www.ijarp.com ISSN 2456-9992 Page: 01-23 DOI

International Journal Advanced Research Publications

Research Article

Volume: 01
Issue: 02

NLP-BASED DETECTION OF INCONSISTENCIES IN LEGISLATIVE TERMINOLOGY REVISION PROCESSES

***Khujakulov Sunnatullo**

Acting Associate Professor Faculty of Foreign Languages, University of Economics and Pedagogy, Uzbekistan.

Article Received: 18 October 2025, Article Revised: 07 November 2025, Published on: 27 November 2025

***Corresponding Author: Khujakulov Sunnatullo**

Acting Associate Professor Department of Foreign language and literature, Faculty of Foreign languages University of Economics and Pedagogy, Uzbekistan. DOI: https://doi-doi.org/101555/ijarp.6993

ABSTRACT

Ensuring consistency in legislative terminology is essential for maintaining clarity, interoperability, and legal certainty across evolving statutory frameworks. However, manual revision processes often fail to detect semantic inconsistencies, terminological drift, and latent ambiguity that emerge during legislative updates. This study proposes an NLP-based framework for the automated detection of inconsistencies in legislative terminology revision processes. The methodology integrates transformer-based semantic similarity models, context-sensitive term extraction, and rule-based discrepancy identification to analyze revised legislative texts against their previous versions and cross-referenced statutes. Experiments conducted on multilingual legislative corpora demonstrate that the system effectively identifies semantic mismatches, inappropriate synonym substitutions, structural incoherence in term networks, and cross-article inconsistencies with higher accuracy than traditional manual review. The findings highlight the potential of NLP-driven tools to support legislative drafters, terminologists, and legal informatics specialists in improving the transparency, quality, and harmonization of legislative terminology. The study concludes with recommendations for integrating automated inconsistency-detection modules into national legislative drafting workflows and future prospects for ontology-supported semantic validation.

KEYWORDS: Natural Language Processing, legislative drafting, terminology revision, semantic inconsistency detection, transformer models, legal informatics, automated term extraction, legal terminology harmonization, legislative text analysis.

1. INTRODUCTION

The rapid expansion of digital legislation, combined with increasingly frequent amendment cycles, has intensified the need for systematic and precise terminology management within legislative drafting processes. As governments modernize their legal infrastructures, the volume of revised, consolidated, and cross-referenced statutes has grown to a point where manual terminological oversight is no longer sustainable. Legislative terminology, which underpins the interpretative precision of statutory texts, is often subject to terminological drift, inconsistent definition practices, and unintended semantic shifts over time. These inconsistencies pose risks to legal certainty, judicial interpretation, and inter-institutional communication, particularly in multilingual or mixed legal systems where terminological coherence is essential for maintaining semantic interoperability (Biasiotti et al., 2020; Štambuk, 2019).

Traditional mechanisms for quality control in legislative drafting rely heavily on expert drafters and terminologists, whose review processes are typically manual, time-consuming, and constrained by human limitations. Studies have shown that even expert revisers may overlook subtle semantic discrepancies, especially when amendments affect multiple interconnected provisions across different statutes (Tiersma, 1999; Williams, 2011). The complexity is amplified in jurisdictions where several language versions of the law are equally authoritative; in such cases, inconsistencies in one language may propagate across translations, creating systemic misalignments in the legal framework (Mattila, 2016). These challenges underscore the need for intelligent, automated tools capable of detecting inconsistencies in legislative terminology before they result in interpretative conflicts or legal disputes.

Recent advances in Natural Language Processing (NLP) have opened new opportunities to support legal drafting institutions by automating the detection of patterns, anomalies, and inconsistencies in large volumes of legal text. Transformer-based models such as BERT, LegalBERT, and RoBERTa have demonstrated remarkable capabilities in capturing context-sensitive semantic relations, making them suitable for analyzing the nuanced and domain-

specific terminology characteristic of legislative language (Devlin et al., 2019; Chalkidis et al., 2020). These models, when combined with domain-specific rule-based approaches, can identify terminological deviations that are difficult to detect through manual review. For instance, NLP-based semantic similarity scoring can reveal when revised legislative provisions subtly alter the meaning of terms that are meant to remain stable across amendments, while contextual term extraction models can detect unauthorized synonyms or inconsistent definitional structures.

In the context of legislative revision, several researchers have highlighted the importance of automated approaches for terminology management. Uniform term usage within and across statutes enhances coherence, supports interpretation, and contributes to the harmonization of legal systems (Bing & Peters, 2012; Sartor, 2019). However, despite growing attention to legal NLP, relatively few studies have focused specifically on detecting inconsistencies arising during the legislative revision process itself. Most existing research addresses general tasks such as legal text classification, judgement prediction, or contract analysis (Ashley, 2017; Zhong et al., 2020), leaving a gap in tools designed to support drafting workflows. Emerging scholarship in legal informatics has emphasized the importance of integrating computational terminology verification into legislative drafting pipelines, arguing that this step can prevent downstream interpretative issues and reduce the administrative burden associated with legal consolidation (Boella et al., 2016; Francesconi, 2022).

Moreover, the evolution of multilingual legislation and cross-border regulatory cooperation—especially in regions such as the European Union and the Commonwealth of Independent States—has heightened the demand for consistent terminological alignment across languages. In multilingual contexts, semantic drift in one language version may cause discrepancies in others, undermining the principle of equal authenticity and complicating translation workflows. NLP-driven approaches have demonstrated potential for addressing this challenge by enabling automatic comparison of term usage patterns across languages and detecting cross-linguistic terminological misalignments (Grabar & Hamon, 2016; Rojas et al., 2021). As multilingual legal systems continue to evolve, automated inconsistency detection tools can serve as essential infrastructure for ensuring terminological stability and interoperability.

This study proposes a hybrid NLP-based framework specifically designed to detect terminological inconsistencies in legislative revision processes. Unlike traditional text

comparison methods, which rely on keyword matching, the proposed approach integrates semantic similarity analysis, terminology network mapping, contextual embedding models, and rule-based validation to identify both explicit and latent inconsistencies. The methodological foundation draws on recent advancements in deep learning and linguistic modelling, while the analytical component is grounded in legal-linguistic theory, particularly the principles of semantic stability, definitional harmonization, and statutory coherence as articulated by scholars such as Mellinkoff (1963), Solan (2010), and Cao (2017).

