to augment, replace, or unify several components of these systems. We identify the development of such LLM-based tools and analysis of their performance and trade-offs as a promising research direction.[7]

The problem of checking compliance of a software system with only opaque access to its running instance or decompiled machine-level code arises in the context of end-users, regulatory authorities, or software marketplaces (such as the Google Play Store[8] and the Apple App Store[9]) verifying whether the software satisfies its stated privacy policies and complies with applicable regulations. Techniques commonly used for this task include reverse-engineering or decompiling the bundled software to analyze its use of protected data, running it in an isolated sandbox environment, and monitoring network calls, operating system calls, and information written to disk.

Slavin et al. (2016) manually curated a map of privacy policy phrases and Android API[10] endpoints providing access to protected data attributes mentioned in those phrases. They used this map to scan the decompiled bytecode of Android apps for calls to API endpoints that provide access to data attributes not disclosed, or only vaguely disclosed, in the privacy policy. Similarly, Story et al. (2019) and Zimmeck et al. (2019) compared the decompiled bytecode of Android apps with privacy practices identified in their privacy policies using ML-based classifiers. Their analysis tracks sensitive Android API calls, requested permissions, and the first or third-party libraries those calls originate from, using a simplified threat model that flags data as compromised upon any sensitive API access. Fan et al. (2020) compared the protected data attributes requested by health-focused Android apps from users with those explicitly mentioned in their privacy policies. Additionally, they examined the use of Transport Layer Security (TLS) protocols in the network calls that transmit this protected data.

However, these methods suffer from limitations such as the inability of code decompilers to effectively analyze native or dynamically loaded libraries (Cao et al., 2024) and the difficulty in reconstructing optimized and obfuscated code (Dramko et al., 2024). Further, proprietary software license agreements often prohibit reverse engineering or decompilation of software,[11] which may also be restricted by applicable copyright protection laws, such as the Digital Millennium Copyright Act (United States Congress, 1998) in the U.S.

[6] Examples of open-source compliance verification software include OpenSCAP (https://www.open-scap.org), Chef Inspec (https://github.com/inspec/inspec) and OpenVAS (https://www.openvas.org), among others.

[7] While proprietary compliance tools like Google Checks (https://checks.google.com) have introduced the use of LLMs for compliance as a beta release at the time of writing, we could not identify significant scholarly research or open-source software specifically designed for this task.

[8] https://play.google.com

[9] https://www.apple.com/app-store

[10] The Android Platform Application Programming Interface (API) is a set of libraries that allow Android apps to interact with the underlying Android operating system (https://developer.android.com/reference).

[11] As an example, Microsoft's services agreement (https://www.microsoft.com/en-us/servicesagreement) contains such a clause at the time of writing.

