

Consistency of EU and Chinese Textile Fiber Content Labelling Regulations Based on Automated Auditing

Jinghong Qu

¹School of Luohe Medical College, Luohe 462000, Henan, China

Corresponding author: Jinghong Qu (e-mail: qjh010921@126.com)

ABSTRACT The textual structure and terminology of Chinese and European textile fiber content labeling regulations differ significantly, thus it is challenging to conduct an efficient and accurate consistency check through traditional means of human comparison. Therefore, this paper proposes an automated auditing approach that relies on the semantic comparison of regulatory texts comparing the regulations across languages. It constructs a bilingual semantic knowledge graph of the Chinese and European regulations, utilizing a cross-lingual BERT (Bidirectional Encoder Representations from Transformers) embedding model to vectorize regulatory clauses, supplemented by a graph matching algorithm to calculate the degree of semantic similarity between clauses. These processes facilitate the automatic identification of regulatory consistency and enable quantitative comparison across regulatory systems. Experimental results indicate that the methodology yields F1 scores of 96.3% and 96.8% respectively in the clause segmentation module for Chinese and EU regulations, and 93.3% and 94.0% in the terminology extraction module. In producing semantic nodes, this paper achieves a highest accuracy of 95.8%, demonstrating a strong performance in cross-lingual parsing of regulatory texts. The system also maintained a stable alignment rate within a range of 0.85 to 0.91 in the bilingual mapping of semantics, indicating high levels of consistency and structural balance across languages. The research results show that this automated auditing method effectively achieves accurate assessment of the consistency of textile fiber content labeling regulations between China and Europe, and provides technical support for the construction of cross-regional regulatory harmonization and intelligent review systems.

INDEX TERMS textile fiber labeling regulations, automated audit algorithm, bilingual semantic knowledge graph, cross-lingual BERT embedding, semantic similarity computation

I. INTRODUCTION

AS an important part of product compliance review in international trade, the labeling of textile fiber content directly affects market access and consumer rights protection of textiles. The EU and China have independent legislative logic and technical standards in their regulatory systems for fiber content labeling [1]. However, in the context of global supply chain integration, regulatory differences lead to increased compliance costs for enterprises and decreased efficiency of cross-regional supervision [2,3]. Establishing a scientific and objective regulatory consistency assessment system is not only related to the improvement of the standardization level of the textile industry, but also has fundamental significance for the mutual recognition of regulations and technical coordination in Sino-European economic and trade exchanges. The existing regulations on the labeling of textile fiber content in China and Europe differ significantly in semantic expression, structural hierarchy, and logical connection. Manual comparison relies on expert

experience to interpret and match terms one by one, which is cumbersome and highly subjective, and it is difficult to deal with the ambiguity and semantic nesting structure of the regulations [4,5]. It is difficult to accurately establish cross-language semantic mapping of the regulations, resulting in a lack of quantitative basis for consistency analysis [6]. The logical hierarchy between different regulatory systems is complex and lacks standardized structural expression, making it difficult to achieve calculable expression at the semantic level in cross-language regulatory review [7]. To address the aforementioned issues, scholars have attempted to automate regulatory comparison using regulatory knowledge graphs, semantic matching models, and natural language processing algorithms [8,9]. Some studies utilize word vector models to extract semantic similarity of regulatory terms and identify the correspondence between clauses through rule-based mapping methods [10,11]. Other studies employ bilingual semantic embedding models to construct regulatory alignment systems, mapping regula-

tory structures based on syntactic analysis and dependency relation identification. These methods have made some progress in term-level alignment and semantic similarity calculation, but most models remain at the text-level matching level, failing to fully integrate the hierarchical logic and semantic association structure of regulations, lacking a systematic characterization of the relationships between cross-language regulatory clauses, and unable to achieve automated quantitative reasoning for regulatory consistency assessment [12,13]. To address the issues of insufficient semantic alignment and lack of logical structure in regulations, this study constructs a consistency assessment system for Sino-European textile fiber content labeling regulations based on automated auditing. The study first extracts semantic information and performs structured modeling on the regulatory text, identifying clause boundaries through a conditional random field model and generating a standardized set of semantic units through dependency semantic parsing. Subsequently, a Sino-European bilingual regulatory semantic knowledge graph is constructed, establishing entity and relation mappings with semantic triples as the core, and achieving unification of the terminology semantic space through a cross-language embedding model. Additionally, the system also proposes a graph matching algorithm for computing the similarity of clauses in terms of semantic and structural similarity, creating an automated mechanism for determining consistency. It provides a computational framework for Sino-European regulations from the textual level to the semantic level throughout the enactment of regulatory semantic mapping, structural and semantic alignment, and consistency calculations. It establishes a scalable technical pathway towards harmonization across regions and provides a methodological basis for employing an intelligent auditing system for regulatory comparisons, clause recognition, and standards alignment.