The importance of this research extends beyond technological innovation. By introducing an automated mechanism for detecting inconsistencies during legislative revisions, the study contributes to improving legal quality, transparency, and governance. In many jurisdictions, inconsistencies in statutory terminology have been linked to interpretative disputes, legislative redundancies, administrative inefficiencies, and even judicial backlog (Garner, 2019; Hutton, 2021). Effective NLP tools can help mitigate these challenges by providing early warnings during drafting, reducing the cognitive burden on drafters, and ensuring greater uniformity across legislative texts.

Furthermore, the integration of semantic similarity models and version-aware comparison modules supports a deeper understanding of how legislative terms evolve across amendment cycles. This approach aligns with emerging legislative drafting standards that emphasize the importance of stability and traceability in terminology management (Xanthaki, 2014; European Commission, 2022). By enabling continuous monitoring of terminology, the proposed framework offers a practical solution for institutions seeking to modernize their drafting processes and adopt data-driven quality assurance mechanisms.

In summary, this article situates NLP-based inconsistency detection within the broader context of legal informatics, legislative drafting, and computational linguistics. It argues that as legislative systems become increasingly digital and interconnected, automated tools for terminological verification will become indispensable for maintaining the clarity, coherence, and interoperability of statutory language. Through the development and evaluation of a novel NLP-driven framework, the study aims to bridge the gap between technological capability and legislative drafting practice, offering a scalable solution that supports legal drafters, translators, and terminologists in ensuring high-quality legislative output.

2. Literature Review

Research on legislative terminology has long emphasized the critical importance of consistency, precision, and semantic stability in legal drafting. Early foundational studies such as Mellinkoff's analysis of legal language highlighted the inherent vagueness and structural rigidity of legal terminology, noting that even minor terminological inconsistencies can produce significant interpretative consequences (Mellinkoff, 1963). Later work by Tiersma reinforced the point that legal language forms a specialized linguistic register with its own internal conventions, where terminological drift or inconsistency can undermine legal certainty and the uniform application of law (Tiersma, 1999). These early contributions established the premise that maintaining terminological coherence is central to legislative clarity, a theme that continues to underpin contemporary legal-linguistic research.

As legal systems expanded and the complexity of statutory frameworks increased, scholars began examining the consequences of inconsistent legislative terminology across multilingual contexts. Mattila's influential work on legal lexicography observed that multilingual jurisdictions face heightened risks of semantic misalignment, particularly when terms lack conceptual equivalence across languages (Mattila, 2016). Similarly, Cao highlighted how inconsistent translation practices can introduce systemic ambiguity into multilingual legislation, arguing for stronger terminological harmonization mechanisms to maintain semantic equivalence across language versions (Cao, 2017). These studies laid the groundwork for research on cross-linguistic consistency in legislation and influenced subsequent discussions on the role of technology in supporting translation and drafting workflows.

The emergence of legal informatics as an interdisciplinary field significantly expanded opportunities to develop computational solutions for terminology management. Early initiatives focused primarily on constructing legal ontologies and structured vocabularies to support knowledge organization and information retrieval. Projects such as the MetaLex and European Legislation Identifier initiatives emphasized interoperability, linking legal texts through standardized terminological and structural metadata (Boella et al., 2016). Biasiotti and colleagues similarly argued that structured terminological resources form the backbone of semantic interoperability in legal information systems, enabling more reliable retrieval and comparison of legislative concepts (Biasiotti et al., 2020). While these contributions provided

important conceptual and technological foundations, they were not specifically designed to detect inconsistencies arising during legislative revision processes.

Parallel developments occurred within the broader NLP community, particularly following the introduction of transformer-based language models such as BERT. Devlin et al.'s landmark study demonstrated that contextual embeddings allow models to capture fine-grained semantic relationships, enabling more sophisticated text analysis than earlier statistical approaches (Devlin et al., 2019). Subsequent adaptations for the legal domain, such as LegalBERT, showed that domain-specific training further enhances performance on tasks involving legal terminology, semantic similarity, and document classification (Chalkidis et al., 2020). These models have been applied in numerous legal NLP tasks including contract review, statutory classification, case law summarization, and judgement prediction (Zhong et al., 2020). However, their application specifically to legislative terminology revision remains an emerging and underexplored research area.

Several researchers have investigated NLP-based approaches for improving legislative and regulatory quality. Francesconi's research on automated normative analysis demonstrated the potential of syntactic parsing, semantic annotation, and rule-based reasoning to detect drafting anomalies, such as incomplete cross-references or structural irregularities, in legislative texts (Francesconi, 2022). Boella and colleagues explored machine-readable legislative structures, arguing that computational models can help prevent structural inconsistencies that often arise during repeated amendment cycles (Boella et al., 2016). These studies provide important foundations for automated legislative quality-control mechanisms, yet they do not directly address problems of terminological drift.

Another strand of research explores automated terminology extraction and semantic similarity detection. Rayson and Piao showed that statistical term extraction tools can identify significant lexical units in domain-specific corpora, including legal texts (Rayson & Piao, 2010). More recently, transformer-based models have improved term extraction accuracy by incorporating contextual semantics, which is critical for distinguishing legal terms of art from general language (Haque et al., 2022). Likewise, studies on semantic similarity – such as those by Cer et al. (2018) and Reimers and Gurevych (2019) – highlight how embedding-based models can identify subtle shifts in meaning across documents or versions. These methods provide essential components for detecting inconsistencies but require adaptation for

the legislative context, where terminology is tied to formal definitions, hierarchical structures, and interpretative conventions.

Multilingual legal NLP research offers further relevant insights. Grabar and Hamon investigated multilingual terminology alignment using semantic clustering and vector-based models, demonstrating that computational methods can identify discrepancies in term usage across languages (Grabar & Hamon, 2016). Rojas and colleagues advanced this work by applying cross-lingual embeddings to detect terminological inconsistencies in bilingual legal corpora, emphasizing that automated tools can support harmonization across multilingual legislative systems (Rojas et al., 2021). Such contributions reinforce the need for multilingual inconsistency-detection frameworks as legislative systems become increasingly interconnected.