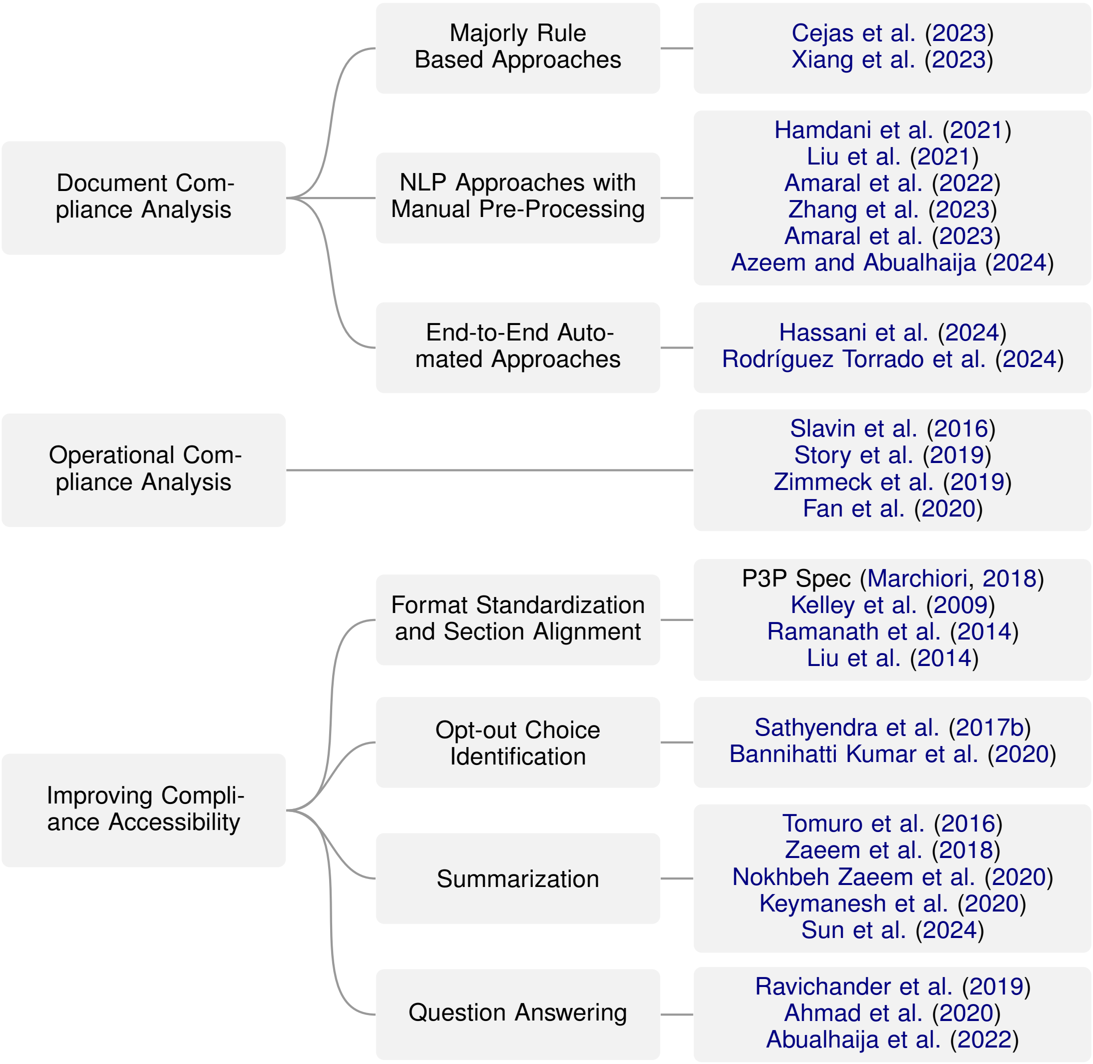
Document Compliance Analysis
Majorly Rule Based Approaches
Cejas et al. (2023)
Xiang et al. (2023)
NLP Approaches with Manual Pre-Processing
Hamdani et al. (2021)
Liu et al. (2021)
Amaral et al. (2022)
Zhang et al. (2023)
Amaral et al. (2023)
Azeem and Abualhaija (2024)
End-to-End Automated Approaches
Hassani et al. (2024)
Rodríguez Torrado et al. (2024)
Operational Compliance Analysis
Slavin et al. (2016)
Story et al. (2019)
Zimmeck et al. (2019)
Fan et al. (2020)
Improving Compliance Accessibility
Format Standardization and Section Alignment
P3P Spec (Marchiori, 2018)
Kelley et al. (2009)
Ramanath et al. (2014)
Liu et al. (2014)
Opt-out Choice Identification
Sathyendra et al. (2017b)
Bannihatti Kumar et al. (2020)
Summarization
Tomuro et al. (2016)
Zaeem et al. (2018)
Nokhbeh Zaeem et al. (2020)
Keymanesh et al. (2020)
Sun et al. (2024)
Question Answering
Ravichander et al. (2019)
Ahmad et al. (2020)
Abualhaija et al. (2022)

Figure 2: Organization of papers according to their problem area and technique used.

3.3 Improving Compliance Accessibility

Privacy policies are notoriously hard to read, and the time required to read them for all software and online services that an individual may use may be prohibitive (McDonald and Cranor, 2009; Cate, 2010; Cranor et al., 2006). Consequently, many users often encounter but do not read privacy policies (Pew Research Center, 2019), undermining their effectiveness in informing users about their data-related rights and choices. As a result, assistive technologies that simplify understanding privacy policies for users or suggest actions based on their privacy preferences are crucial steps forward to empower users to make informed decisions about their data. We present key developments in this area, and refer readers to Ravichander et al.'s (2021) survey for a more detailed review.

Format Standardization and Section Alignment The World Wide Web Consortium's Platform for Privacy Preferences (P3P) Specification (Marchiori, 2018) was an early attempt at defining a standardized, machine-readable format in which websites could express their privacy policies. P3P user agents such as *PrivacyBird* (Cranor et al., 2006) could interpret privacy policies in the P3P format, compare them against the user's pre-specified privacy preferences, and display the results to the user or take action such as blocking cookies. The specification, however, did not gain widespread use and was marked as obsolete in 2018. Kelley et al.'s (2009) "nutrition labels" for privacy policies are another attempt at displaying privacy information in a digestible format to users. While these works focus on defining structured formats with desirable properties, the need to manually encode privacy policies likely contributed to their limited adoption.

More recently, Ramanath et al. (2014) and Liu et al. (2014) used HMMs to align similar sections across privacy policies. This alignment can be used by end users to compare policies of multiple providers of the same service or as an input to information extraction and summarization pipelines.

Choice Identification A recent study by the Pew Research Center (2019) suggests that US adults prefer better tools for allowing people to control their personal information themselves to stronger laws governing the use of their information by data controllers. One such tool could be for easily identifying opportunities to opt-out of data collection, which are otherwise buried in privacy policy text. Sathyendra et al. (2017b) trained logistic regression classifiers for identifying such choice statements, the party offering the choice, and its purpose in privacy policies. Bannihatti Kumar et al. (2020) focused on finding opt-out hyperlinks and developed a browser extension, *Opt-Out Easy*, for surfacing this information to the user.