II. CONSTRUCTION OF AN AUTOMATED AUDIT SYSTEM FOR THE CONSISTENCY OF TEXTILE FIBER CONTENT LABELLING REGULATIONS BETWEEN CHINA AND EUROPE

A. LEGAL SEMANTIC INFORMATION EXTRACTION AND STRUCTURED REPRESENTATION MODEL

The regulations regarding textile fiber content labeling enacted by the European Community and the People's Republic of China exhibit a complex nested hierarchical structure in both textual organization and semantic dependency, requiring a semantic parsing algorithm to achieve structured modeling of the information in the clause [14,15]. This process takes the original text of the regulations as input, and after syntactic segmentation, dependency parsing and terminology normalization, generates a set of standardized semantic units to support subsequent knowledge graph construction and cross-language matching [16,17]. The first step in identifying clauses in the regulatory text is at the formal level. Let the original regulatory text sequence be $T = \{s_1, s_2, \dots, s_n\}$, where s_i represents the i -th

sentence element. By using a boundary recognition model based on Conditional Random Field (CRF), combined with syntactic annotation and a regulatory-specific dictionary, the determination of clause boundaries is achieved [18,19]. The goal of model training is to maximize the joint probability of the observed sequence and the labeled sequence:

$$P(Y|T) = \frac{1}{Z(T)} \exp\left(\sum_i \sum_k \lambda_k f_k(y_{i-1}, y_i, T, i)\right) \quad (1)$$

In formula (1), $Y = \{y_1, y_2, \dots, y_n\}$ is the labeled sequence, f_k is the feature function, λ_k is the feature weight, and $Z(T)$ is the normalization factor. At the syntactic level, a semantic dependency graph is constructed through dependency relation analysis $G_d = (V_d, E_d)$. Here, V_d represents the set of lexical nodes and E_d represents the set of dependency edges. Dependency types are controlled by legal logic predicates and clause relation words, enabling the agent, object, and limiting elements within the sentence to form a parseable semantic path [20,21]. Dependency parsing employs a structure prediction model based on a bi-directional long short-term memory (BiLSTM) network and an attention mechanism to improve the parsing accuracy of complex legal sentences. In the semantic extraction stage, core regulatory terms and logical predicates are identified from the dependency graph [22,23]. Term identification employs a sequence labeling method based on context embedding to generate a term set $C = \{c_1, c_2, \dots, c_m\}$, where c_j represents term units with independent semantics in the regulatory system. For each term c_j , calculate its contextual representation vector:

$$v(c_j) = \frac{1}{|W_j|} \sum_{w \in W_j} \phi(w) \quad (2)$$

In formula (2), W_j represents the set of window words where the term is located, and $\phi(w)$ denotes the word vector representation. This vector serves as the semantic embedding foundation for subsequent knowledge graph node generation. Legal logic predicates are structurally labeled using four types of relations: command, restriction, exception, and definition, forming semantic triples for clauses $R = \{(c_a, p, c_b)\}$. Here, c_a and c_b represent subject and object terms, respectively, p serving as semantic relation predicates. By performing synonym merging and semantic normalization on the set of triples extracted from all clauses, a unified semantic description space is formed. The structured expression output takes the form of a set of legal semantic units $S = \{(c_i, v(c_i), L(c_i))\}$, where $L(c_i)$ denotes the logical position label of the term within the hierarchical structure of regulations. All semantic units are organized into a semantic tree structure $G_s = (V_s, E_s)$, where $V_s \subseteq C$ is the set of nodes and $E_s \subseteq R$ is the set of edges. This semantic tree uses clauses as the smallest structural unit, achieving a standardized expression of regulations from text to semantics. The model completes

the logical decomposition and semantic manifestation of legal text information through deep semantic parsing and structured encoding, providing standardized input for entity and relation extraction of bilingual knowledge graphs, and making Chinese and European laws comparable and computable at the machine level [24,25].