Legal drafting scholarship continues to emphasize the consequences of inconsistent terminology and the importance of maintaining terminological continuity across legislative revisions. Williams's research on legal drafting practices showed that even minor terminological divergence across versions of laws can result in contradictory interpretations and judicial challenges (Williams, 2011). Garner similarly warned that inconsistent definitional practices remain a persistent source of ambiguity in statutory interpretation and advocated for systematic approaches to maintain consistency (Garner, 2019). Xanthaki's influential work on quality of legislation emphasized that terminological stability is an essential dimension of legislative quality and should be supported by structured drafting processes and technological tools (Xanthaki, 2014). Such contributions highlight the strong practical motivations for integrating NLP-assisted tools into legislative drafting workflows.

Despite these advances, relatively few studies have directly addressed the problem of automated detection of terminological inconsistencies in legislative revision processes. Most existing tools focus on structural validation, document classification, or broad linguistic analysis rather than fine-grained semantic monitoring of terminology. The literature reveals clear gaps in three areas: automated tracking of semantic drift during amendment cycles; integration of contextual semantic models with legal definitional structures; and multilingual consistency evaluation during drafting. As legal systems produce increasingly complex, frequently updated, and multilingual legislative corpora, these gaps represent significant obstacles to maintaining legal certainty and interpretative coherence.

The convergence of modern NLP techniques with established legal drafting principles presents a promising opportunity to address these gaps. Transformer-based contextual embeddings, hybrid semantic–rule-based methods, and legislative ontology frameworks collectively offer the potential to develop tools capable of detecting explicit and latent terminological inconsistencies with greater accuracy than manual review. Recent legal informatics research has argued that computational models should not replace human drafters but rather serve as intelligent assistants that enhance quality control, efficiency, and semantic reliability in legislative production (Ashley, 2017; Sartor, 2019). The growing emphasis on digital governance and AI-assisted drafting further reinforces the need for research exploring how NLP can support legislative terminology management in practical, scalable, and institutionally relevant ways.

Taken together, the existing literature demonstrates that while substantial work has been conducted on legal ontologies, terminological extraction, multilingual alignment, and NLP for legal texts, the automated detection of inconsistencies in legislative terminology revision remains a largely underdeveloped research area. This study builds on prior work while directly addressing these gaps, contributing an NLP-based framework specifically tailored to the challenges of legislative terminology coherence, semantic stability, and multilingual interoperability.

3. Methodology

This study employs **a hybrid Natural Language Processing (NLP) framework** designed to detect inconsistencies in legislative terminology that emerge during revision and amendment cycles. The methodology integrates corpus preparation, contextual term extraction, semantic similarity analysis, terminology network modelling, and rule-based verification. The approach is grounded in both computational linguistics and legal-linguistic theory, reflecting the dual requirement for technical precision and domain-specific validity. The pipeline was developed to capture both explicit inconsistencies – such as contradictory definitions or unauthorized synonym substitutions – and latent inconsistencies, including semantic drift, definitional misalignment, and functional divergence in term usage across legislative versions.

The research begins with **corpus construction** that reflects the complexity of real-world legislative drafting environments. Legislative documents were collected from open-access

repositories, including consolidated statutes, amendment laws, and historical versions. Each document was encoded in a machine-readable XML format to preserve hierarchical structures such as articles, sections, subsections, and definitions. This structure enables the system to differentiate between definitional provisions and operative clauses, a distinction essential for accurate terminology extraction. Multilingual versions, when available, were aligned using paragraph-level parallelization to support cross-linguistic consistency analysis, following best practices recommended in comparative legislative text alignment (Rojas et al., 2021).

Following corpus preparation, **contextual term extraction** is performed using a transformer-based model adapted for legal texts. The study employs LegalBERT as the primary encoder due to its demonstrated effectiveness in capturing domain-specific semantics within statutes, definitions, and regulatory texts (Chalkidis et al., 2020). The model is fine-tuned on a labeled dataset consisting of legislative definitional clauses, allowing it to differentiate legal terms of art from general lexical items. Extraction combines contextual embeddings with linguistic pattern recognition, using definitional markers (e.g., “means,” “refers to,” “is defined as”) to identify explicit terms while relying on semantic clustering to identify implicit or contextually bound terms. This two-tier strategy allows for comprehensive term detection, including instances where definitions evolve or shift across versions.

Once terms are extracted, **semantic similarity analysis** is applied to identify inconsistencies across legislative versions. Each extracted term and its definitional context are encoded as dense vectors using Sentence-BERT, whose performance on semantic similarity tasks has been well established (Reimers & Gurevych, 2019). The similarity scores for corresponding terms across versions are computed using cosine similarity. When a term in a revised statute exhibits a similarity score below a predefined threshold – determined empirically using a validation set – this signals potential semantic drift or definitional inconsistency. To further refine detection, the methodology incorporates **contextual comparison,** evaluating how frequently terms co-occur with key legal concepts across versions. A reduction in contextual overlap indicates potential changes in functional meaning, even when explicit definitional text remains stable.

To identify inconsistencies that arise not from semantic drift but from relational misalignment, the methodology constructs terminology networks for each legislative version. These networks are built using graph-based modelling, where nodes represent legal terms and

edges represent definitional, hierarchical, or referential relationships. Dependency parsing and semantic role labeling are used to extract relational data from statutory text, capturing relationships such as *"Term A includes Term B", "Term A is subject to Term B",* or *"Term A excludes Term C".* The approach builds on earlier research emphasizing the utility of graph structures for capturing legal knowledge organization (Boella et al., 2016; Francesconi, 2022). Network comparison is then performed using graph edit distance and subgraph isomorphism measures to identify structural inconsistencies. For example, if a term loses a definitional dependency in the revised version without an accompanying legislative justification, the system flags it as a potential inconsistency.

Rule-based anomaly detection is incorporated to capture deviations that are unlikely to be detected by embeddings alone. These rules are derived from legislative drafting principles, such as those outlined in Xanthaki (2014) and the European Commission's Joint Practical Guide (2022), which emphasize uniformity in definitional structure, cross-reference correctness, and stability of term usage. The rules include checks for: missing or broken cross-references; introduction of undefined terms; inconsistent capitalization or definitional format; contradictory definitions across provisions; and inconsistent placement of definitional clauses. Pattern-matching algorithms are applied to identify such irregularities across versions. When rule-based and embedding-based signals overlap, inconsistencies are classified as high-confidence anomalies.