Summarization Automated text summarization offers a promising approach to distill lengthy privacy policies into brief, accessible summaries. Since most text summarization techniques are domain-agnostic or readily adaptable, we direct readers to Zhang et al.'s (2024) comprehensive survey, focusing here only on approaches specific to privacy policies. Tomuro et al. (2016) uses an ensemble of keyword identification-based and machine learning based techniques to extract and classify important sentences in privacy policies. Zaeem et al. (2018); Nokhbeh Zaeem et al. (2020); Keymanesh et al. (2020) focus on identifying and categorizing "risky" statements or sections in privacy policies – statements that could pose a privacy risk to the end user. More recently, Sun et al. (2024) explored the use of LLM agents for digital privacy management, including privacy policy summarization, and reported promising results.

Question Answering QA systems complement summarization by allowing users to explore specific topics in detail, addressing a key limitation of summary-based approaches (Sathyendra et al., 2017a). Ravichander et al. (2019) and Ahmad et al. (2020) introduced two privacy policy QA datasets, PrivacyQA and PolicyQA, respectively (see Section 4). Their work demonstrates that BERT-like transformers (Devlin et al., 2019) struggle with privacy policy QA, especially when evidence spans are longer or while identifying unanswerable questions. Abualhaija et al. (2022) focused on question answering over compliance documents to simplify the task of requirements engineering, although their approach would also be applicable to privacy policies. They used text similarity measures to identify chunks of the document relevant to the query, which they passed through BERT-based models to generate answers. Similarly, though Rodríguez Torrado et al. (2024) developed their LLM-based privacy policy QA system for compliance checking, it can also help users understand policies through direct questioning.

4 Datasets

The availability of high-quality datasets is crucial for advancing automation of compliance processes. Table 1 summarizes datasets commonly used in regulatory data compliance research, categorized by focus:

Privacy Policy Understanding Datasets like the OPP-115 Corpus (Wilson et al., 2016), MAPP Corpus (Arora et al., 2022), and MAPS Policies Dataset (Zimmeck et al., 2019) provide annotated privacy policies at varying scales, useful for developing tools that analyze privacy policies, summarize key practices, and improve user understanding.

Regulatory Alignment The Privacy Law Corpus (Gupta et al., 2022) and OPP-115/GDPR (Poplavska et al., 2020) link privacy policy content to regulations like GDPR, facilitating automated data compliance checks.

User Choices and Controls The Opt-out Choice datasets (Sathyendra et al., 2017a; Bannihatti Kumar et al., 2020) focus on identifying user options for data collection and sharing, such as opt-out mechanisms.

Question Answering Datasets like the PrivacyQA Corpus (Ravichander et al., 2019), PolicyQA corpus (Ahmad et al., 2020), and GenAIPABench (Hamid et al., 2023) support the development of interactive systems for answering user questions about privacy policies and regulations.

Large-scale Policy Retrieval MAPS Policies Dataset (Zimmeck et al., 2019) and ACL/COLING 2014 Corpus (Ramanath et al., 2014; Liu et al., 2014) enable large-scale analysis of privacy policies for pattern detection and policy comparison.

Table 1: Datasets for Automated Regulatory Compliance. The datasets contain English language texts unless specified otherwise.

Dataset	Description	Documents	Labels	Size
OPP-115 (Wilson et al., 2016)	Annotated online privacy policies by law students, focusing on data practices.	Online Privacy Policies	Yes	115 policies
MAPP (Arora et al., 2022)	App privacy policies annotated for data practices (English and German text).	Mobile App Privacy Policies	Yes	155 policies
MAPS Policies (Zimmeck et al., 2019)	Privacy policy URLs from Google Play Store apps.	Mobile App Privacy Policies	No	441,626 policies
Privacy Law (Gupta et al., 2022)	Privacy laws and guidelines from 183 jurisdictions in plain text and translations.	Global Privacy Laws	No	1,043 laws
OPP-115/GDPR (Poplavska et al., 2020)	Aligns OPP-115 annotations with GDPR principles.	Online Privacy Policies, GDPR	Yes	115 policies
Opt-out Choice (Bannihatti Kumar et al., 2020)	Annotated privacy policies for identifying opt-out mechanisms and data categories.	Online Privacy Policies	Yes	236 policies
PrivacyQA (Ravichander et al., 2019)	QA pairs annotated for understanding mobile app privacy policies.	Mobile App Privacy Policies	Yes	1,750 QA pairs
PolicyQA (Ahmad et al., 2020)	QA examples curated from privacy policies with human-annotated questions.	Online Privacy Policies	Yes	25,017 QA pairs
GenAIPABench (Hamid et al., 2023)	Benchmark for evaluating generative AI privacy assistants using annotated questions.	GDPR, CCPA, Online Privacy Policies	Yes	5 policies, 38 QA pairs
APP-350 (Zimmeck et al., 2019)	Android app privacy policies annotated for privacy practices.	Mobile App Privacy Policies	Yes	350 policies
Ramanath et al. (2014); Liu et al. (2014)	Website privacy policies annotated and segmented for analysis.	Privacy Policies	Yes	1,010 policies

5 Discussion and Desiderata

Automated regulatory data compliance has evolved significantly in recent years, driven by advances in NLP and the emergence of LLMs. Current systems show promise in several areas: they can analyze the compliance and completeness of documents, compare them to the behavior of software systems, and help stakeholders understand complex legal language (Zaeem et al., 2018; Nokhbeh Zaeem et al., 2020; Sun et al., 2024). However, significant challenges remain unsolved. Systems struggle to adapt to regulatory changes and often fail to capture nuanced legal requirements. Most research focuses narrowly on specific documents and regulations like privacy policies and GDPR, limiting generalization across jurisdictions. We lack reliable methods to verify if software systems meet regulatory requirements under realistic threat models, or to ensure consistency between different compliance documents. The field also faces broader challenges around trust and reliability – systems need to explain their decisions and maintain accuracy across different contexts. These limitations point to several future research directions.