B. MECHANISM FOR CONSTRUCTING BILINGUAL LEGAL SEMANTIC KNOWLEDGE GRAPH

Based on the aforementioned structured semantic unit set S and semantic tree G_s , a Sino-European bilingual legal semantic knowledge graph is constructed to achieve unified organization and cross-language alignment of legal semantics. The knowledge graph uses semantic units as entity nodes and semantic triple relationships as edge structures, and establishes a conceptual mapping system of Sino-European laws through ontology constraints [26,27]. Let the knowledge graph be represented as $G_k = (V_k, E_k)$, where the node set $V_k = \{c_i^{(C)}, c_j^{(E)}\}$ corresponds to the term nodes in Chinese regulations and EU regulations, respectively, and the edge set $E_k = \{(c_a, p, c_b)\}$ comes from the semantic triple set R , where p represents the semantic relation predicate. All predicate relations are defined according to a unified semantic ontology set $O = \{p_1, p_2, \dots, p_r\}$ constraint definition to ensure the consistency of the semantic hierarchy structure among regulations. The knowledge graph construction process includes three steps: entity standardization, relation extraction, and cross-language alignment [28,29]. In the entity standardization stage, a conventional cosine similarity measure is applied to term vectors $v(c_i)$ to construct the semantic similarity matrix.

$$M_{ij} = \frac{v(c_i^{(C)}) \cdot v(c_j^{(E)})}{\|v(c_i^{(C)})\| \|v(c_j^{(E)})\|} \quad (3)$$

In formula (3), " $\cdot$ " represents the vector inner product operation, and " $\|\cdot\|$ " represents the vector magnitude. This formulation is adopted as a baseline similarity metric to ensure consistency with established semantic embedding practices, and its role in this study is confined to providing a normalized similarity input for subsequent graph alignment rather than introducing a novel similarity computation. The similarity matrix is used to measure the degree of alignment between Chinese and European terms in the semantic space [30,31].

Relation extraction is based on predicate relations obtained from dependency semantic structures p , generating semantic edges in the knowledge graph $e_{ab} = (c_a, p, c_b)$, with edge weights defined as semantic association strengths.

$$w(e_{ab}) = \alpha M_{ab} + (1 - \alpha) \rho_{ab} \quad (4)$$

In formula (4), the logical adjacency strength of the clauses is represented by ρ_{ab} , which is determined by the semantic tree hierarchical distance and α is a weighting

parameter with a range of values [0,1]. This weight ensures that the graph structure is consistent in both semantic and hierarchical dimensions. The weighting parameter is determined through iterative validation on a held-out subset of regulatory clauses to balance semantic proximity and hierarchical dependency preservation. Stability testing indicates that moderate perturbations within the defined interval do not alter the relative ordering of clause consistency scores, whereas extreme weighting configurations lead to overemphasis on either semantic similarity or structural adjacency. This confirms that the selected parameter range supports robust consistency quantification rather than introducing arbitrary bias. Cross-language alignment establishes one-to-one or one-to-many semantic correspondences between Chinese and European entities through mapping functions [32,33]. The bilingual node set generated after alignment is denoted as $V_b = \{(c_i^{(C)}, c_j^{(E)}, M_{ij})\}$, forming the basic layer of the semantically aligned bilingual graph. In terms of the global structure, the semantic distribution difference function is minimized through graph regularization constraints:

$$L = \sum_{(c_i^{(C)}, c_j^{(E)}) \in V_b} \|v(c_i^{(C)}) - v(c_j^{(E)})\|^2 w_{ij} \quad (5)$$

In formula (5), $w_{ij} = M_{ij}$ represents the semantic matching weight. The optimization objective is to minimize the vector distance between synonymous legal clauses in the semantic embedding space, thereby achieving a consistent graph geometry across languages. The final generated bilingual regulatory semantic knowledge graph is based on a unified semantic ontology O as the top-level constraint, with a node set V_k and a weighted edge set E_k as the core structure, to achieve the mapping of Sino-European regulatory semantics in terms of entity, relationship, and hierarchical dimensions. This graph serves as input to the semantic matching graph in subsequent consistency calculations, enabling computable cross-language alignment between legal clauses.