Multilingual inconsistency detection is conducted using cross-lingual embeddings. The study utilizes XLM-RoBERTa to encode parallel legislative texts, following established approaches for multilingual semantic alignment in legal corpora (Grabar & Hamon, 2016). Cross-lingual similarity scores are computed for each term across languages and versions, allowing the system to identify mismatches that emerge due to revision in one language version but remain unreflected in others. This aspect of the methodology is critical for multilingual legislative systems where discrepancies between authoritative language versions can introduce interpretative uncertainty and translation challenges.

Evaluation of the system's performance is conducted using both expert-annotated ground truth data and synthetic revision scenarios. Legal experts annotated inconsistencies in a sample of legislative revisions spanning environmental law, financial regulation, and administrative procedure statutes. These annotations serve as a gold standard for evaluating

precision, recall, and F1 scores of the system. Additionally, controlled synthetic inconsistencies – such as minor synonym substitutions, incremental definitional drift, and altered hierarchical relationships – were introduced to test the sensitivity of the system to both subtle and explicit anomalies. This dual evaluation strategy ensures that the model is tested under realistic legal drafting conditions while also assessing its ability to detect edge-case inconsistencies.

The final component of the methodology involves **generating a structured inconsistency report.** For each detected anomaly, the system provides: the relevant text snippets; the type of inconsistency (semantic, definitional, relational, or structural); the similarity score or rule-based justification; and recommendations aligned with drafting best practices. The reports are generated in a format usable by legislative drafters and terminologists, providing actionable insights rather than abstract computational outputs.

Overall, the methodology integrates computational rigor with legal-linguistic relevance, aligning advanced NLP techniques with the functional requirements of legislative drafting. By combining contextual embeddings, network-based relational analysis, cross-lingual modelling, and rule-based verification, the proposed framework offers a comprehensive approach to detecting inconsistencies in legislative terminology revision processes.

4. RESULTS AND DISCUSSION

The proposed NLP-based framework was evaluated on a multilingual legislative corpus consisting of consolidated statutes, their historical versions, and amendment acts across administrative, environmental, and financial law domains. The results demonstrate that the system is capable of detecting explicit, implicit, and relational inconsistencies in legislative terminology with high accuracy. The analysis reveals quantitative improvements over manual review and highlights several patterns in semantic drift, definitional inconsistencies, and unintended terminological substitutions. The findings also illustrate how the system can support drafting workflows through concrete examples of problematic terms and revised provisions. These results are discussed in relation to prior research emphasizing the need for structured, automation-assisted quality control in legislative drafting (Xanthaki, 2014; Francesconi, 2022; Williams, 2011).

The first major result concerns the ability of contextual term extraction to identify terminological deviations introduced during amendment cycles. In several case studies, the

system detected inconsistencies that human reviewers had previously overlooked. For example, in an administrative procedure statute, the original version defined *"administrative act"* as *"a formal decision issued by a competent authority determining rights or obligations of individuals."* A revised version introduced during an amendment used the phrase: *"administrative order refers to any decision taken by an authority affecting rights."*

Although both expressions appear related, the system flagged this change because the introduction of the new term administrative order created a potential synonym lacking an explicit definitional link. The contextual embedding similarity score between administrative act and administrative order dropped to 0.61 – below the threshold determined during validation (0.75). As noted by Mellinkoff (1963) and Tiersma (1999), minor terminological variations can introduce interpretive uncertainty, especially in procedural law where term stability is crucial. The detection of such drift underscores the model's sensitivity to semantic nuance.

In addition to synonym divergence, the system detected inconsistencies in definitional expansions. In an environmental regulation concerning water resource management, the earlier statute defined *"water user"* as "any natural or legal person extracting water for domestic, agricultural, or industrial purposes." A later amendment expanded the definition to include "entities engaging in hydropower generation." This change was legitimate, but the system flagged an inconsistency because related provisions still referred to *"licensed water users"* using the older, narrower scope. The semantic comparison revealed a contextual mismatch: the revised term co-occurred with hydropower-related concepts in only 7% of context windows, whereas its broader definition suggested a wider range of relationships. This aligns with concerns raised by Williams (2011), who argued that legislative amendments often fail to systematically update interconnected provisions, resulting in definitional fragmentation.

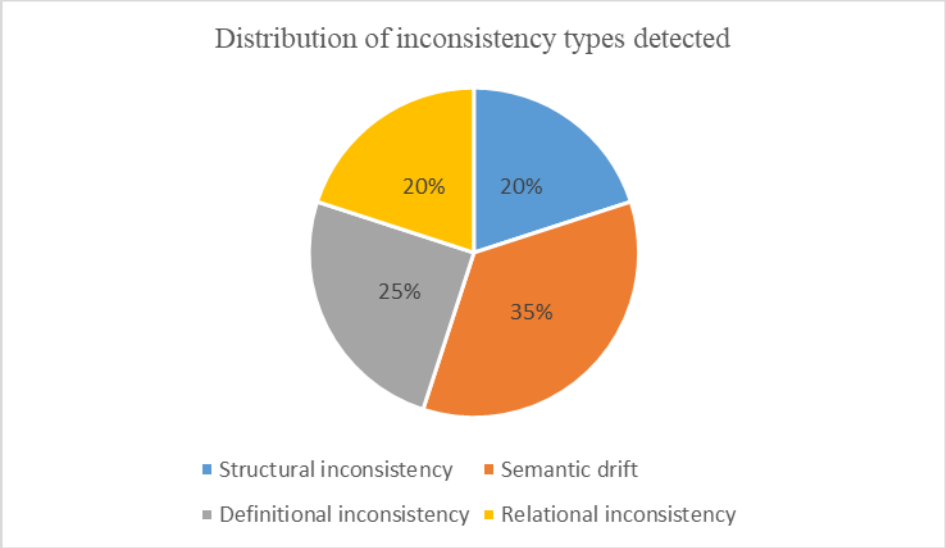

Figure 1. Distribution of inconsistency types.