Adapting to Regulatory Changes Current automated compliance techniques often rely on static features and taxonomies that are manually extracted from regulations, as discussed in Section 3. When regulations change, these features must be updated—a process that is both time-consuming and prone to errors. This challenge of adapting to change extends beyond existing regulations. Organizations invest significant resources in preparing for upcoming regulatory changes.[12] Yet, the potential role of automation in anticipating and preparing for such changes remains largely unexplored. The problem becomes more complex for organizations operating across borders – they must navigate different, often conflicting regulatory requirements.[13] Most current techniques and their evaluations focus narrowly on specific regulations like the GDPR and English language text, limiting their applicability to this broader compliance landscape.

The emergence of LLM-based approaches that do not require manual feature extraction (Hassani

[12] https://kpmg.com/us/en/articles/2023/cco-survey-2023-gated.html

[13] https://blog.cscglobal.com/trends-in-global-compliance/

et al., 2024; Rodríguez Torrado et al., 2024) already represents the first steps in this direction. As the multilingual capabilities of LLMs improve (Ahuja et al., 2024), the applicability of these approaches to jurisdictions worldwide will continue to grow. Alongside these improvements, the evaluation and benchmarking of these approaches in multiple languages and on documents pertaining to various jurisdictions and regulatory changes is also necessary to garner confidence in their reliability.

Creating Standardized Benchmarks The lack of standardized evaluation methods presents a major challenge to progress in automated data compliance. Current research evaluates techniques on different datasets and with variable evaluation criteria, making it difficult to compare approaches effectively. Without standardized benchmarks, it remains impossible to identify state-of-the-art approaches for specific compliance tasks, leaving practitioners to implement and evaluate methods independently. We need comprehensive benchmark datasets that capture the breadth of data compliance tasks, spanning multiple languages and jurisdictions, and including complex scenarios that require understanding regulatory nuance.

Research in benchmark development (Reuel et al., 2024; Cao et al., 2025; Weber et al., 2019) has established best practices for creating AI-focused evaluation frameworks. The process begins with defining tasks and scope, as outlined in Sections 3 and 4. The benchmark should include regulatory documents from various jurisdictions, such as South Korea's Personal Information Protection Act, Japan's Act on the Protection of Personal Information, and Brazil's General Data Protection Law, in their original languages and English translations. After identifying relevant tasks and collecting aligned data, the next step is standardizing evaluation metrics. Since most automated compliance tasks map to established NLP problems like question answering and text classification, existing evaluation metrics can be adapted (Blagec et al., 2022).

Improving Operational Compliance Automation The challenge of verifying whether software systems comply with regulations remains underexplored, as discussed in Section 3.2. Current approaches rely on a combination of manual review and basic automated checks, with few comprehensive automated solutions in the scientific literature. Recent advances in LLMs for code understanding offer promising directions for automating compliance verification. However, research exploring their application to compliance checking remains limited. In particular, we need methods that can understand complex regulatory requirements and verify their implementation across large codebases. Another crucial gap lies in analyzing bundled software—where source code may not be available—against its stated policies. Current techniques employ oversimplified threat models that may miss compliance violations; more sophisticated approaches are needed to detect potential violations under realistic scenarios.

Building Trust in Compliance Systems For automated compliance checking to move beyond research prototypes and into widespread adoption, systems must earn the trust of regulatory bodies, organizations, and end users. This requires progress in several areas. First, compliance systems must demonstrate reliable and consistent performance. Many current approaches rely on LLMs, which can produce hallucinations or inconsistent outputs when analyzing complex legal text (Huang et al., 2024). We need techniques that achieve demonstrably high accuracy on benchmarks and provide consistent results.

Second, systems must explain their decisions clearly. When a compliance violation is identified, the system should point to the specific regulatory requirement that was violated and explain how it arrived at this conclusion. While recent work has shown that LLMs can generate plausible explanations for their decisions (Agarwal et al., 2024), the critical challenge for regulatory compliance is ensuring these explanations are not just convincing, but also faithful to the underlying reasoning. This is particularly important where compliance decisions can have significant legal or financial implications.

However, recent studies indicate that reasoning produced by LLMs is frequently unfaithful to the computations that actually drive their predictions (Chen et al., 2025; Turpin et al., 2023). Promising interpretability approaches—ranging from token-level SHAP attributions (Goldshmidt and Horovicz, 2024; Mosca et al., 2022) to analyses that trace decisions to model internals (Lindsey et al., 2025; Wu et al., 2024)—aim to bridge this faithfulness gap by grounding explanations in verifiable model evidence. Continued progress on these methods and their integration into compliance pipelines is essential to improve trust in

LLM-based compliance automation.

FATE Considerations The automation of regulatory compliance raises distinct challenges around fairness, accountability, transparency, and ethics (FATE). While these principles apply to all AI systems, compliance automation's legal implications demand particular attention. Organizations need to demonstrate to regulators that their compliance systems operate reliably. This includes maintaining clear documentation of automated vs. human decisions, and tracing decisions back to specific regulatory requirements. When an AI system makes a compliance decision with legal or financial consequences, clear chains of responsibility become crucial. This necessitates frameworks for human oversight that specify when human review is required and establish procedures for handling system errors. These considerations highlight why full automation of compliance processes may be neither feasible nor desirable – while automation can handle routine tasks, human judgment remains essential for interpreting complex requirements and taking responsibility for critical decisions.

Limitations

This survey has two key limitations. First, there is limited visibility into how large organizations implement compliance for complex systems in practice, including proprietary software tools and industry case studies. Understanding these practices would require collaborative efforts between industry, regulatory bodies, and academia. Second, while research in the fields of law, software engineering, and software security offers valuable insights about the problems discussed in this survey, it remains outside the scope of this survey as we focus solely on machine learning and natural language processing research.

References

Sallam Abualhaija, Chetan Arora, Amin Sleimi, and Lionel C. Briand. 2022. Automated question answering for improved understanding of compliance requirements: A multi-document study. In *2022 IEEE 30th International Requirements Engineering Conference (RE)*, pages 39–50.

Chirag Agarwal, Sree Harsha Tanneru, and Himabindu Lakkaraju. 2024. Faithfulness vs. plausibility: On the (un)reliability of explanations from large language models. *Preprint*, arXiv:2402.04614.

Wasi Ahmad, Jianfeng Chi, Yuan Tian, and Kai-Wei Chang. 2020. PolicyQA: A reading comprehension dataset for privacy policies. In *Findings of the Association for Computational Linguistics: EMNLP 2020*, pages 743–749, Online. Association for Computational Linguistics.

Sanchit Ahuja, Divyanshu Aggarwal, Varun Gumma, Ishaan Watts, Ashutosh Sathe, Millicent Ochieng, Rishav Hada, Prachi Jain, Mohamed Ahmed, Kalika Bali, and Sunayana Sitaram. 2024. Benchmarking large language models across languages, modalities, models and tasks. In *North American Chapter of the Association for Computational Linguistics*.

Orlando Amaral, Sallam Abualhaija, and Lionel Briand. 2023. Ml-based compliance verification of data processing agreements against gdpr. In *2023 IEEE 31st International Requirements Engineering Conference (RE)*, pages 53–64.

Orlando Amaral, Sallam Abualhaija, Mehrdad Sabetzadeh, and Lionel Briand. 2021. A model-based conceptualization of requirements for compliance checking of data processing against gdpr. In *2021 IEEE 29th International Requirements Engineering Conference Workshops (REW)*, pages 16–20.

Orlando Amaral, Sallam Abualhaija, Damiano Torre, Mehrdad Sabetzadeh, and Lionel C. Briand. 2022. Ai-enabled automation for completeness checking of privacy policies. *IEEE Transactions on Software Engineering*, 48(11):4647–4674.

John Armour, Colin Mayer, and Andrea Polo. 2015. Regulatory sanctions and reputational damage in financial markets. Research Paper 62/2010, Oxford Legal Studies Research Paper. European Corporate Governance Institute (ECGI) - Finance Working Paper No. 300/2010.

Siddhant Arora, Henry Hosseini, Christine Utz, Vinayshekhar K Bannihatti, Tristan Dhellemmes, Abhilasha Ravichander, Peter Story, Jasmine Mangat, Rex Chen, Martin Degeling, et al. 2022. A tale of two regulatory regimes: Creation and analysis of a bilingual privacy policy corpus. In *LREC proceedings*.

Muhammad Ilyas Azeem and Sallam Abualhaija. 2024. A multi-solution study on gdpr ai-enabled completeness checking of dpas. *Empirical Software Engineering*, 29(4).

Vinayshekhar Bannihatti Kumar, Roger Iyengar, Namita Nisal, Yuanyuan Feng, Hana Habib, Peter Story, Sushain Cherivirala, Margaret Hagan, Lorrie Cranor, Shomir Wilson, et al. 2020. Finding a choice in a haystack: Automatic extraction of opt-out statements from privacy policy text. In *Proceedings of The Web Conference 2020*, pages 1943–1954.

Kathrin Blagec, Georg Dorffner, Milad Moradi, Simon Ott, and Matthias Samwald. 2022. A global analysis of metrics used for measuring performance in natural language processing. In *Proceedings of NLP Power!*

The First Workshop on Efficient Benchmarking in NLP, pages 52–63, Dublin, Ireland. Association for Computational Linguistics.

Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. 2020. Language models are few-shot learners. In *Proceedings of the 34th International Conference on Neural Information Processing Systems*, NIPS '20, Red Hook, NY, USA. Curran Associates Inc.

California State Legislature. 2018. California consumer privacy act of 2018. https://leginfo.legislature.ca.gov/faces/codes_displayText.xhtml?division=3.&part=4.&lawCode=CIV&title=1.81.5. Cal. Civ. Code § 1798.100 et seq.