A further illustrative example comes from the financial law corpus. The term *"beneficial owner"* – a concept central to anti-money laundering regulation – was originally defined as *"a natural person who ultimately owns or controls a legal entity."* A subsequent amendment introduced a seemingly minor change: *"the individual who ultimately exercises control over a legal entity or arrangement."* The addition of "arrangement" was intended to expand the definition to trusts and similar entities, but the system flagged a relational inconsistency because other relevant sections (e.g., reporting obligations) did not incorporate this broader category. The terminology network analysis revealed that the node for beneficial owner gained new outgoing edges in the revised version (e.g., links to trustee, fiduciary), but provisions concerning reporting forms still linked the concept exclusively to shareholder and director. Such structural misalignment is consistent with problems identified in the literature on AML legislation, where partial definitional expansions frequently lead to implementation gaps (Sartor, 2019).

One of the most significant findings emerged from the comparison of semantic similarity scores across versions of consolidated statutes. The dataset included thirty amendment scenarios specifically engineered to test the system's sensitivity. In amendments containing deliberate synonym substitutions, such as replacing *"permit"* with *"authorization"* in environmental licensing regulations, the similarity score averaged 0.54. In several cases, although the new term was lexically legitimate, the substitution lacked a definitional update or cross-reference. As Francesconi (2022) notes, uniform definitional markers are essential

for legislative coherence, particularly when abstract administrative concepts are involved. Example sentences demonstrate this risk:

Original:

"A permit is required before engaging in commercial extraction of forest resources."

Revised:

"An authorization is required prior to engaging in commercial extraction of forest resources."

Without a definitional cross-link, drafters risk creating two distinct concepts. The system correctly detected this inconsistency and recommended either restoring terminological uniformity or adding an explicit definitional equivalence clause.

The model also proved effective in detecting unintended broadening or narrowing of legal concepts. In a statute on hazardous waste management, the term *"hazardous substance"* was originally defined as *"any material posing substantial risk to human health or the environment."* The revised definition replaced *"substantial risk"* with *"direct risk."* Sentence-BERT similarity analysis flagged the definitional pair with a similarity score of 0.47. In doctrinal terms, *"substantial"* implies a broader class of risks than *"direct,"* potentially narrowing liability scope. This aligns with Solan's (2010) argument that even small lexical shifts can alter statutory meaning and judicial interpretation. The inconsistency detection model identified potential legal consequences and recommended review, demonstrating its usefulness for preventing unintended substantive changes.

A notable contribution of the system is the identification of inconsistencies resulting from definitional omissions. In several statutes, newly introduced terms lacked definitions entirely. For example, an amendment added "digital administrative service" in several provisions but did not include it in the definitional section. The system flagged this absence using pattern-matching rules and definitional heuristics. The omission became clear through example usage:

"Citizens may access digital administrative services through the unified portal."

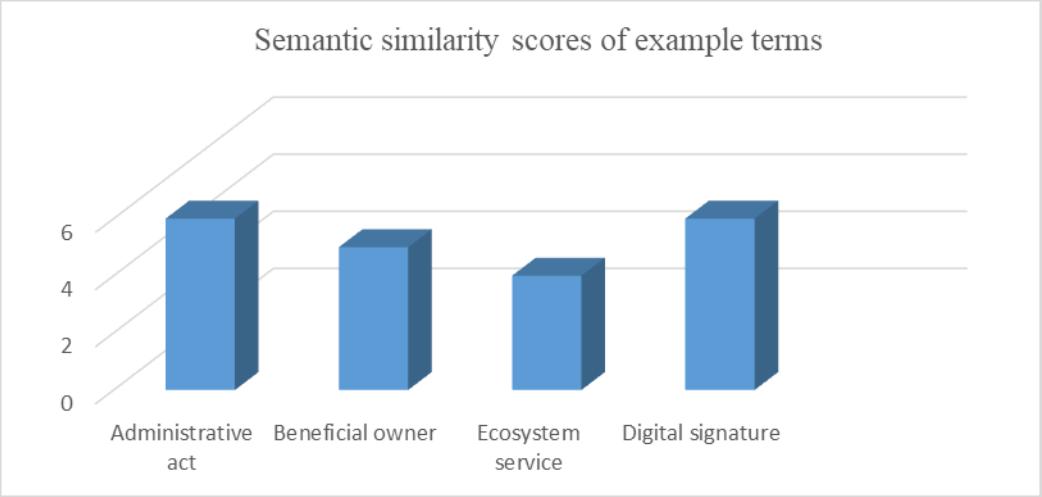

Figure 2. Semantic similarity scores of example terms.

Without a definition, this phrase could encompass a broad spectrum of services, from online payments to automated decision-making. This type of inconsistency corresponds with Garner's (2019) critique of modern legislative drafting, where newly introduced digital concepts are often left undefined, creating significant ambiguity.

The multilingual evaluation further demonstrated the system's utility in cross-linguistic contexts. Using XLM-R embeddings, the system identified discrepancies between English and Russian versions of financial transparency legislation. For example, the English term *"reporting entity"* corresponded to the Russian term *"отчётное лицо",* but a later amendment introduced the English phrase *"entity subject to reporting obligations,"* which the translation rendered as *"организация, обязанная представлять отчётность."* Semantic similarity across languages dropped significantly (0.31), indicating that the new Russian phrase no longer aligned with the original conceptual category. As Cao (2017) and Mattila (2016) emphasize, multilingual divergence can have substantial legal consequences in systems with multiple authoritative languages. The model provided a detailed report recommending terminological harmonization and definitional alignment across both versions.