Jialun Cao, Yuk-Kit Chan, Zixuan Ling, Wenxuan Wang, Shuqing Li, Mingwei Liu, Ruixi Qiao, Yuting Han, Chaozheng Wang, Boxi Yu, Pinjia He, Shuai Wang, Zibin Zheng, Michael R. Lyu, and Shing-Chi Cheung. 2025. How should i build a benchmark? revisiting code-related benchmarks for llms. *Preprint*, arXiv:2501.10711.

Ying Cao, Runze Zhang, Ruigang Liang, and Kai Chen. 2024. Evaluating the effectiveness of decompilers. In *Proceedings of the 33rd ACM SIGSOFT International Symposium on Software Testing and Analysis*, ISSTA 2024, page 491–502, New York, NY, USA. Association for Computing Machinery.

Fred H. Cate. 2010. The limits of notice and choice. *IEEE Security & Privacy*, 8(2):59–62.

Orlando Amaral Cejas, Muhammad Ilyas Azeem, Sallam Abualhaija, and Lionel C. Briand. 2023. Nlp-based automated compliance checking of data processing agreements against gdpr. *IEEE Transactions on Software Engineering*, 49(9):4282–4303.

Anupam Chander, Meaza Abraham, Sandeep Chandy, Yuan Fang, Dayoung Park, and Isabel Yu. 2021. Achieving privacy: Costs of compliance and enforcement of data protection regulation. Georgetown Law Faculty Publications and Other Works.

Mark Chen, Jerry Tworek, Heewoo Jun, Qiming Yuan, Henrique Ponde de Oliveira Pinto, Jared Kaplan, Harri Edwards, Yuri Burda, Nicholas Joseph, Greg Brockman, Alex Ray, Raul Puri, Gretchen Krueger, Michael Petrov, Heidy Khlaaf, Girish Sastry, Pamela Mishkin, Brooke Chan, Scott Gray, Nick Ryder, Mikhail Pavlov, Alethea Power, Lukasz Kaiser, Mohammad Bavarian, Clemens Winter, Philippe Tillet, Felipe Petroski Such, Dave Cummings, Matthias Plappert, Fotios Chantzis, Elizabeth Barnes, Ariel Herbert-Voss, William Hebgen Guss, Alex Nichol, Alex Paino, Nikolas Tezak, Jie Tang, Igor Babuschkin, Suchir Balaji, Shantanu Jain, William Saunders, Christopher Hesse, Andrew N. Carr, Jan Leike, Josh Achiam, Vedant Misra, Evan Morikawa, Alec Radford, Matthew Knight, Miles Brundage, Mira Murati, Katie Mayer, Peter Welinder, Bob McGrew, Dario Amodei, Sam McCandlish, Ilya Sutskever, and Wojciech Zaremba. 2021. Evaluating large language models trained on code. *Preprint*, arXiv:2107.03374.

Yanda Chen, Joe Benton, Ansh Radhakrishnan, Jonathan Uesato, Carson Denison, John Schulman, Arushi Somani, Peter Hase, Misha Wagner, Fabien Roger, et al. 2025. Reasoning models don't always say what they think. *arXiv preprint arXiv:2505.05410*.

Lorrie Faith Cranor, Praveen Guduru, and Manjula Arjula. 2006. User interfaces for privacy agents. *ACM Trans. Comput.-Hum. Interact.*, 13(2):135–178.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.

Luke Dramko, Jeremy Lacomis, Edward J. Schwartz, Bogdan Vasilescu, and Claire Le Goues. 2024. A taxonomy of c decompiler fidelity issues. In *Proceedings of the USENIX Security Symposium*.

European Parliament. 2016. Regulation (eu) 2016/679 of the european parliament and of the council of 27 april 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing directive 95/46/ec (general data protection regulation) (text with eea relevance).

Ming Fan, Le Yu, Sen Chen, Hao Zhou, Xiapu Luo, Shuyue Li, Yang Liu, Jun Liu, and Ting Liu. 2020. An empirical evaluation of gdpr compliance violations in android mhealth apps. In *2020 IEEE 31st International Symposium on Software Reliability Engineering (ISSRE)*, pages 253–264.

Roni Goldshmidt and Miriam Horovicz. 2024. Tokenshap: Interpreting large language models with monte carlo shapley value estimation. *Preprint*, arXiv:2407.10114.

Sonu Gupta, Ellen Poplavska, Nora O'Toole, Siddhant Arora, Thomas Norton, Norman Sadeh, and Shomir Wilson. 2022. Creation and analysis of an international corpus of privacy laws. *arXiv preprint arXiv:2206.14169*.

Rajaa El Hamdani, Majd Mustapha, David Restrepo Amariles, Aurore Troussel, Sébastien Meeùs, and Katsiaryna Krasnashchok. 2021. A combined rule-based and machine learning approach for automated gdpr compliance checking. In *Proceedings of the Eighteenth International Conference on Artificial Intelligence and Law*, pages 40–49.

Aamir Hamid, Hemanth Reddy Samidi, Tim Finin, Primal Pappachan, and Roberto Yus. 2023. Genaipabench: A benchmark for generative ai-based privacy assistants. *arXiv preprint arXiv:2309.05138*.