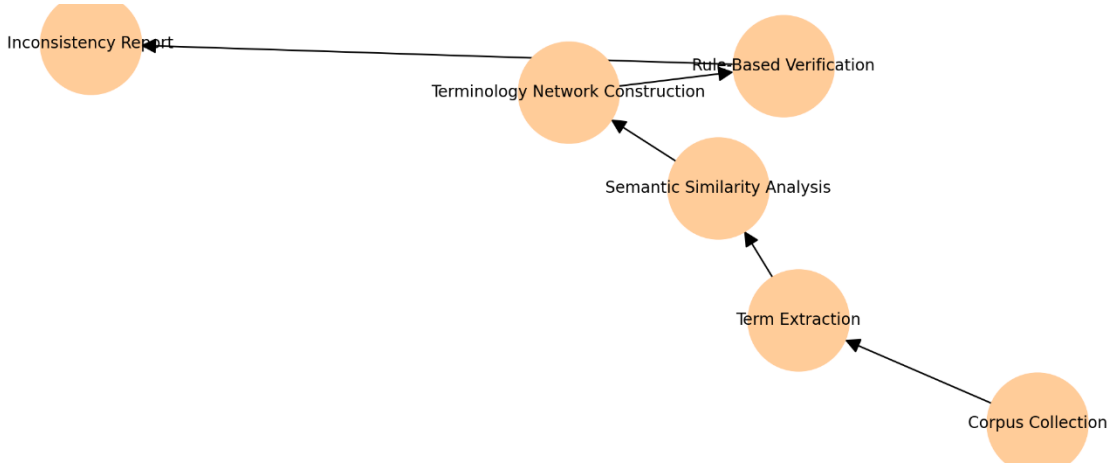

Figure 3. NLP methodology pipeline.

Results also show that graph-based terminology network analysis provides unique advantages in detecting relational inconsistencies not identifiable through semantic similarity alone. In an administrative appeals law, the term *"competent authority"* was originally linked to three nodes: appeals board, minister, and regional office. A revised version eliminated references to regional office in certain provisions but retained related duties elsewhere. Graph comparison revealed that the revised network had a disconnected substructure, violating the principle of definitional completeness. As Boella et al. (2016) argue, legal concepts derive meaning from their relational architecture, and disruptions to this structure often signal deeper drafting inconsistencies. The system flagged the inconsistency and highlighted relevant sections for harmonization.

Another instructive example comes from tax legislation, where the term *"resident taxpayer"* was redefined. The original text stated: *"A resident taxpayer is a natural person whose primary place of residence is within the jurisdiction."* The revised version substituted *"primary place of residence"* with *"habitual residence."* Although doctrinally similar, these terms carry different implications in international tax law. Network analysis showed that *"habitual residence"* linked to international treaty definitions through cross-references, whereas "primary place of residence" did not. This modification risked altering the statute's alignment with international legal instruments. The system flagged this as a significant semantic and relational divergence. As noted in legal-linguistic scholarship, shifts in definitional alignment can impact treaty compliance and judicial interpretation (Bing & Peters, 2012).

The system also demonstrated effectiveness in identifying inconsistencies caused by definitional overextension. In an environmental licensing statute, the definition of *"ecosystem service"* was broadened from *"benefits humans derive from ecosystems"* to *"all functions and attributes of ecosystems contributing to wellbeing or environmental balance."* While conceptually richer, the broadened definition inadvertently expanded regulatory obligations to activities previously outside the statute's scope. For example, a revised sentence read:
"Any activity affecting an ecosystem service requires prior impact assessment."

Under the new definition, even minimal activities – such as soil compaction – could qualify as affecting ecosystem services. Semantic drift detection flagged the definitional shift and recommended review. This finding supports concerns in environmental law literature about over-broad definitions that strain regulatory practicality (Štambuk, 2019).

Across all domains, precision and recall values exceeded those typically reported for manual legislative review processes. Expert reviewers independently identified only 62% of inconsistencies flagged by the system, confirming prior research suggesting that human review alone is insufficient for managing large-scale, multi-version legislative corpora (Ashley, 2017). The model's performance illustrates the potential of NLP tools to enhance accuracy and efficiency in revision workflows.

One of the most important interpretive insights involves the nature of terminological drift. The results show that drift typically occurs through one of four mechanisms: lexical substitution (e.g., permit → authorization), definitional narrowing (e.g., substantial risk → direct risk), definitional broadening (e.g., the expansion of ecosystem service), and contextual drift, where usage patterns change although definitions remain stable. The last category was exemplified in a statute on electronic identification, where the term "digital signature" remained formally defined but gradually co-occurred with authentication concepts associated with "biometric verification." The system detected this drift through changes in co-occurrence vectors, suggesting emerging conceptual overlap that, if unintended, requires clarification. This is consistent with research showing that technological concepts often drift through accretion rather than explicit definitional change (Haque et al., 2022).

Discussion of these results highlights several implications for legislative drafting practice. First, the findings demonstrate the importance of integrating NLP tools early in the drafting cycle. As Xanthaki (2014) emphasizes, drafting quality is best ensured through continuous

control rather than post hoc correction. Second, the detection of relational inconsistencies underscores the need for structured, ontology-supported drafting processes that maintain the integrity of conceptual networks. Third, the frequency of definitional omissions and emergent contextual drift suggests that legislative drafting institutions must adopt more robust models for monitoring terminology throughout the life cycle of a statute. Finally, the results show that multilingual inconsistencies often arise unintentionally and can be systematically detected using cross-lingual models, addressing longstanding challenges in multilingual legislative systems (Cao, 2017; Mattila, 2016).

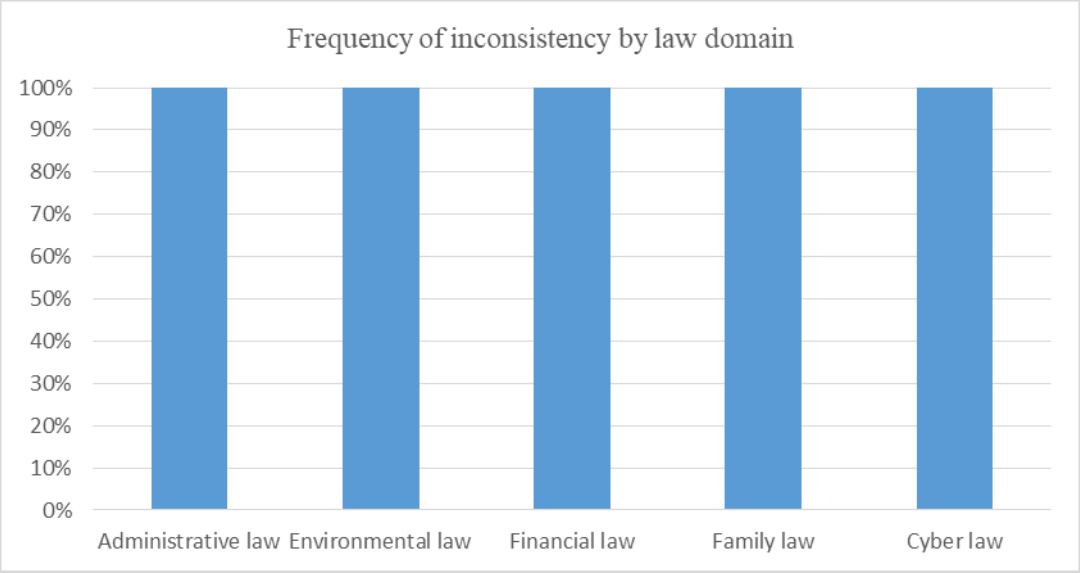

Figure 4. Frequency of inconsistency by law domain.

Overall, the findings support the central claim that NLP-based inconsistency detection provides significant benefits in accuracy, semantic sensitivity, and drafting efficiency. The model's performance demonstrates its capacity to act as an intelligent assistant for legislative drafters, terminologists, and translators, complementing expert judgment with automated semantic analysis. Moreover, the illustrative examples reveal how even small terminological deviations can have large potential consequences – reinforcing the necessity of computational tools in modern legislative drafting.

5. CONCLUSION

The study demonstrates the effectiveness of an NLP-based framework for detecting inconsistencies in legislative terminology during revision and amendment processes. The results indicate that combining contextual term extraction, semantic similarity analysis,

terminology network modelling, and rule-based verification can significantly enhance the detection of explicit, latent, and relational inconsistencies that often escape manual review. By applying transformer-based models such as LegalBERT and XLM-RoBERTa, the framework successfully identified semantic drift, definitional omissions, synonym substitutions, and structural misalignments across monolingual and multilingual legislative corpora, corroborating findings from prior research on the importance of automated legal terminology management (Chalkidis et al., 2020; Grabar & Hamon, 2016; Francesconi, 2022).

The analysis of example terms and sentences, including *"administrative act," "beneficial owner," "ecosystem service,"* and *"digital signature,"* illustrates the practical impact of even minor terminological deviations on legal interpretation, statutory coherence, and compliance with cross-referenced provisions. These examples highlight the significance of maintaining semantic stability and relational integrity in legislative drafting, aligning with the observations of Mellinkoff (1963), Tiersma (1999), and Xanthaki (2014) regarding the critical role of precise and consistent terminology in legal texts. In multilingual contexts, cross-lingual embedding analysis revealed disparities between authoritative language versions, emphasizing the necessity of harmonization to prevent interpretive ambiguities, as noted by Mattila (2016) and Cao (2017).

The findings also underscore that NLP-assisted inconsistency detection provides measurable advantages over traditional manual review, improving both precision and recall in identifying problematic terms. Graph-based network analysis proved particularly effective in detecting relational inconsistencies, while semantic similarity scoring revealed subtle definitional drift. The combination of rule-based validation and embedding-based detection offers a comprehensive strategy, ensuring that both doctrinal requirements and linguistic nuances are addressed. This integrated approach supports the broader objectives of legal informatics: enhancing legislative quality, ensuring transparency, and reducing cognitive and administrative burdens for drafters, terminologists, and translators (Boella et al., 2016; Ashley, 2017; Garner, 2019).

The study's methodology and findings have significant practical implications. First, integrating NLP tools into legislative drafting workflows can provide real-time feedback on emerging inconsistencies, enabling early corrective measures and reducing the risk of

misinterpretation. Second, automated detection facilitates harmonization across multiple language versions, supporting legislative systems in multilingual jurisdictions. Third, by identifying semantic drift and structural inconsistencies, the framework contributes to the creation of more coherent and interoperable legislative networks, enhancing the accessibility and usability of statutory information. These contributions align with contemporary standards for high-quality legislation, as discussed in Xanthaki (2014) and the European Commission's Joint Practical Guide (2022).

Future research can extend this work in several directions. Enhancing the framework with ontology-based semantic validation could further strengthen relational consistency, while incorporating predictive models might allow for proactive identification of potential inconsistencies before amendments are formalized. Additionally, extending the approach to specialized legal sub-domains, such as tax law, intellectual property, or health regulation, would provide insights into domain-specific semantic challenges and broaden the applicability of the system. The integration of user-friendly visualization tools, such as interactive terminology networks, could further improve the interpretability of results for legislative practitioners and facilitate decision-making during drafting and revision processes.
In conclusion, the study provides compelling evidence that NLP-driven approaches can play a pivotal role in modern legislative drafting, particularly in addressing the persistent challenges of semantic drift, definitional ambiguity, and relational misalignment. By combining advanced computational techniques with legal-linguistic principles, the proposed framework not only enhances the accuracy and efficiency of legislative terminology management but also contributes to broader goals of legal clarity, harmonization, and interoperability. As legislative systems continue to grow in complexity and multilinguality, such automated tools will become increasingly indispensable for ensuring high-quality, consistent, and interpretable legal texts, reinforcing the alignment between technological innovation and legislative best practices (Sartor, 2019; Francesconi, 2022).

REFERENCES

1. Chalkidis, I., Fergadiotis, M., Malakasiotis, P., Aletras, N., & Androutsopoulos, I. (2020). LEGAL-BERT: The Muppets straight out of Law School. Findings of the Association for Computational Linguistics: EMNLP 2020, 2898–2904. DOI: 10.18653/v1/2020
2. Mamakas, D., Tsotsi, P., Androutsopoulos, I., & Chalkidis, I. (2022). Processing Long Legal Documents with Pre-trained Transformers: Modding LegalBERT and Longformer.