Shabnam Hassani, Mehrdad Sabetzadeh, Daniel Amyot, and Jain Liao. 2024. Rethinking legal compliance automation: Opportunities with large language models. *Preprint*, arXiv:2404.14356.

Lei Huang, Weijiang Yu, Weitao Ma, Weihong Zhong, Zhangyin Feng, Haotian Wang, Qianglong Chen, Weihua Peng, Xiaocheng Feng, Bing Qin, and Ting Liu. 2024. A survey on hallucination in large language models: Principles, taxonomy, challenges, and open questions. *ACM Trans. Inf. Syst.* Just Accepted.

Patrick Gage Kelley, Joanna Bresee, Lorrie Faith Cranor, and Robert W. Reeder. 2009. A “nutrition label” for privacy. In *Proceedings of the 5th Symposium on Usable Privacy and Security*, SOUPS ’09, New York, NY, USA. Association for Computing Machinery.

Moniba Keymanesh, Micha Elsner, and Srinivasan Sarthasarathy. 2020. Toward domain-guided controllable summarization of privacy policies. In *NLLP@KDD*.

Jack Lindsey, Wes Gurnee, Emmanuel Ameisen, Brian Chen, Adam Pearce, Nicholas L. Turner, Craig Citro, David Abrahams, Shan Carter, Basil Hosmer, Jonathan Marcus, Michael Sklar, Adly Templeton, Trenton Bricken, Callum McDougall, Hoagy Cunningham, Thomas Henighan, Adam Jermyn, Andy Jones, Andrew Persic, Zhenyi Qi, T. Ben Thompson, Sam Zimmerman, Kelley Rivoire, Thomas Conerly, Chris Olah, and Joshua Batson. 2025. On the biology of a large language model. *Transformer Circuits Thread*.

Fei Liu, Rohan Ramanath, Norman Sadeh, and Noah A Smith. 2014. A step towards usable privacy policy: Automatic alignment of privacy statements. In *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*, pages 884–894.

Shuang Liu, Baiyang Zhao, Renjie Guo, Guozhu Meng, Fan Zhang, and Meishan Zhang. 2021. Have you been properly notified? automatic compliance analysis of privacy policy text with gdpr article 13. In *Proceedings of the Web Conference 2021*, WWW ’21, page 2154–2164, New York, NY, USA. Association for Computing Machinery.

Massimo Marchiori. 2018. The platform for privacy preferences 1.0 (P3P1.0) specification. W3C recommendation, W3C. Https://www.w3.org/TR/2018/OBSL-P3P-20180830/.

Lauren Martin, Nick Whitehouse, Stephanie Yiu, Lizzie Catterson, and Rivindu Perera. 2024. Better call gpt, comparing large language models against lawyers. *Preprint*, arXiv:2401.16212.

Aleecia M. McDonald and Lorrie Faith Cranor. 2009. The cost of reading privacy policies. In *I/S: A Journal of Law and Policy for the Information Society*.

Alan McQuinn and Daniel Castro. 2019. The costs of an unnecessarily stringent federal data privacy law. Technical report, Information Technology and Innovation Foundation.

Edoardo Mosca, Ferenc Szigeti, Stella Tragianni, Daniel Gallagher, and Georg Groh. 2022. SHAP-based explanation methods: A review for NLP interpretability. In *Proceedings of the 29th International Conference on Computational Linguistics*, pages 4593–4603, Gyeongju, Republic of Korea. International Committee on Computational Linguistics.

Razieh Nokhbeh Zaeem, Safa Anya, Alex Issa, Jake Nimergood, Isabelle Rogers, Vinay Shah, Ayush Srivastava, and K. Suzanne Barber. 2020. Privacycheck v2: A tool that recaps privacy policies for you. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*, CIKM ’20, page 3441–3444, New York, NY, USA. Association for Computing Machinery.

Pew Research Center. 2019. Americans and privacy: concerned, confused and feeling lack of control over their personal information.

Ellen Poplavska, Thomas B Norton, Shomir Wilson, and Norman Sadeh. 2020. From prescription to description: Mapping the gdpr to a privacy policy corpus annotation scheme. In *Legal Knowledge and Information Systems*, pages 243–246. IOS Press.

Rohan Ramanath, Fei Liu, Norman Sadeh, and Noah A Smith. 2014. Unsupervised alignment of privacy policies using hidden markov models. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 605–610.

Abhilasha Ravichander, Alan W Black, Thomas Norton, Shomir Wilson, and Norman Sadeh. 2021. Breaking down walls of text: How can NLP benefit consumer privacy? In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 4125–4140, Online. Association for Computational Linguistics.

Abhilasha Ravichander, Alan W Black, Shomir Wilson, Thomas Norton, and Norman Sadeh. 2019. Question answering for privacy policies: Combining computational and legal perspectives. In *Proceedings of* *the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 4947–4958, Hong Kong, China. Association for Computational Linguistics.

Anka Reuel, Amelia Hardy, Chandler Smith, Max Lamparth, Malcolm Hardy, and Mykel J. Kochenderfer. 2024. Betterbench: Assessing ai benchmarks, uncovering issues, and establishing best practices. *Preprint*, arXiv:2411.12990.