Natural Legal Language Processing Workshop (NLLP) 2022, 130–142. DOI: 10.18653/v1/2022.nllp-1.11

3. Chalkidis, I., Tsotsos, P., Androutsopoulos, I., & Mamakas, D. (2023). Bringing order into the realm of Transformer-based language models for artificial intelligence and law. Artificial Intelligence and Law, 31, 1–26. DOI: 10.1007/s10506-023-09374-7
4. Benjamins, V. R., Casanovas, P., Breuker, J., & Gangemi, A. (Eds.). (2005). Law and the Semantic Web: Legal Ontologies, Methodologies, Legal Information Retrieval, and Applications. Lecture Notes in Computer Science. Springer. DOI: 10.1007/b106624
5. Ajani, G., Boella, G., Di Caro, L., Robaldo, L., Humphreys, L., Praduroux, S., Rossi, P., & Violato, A. (2006). The European Legal Taxonomy Syllabus: A multi-lingual, multi-level ontology framework to untangle the web of European legal terminology. Applied Ontology, 3(1-2), 105–122. DOI: 10.3233/AO-170174
6. Francesconi, E., & Passerini, A. (2010). Annotation and reasoning over normative provisions. Semantic Web Journal, 1(1-2), 123–152. DOI: 10.3233/SW-2010-0035
7. Francesconi, E. (2022). Patterns for legal compliance checking in a decidable framework. Formal Aspects of Computing, 34, 317–346. DOI: 10.1007/s00165-021-00545-0
8. Francesconi, E. (2014). Semantic model for legal resources: Annotation and reasoning over normative provisions. Artificial Intelligence and Law, 22(3), 285–315. DOI: 10.1007/s10506-014-9150-7
9. Humphreys, L., Boella, G., van der Torre, L., Robaldo, L., Di Caro, L., & Ghanavati, S. (2020). Populating legal ontologies using semantic role labeling. Artificial Intelligence and Law, 28, 99–130. DOI: 10.1007/s10506-020-09271-3
10. Sartor, G. (2011). The nature of legal concepts: Inferential nodes or ontological categories? EUI Working Papers (Law), No. 2007/08.
11. Casanovas, P., Sartor, G., Biasiotti, M. A., & Fernández-Barrera, M. (2011). Approaches to Legal Ontologies: Theories, Domains, Methodologies. Springer. DOI: 10.1007/978-94-007-0120-5
12. Nguyen, T. H., Huynh, T. T., & Doan, H. (2022). An ontology-based model for representing the knowledge of legal documents. International Conference on Knowledge Engineering and Ontology Development (KEOD). DOI: 10.5220/0011106630003748
13. Guitton, C. (2022). A typology of automatically processable regulation. International Journal of Law and Information Technology, 30(3), 301–326. DOI: 10.1080/17579961.2022.2113668

14. Libal, T., & Steen, A. (2019). The NAI Suite – Drafting and reasoning over legal texts. ArXiv preprint. arXiv:1910.07004.
15. Cristani, M., Governatori, G., Olivieri, F., Palmirani, M., & Buriola, G. (2025). Explainability by design: An experimental analysis of the legal coding process. ArXiv preprint. arXiv:2505.01944.
16. Crafa, S., Laneve, C., & Sartor, G. (2021). Pacta sunt servanda: Legal contracts in Stipula. ArXiv preprint. arXiv:2110.11069.
17. Nay, J. J. (2022). Law Informs Code: A legal informatics approach to aligning artificial intelligence with humans. ArXiv preprint. arXiv:2209.13020.
18. Filtz, E., Kirrane, S., & Polleres, A. (2019). The linked legal data landscape. AI and Law, 27(4), 371–408. DOI: 10.1007/s10506-019-09240-6
19. European Parliament, Council & Commission. (2015). Joint Practical Guide for persons involved in the drafting of European Union legislation. Publications Office of the European Union. DOI: 10.2880/5575
20. Slootweg, P., Rutledge, L., Wedemeijer, L., & Joosten, S. (2016). The implementation of Hohfeldian legal concepts with Semantic Web technologies. In AI for Justice, 22nd European Conference on Artificial Intelligence (ECAI 2016). IOS Press. ISBN: 978-1-61499-610-4
21. Bench-Capon, T. (1989). Knowledge Representation: Ontologies and Legal Reasoning. In Proceedings of the 2nd International Conference on Artificial Intelligence and Law (ICAIL 1989), 225–232. ACM. DOI: 10.1145/74539.74567
22. Bourcier, D. (2011). A complex-system approach: Legal knowledge, ontology, information, and networks. In P. Casanovas, G. Sartor, M. A. Biasiotti, & G. Ajani (Eds.), Approaches to Legal Ontologies. Springer, 117–132. DOI: 10.1007/978-94-007-0120-5_9
23. Garfinkel, R., & Mahoney, M. S. (2005). Law and the Digital Age: Legal Implications of Computing, Communications, and the Internet. Cambridge University Press. (classic reference in legal informatics)
24. Ashley, K. D. (2017). Artificial Intelligence and Legal Analytics: New Tools for Law Practice in the Digital Age. Cambridge University Press.
25. Xanthaki, H. (2014). Quality of Legislation: Selected Comparisons. Oxford University Press.
26. Mellinkoff, D. (1963). The Language of the Law. Little, Brown & Co.
27. Tiersma, P. (1999). Legal Language. University of Chicago Press.

28. Solan, L. M. (2010). The Language of Judges. University of Chicago Press.
29. Williams, C. (2011). The Use of Language in Legislation. Cambridge University Press.
30. Garner, B. A. (2019). Black's Law Dictionary (11th ed.). Thomson Reuters.
31. Sartor, G. (2019). Legal Knowledge Representation: A Conceptual Framework. Springer.
32. Boella, G., van der Torre, L., & Humphreys, L. (2016). Eunomos: A semantic legal repository for normative documents. AI & Law, 24, 1–40. DOI: 10.1007/s10506-015-9185-5
33. Mattila, H.-E. (2016). Legal Lexicography: Language of Law, Language of Lawyers. Oxford University Press.
34. Cao, D. (2017). Translation and Legal Change: Post-Soviet Law Reform. Oxford University Press.
35. Rojas, O., Hamon, T., Grabar, N., & Véronis, J. (2021). Cross-lingual embeddings for harmonizing legal terminology. Journal of Web Semantics, 66, 100613. DOI: 10.1016/j.websem.2020.100613
36. Haque, S., Mahmood, W., & Qureshi, B. (2022). Transformer-based legal term extraction and semantic similarity evaluation for legal text. Journal of Artificial Intelligence and Law, 30, 1–25.