David Rodríguez Torrado, Ian Yang, Jose Del Alamo, and Norman Sadeh. 2024. Large language models: a new approach for privacy policy analysis at scale. *Computing*, 106:3879–3903.

Kanthashree Mysore Sathyendra, Abhilasha Ravichander, Peter Garth Story, Alan W Black, and Norman Sadeh. 2017a. Helping users understand privacy notices with automated query answering functionality: An exploratory study. Technical Report CMU-ISR-17-114R, Carnegie Mellon University.

Kanthashree Mysore Sathyendra, Shomir Wilson, Florian Schaub, Sebastian Zimmeck, and Norman Sadeh. 2017b. Identifying the provision of choices in privacy policy text. In *Proceedings of the 2017 conference on empirical methods in natural language processing*, pages 2774–2779.

Rocky Slavin, Xiaoyin Wang, Mitra Bokaei Hosseini, James Hester, Ram Krishnan, Jaspreet Bhatia, Travis D. Breaux, and Jianwei Niu. 2016. Toward a framework for detecting privacy policy violations in android application code. In *Proceedings of the 38th International Conference on Software Engineering*, ICSE '16, page 25–36, New York, NY, USA. Association for Computing Machinery.

Peter Story, Sebastian Zimmeck, Abhilasha Ravichander, Daniel Smullen, Ziqi Wang, Joel Reidenberg, N Russell, and Norman Sadeh. 2019. Natural language processing for mobile app privacy compliance. In *AAAI Spring Symposium on Privacy Enhancing AI and Language Technologies*.

Bolun Sun, Yifan Zhou, and Haiyun Jiang. 2024. Empowering users in digital privacy management through interactive llm-based agents. *Preprint*, arXiv:2410.11906.

Noriko Tomuro, Steven Lytinen, and Kurt Hornsburg. 2016. Automatic summarization of privacy policies using ensemble learning. In *Proceedings of the Sixth ACM Conference on Data and Application Security and Privacy*, CODASPY '16, page 133–135, New York, NY, USA. Association for Computing Machinery.

Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, Aurelien Rodriguez, Armand Joulin, Edouard Grave, and Guillaume Lample. 2023. Llama: Open and efficient foundation language models. *Preprint*, arXiv:2302.13971.

Miles Turpin, Julian Michael, Ethan Perez, and Samuel R. Bowman. 2023. Language models don't always say what they think: Unfaithful explanations in chain-of-thought prompting. In *Thirty-seventh Conference on Neural Information Processing Systems*.

United States Congress. 1998. Digital millennium copyright act. `https://www.congress.gov/bill/105th-congress/house-bill/2281`. Public Law No: 105-304, 105th Congress (1997-1998).

Lukas M Weber, Wouter Saelens, Robrecht Cannoodt, Charlotte Soneson, Alexander Hapfelmeier, Paul P Gardner, Anne-Laure Boulesteix, Yvan Saeys, and Mark D Robinson. 2019. Essential guidelines for computational method benchmarking. *Genome Biol.*, 20(1):125.

Shomir Wilson, Florian Schaub, Aswarth Abhilash Dara, Frederick Liu, Sushain Cherivirala, Pedro Giovanni Leon, Mads Schaarup Andersen, Sebastian Zimmeck, Kanthashree Mysore Sathyendra, N Cameron Russell, et al. 2016. The creation and analysis of a website privacy policy corpus. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1330–1340.

Zhengxuan Wu, Atticus Geiger, Thomas Icard, Christopher Potts, and Noah D. Goodman. 2024. Interpretability at scale: Identifying causal mechanisms in alpaca. *Preprint*, arXiv:2305.08809.

Anhao Xiang, Weiping Pei, and Chuan Yue. 2023. Policychecker: Analyzing the gdpr completeness of mobile apps' privacy policies. In *Proceedings of the 2023 ACM SIGSAC Conference on Computer and Communications Security*, CCS '23, page 3373–3387, New York, NY, USA. Association for Computing Machinery.

Xiao Yu, Lei Liu, Xing Hu, Jacky Wai Keung, Jin Liu, and Xin Xia. 2024. Where are large language models for code generation on github? *Preprint*, arXiv:2406.19544.

Razieh Nokhbeh Zaeem, Rachel L. German, and K. Suzanne Barber. 2018. Privacycheck: Automatic summarization of privacy policies using data mining. *ACM Trans. Internet Technol.*, 18(4).

Haopeng Zhang, Philip S. Yu, and Jiawei Zhang. 2024. A systematic survey of text summarization: From statistical methods to large language models. *Preprint*, arXiv:2406.11289.

Lu Zhang, Nabil Moukafih, Hamad Alamri, Gregory Epiphaniou, and Carsten Maple. 2023. A bert-based empirical study of privacy policies' compliance with gdpr. In *2023 IEEE Conference on Communications and Network Security (CNS)*, pages 1–6.

Sebastian Zimmeck, Peter Story, Daniel Smullen, Abhilasha Ravichander, Ziqi Wang, Joel Reidenberg, N Cameron Russell, and Norman Sadeh. 2019. Maps:

Scaling privacy compliance analysis to a million apps.
Proceedings on Privacy Enhancing Technologies.