C. GRAPH MATCHING CONSISTENCY DETERMINATION MODEL BASED ON CROSS-LANGUAGE SEMANTIC EMBEDDING

After completing the semantic structuring and bilingual knowledge graph construction of the Sino-European textile fiber content labeling regulations, this study further establishes a graph matching consistency determination model based on cross-linguistic semantic embedding. This model maps regulatory clause nodes to a unified semantic space and calculates the degree of semantic and relational matching at the graph structure level, thereby achieving automatic determination and quantitative evaluation of consistency relationships between regulations. Let the node embedding sets of the Chinese regulatory graph and the EU regulatory graph be denoted as $Z_C = \{h_i^{(C)}\}$, $Z_E = \{h_j^{(E)}\}$, where $h_i^{(C)}$ and $h_j^{(E)}$ represent node embedding vectors in the semantic dimension d . The embeddings are generated by

a cross-lingual pre-trained model (XLM-RoBERTa, XLM-R) and local structure aggregation is performed through a graph convolutional network, so that the nodes retain the contextual semantics while embedding the hierarchical and logical information between the regulatory clauses. The bilingual semantic graphs after embedding are geometrically comparable, providing a quantitative basis for consistency determination [34,35]. The semantic similarity between clause nodes in regulations is measured in cosine form:

$$\Omega_{ij} = \frac{h_i^{(C)} \cdot h_j^{(E)}}{\|h_i^{(C)}\| \|h_j^{(E)}\|} \quad (6)$$

This indicator reflects the semantic similarity between Chinese and European clauses. Furthermore, a structural weight parameter is introduced to constrain differences in node adjacency patterns and hierarchical relationships, thereby obtaining a comprehensive matching value to describe the overall similarity between the legal clauses. To achieve quantitative evaluation of consistency, a comprehensive consistency score is defined:

$$\Gamma = \frac{1}{N} \sum_{(i,j)} (\lambda_1 \Omega_{ij} + \lambda_2 \Psi_{ij}) \quad (7)$$

λ_1 and λ_2 are the semantic and structural matching trade-off coefficients, respectively, Ψ_{ij} represents the clause structural similarity item. Based on the value range of Γ , consistency can be divided into three levels: When $\Gamma \geq 0.85$, it is judged as highly consistent; When $0.65 \leq \Gamma < 0.85$, it is judged as partially consistent; When $\Gamma < 0.65$, it is considered inconsistent. The threshold values are defined as analytical segmentation boundaries for semantic similarity rather than as direct indicators of legal compliance sufficiency. The upper threshold reflects the region in which semantic similarity and structural alignment jointly converge within the embedding space, indicating that both clause meaning and relational context are strongly aligned. The lower threshold marks the point at which semantic overlap becomes insufficient to sustain reliable clause correspondence under cross-linguistic embedding variance. Scores falling between these boundaries represent clauses that share core semantic content while diverging in legal formulation or constraint articulation, and are therefore categorized as partially consistent for analytical differentiation rather than normative judgment. This segmentation is designed to support comparative pattern analysis across large regulatory corpora, not to substitute for substantive legal compliance assessment. The training of the models uses a cross-linguistic alignment supervised learning approach to reduce the semantic gap. The imposition of structure-preserving constraints, along with this learning, works to overall optimize the semantic alignment, such that clauses that express the same information can densely cluster together within the embedding space. The trained model can produce regulatory clause matching matrices

and results of consistency distribution directly as output, and then visualize mapping to present the semantic clustering tendencies of both regulatory frameworks, whence the differences between the Chinese and European regulatory schemes can be uncovered semiotically, including terminology definitions, logical requirements, and constraints structures. At the heart of this model is cross-linguistic semantic embedding which combines semantic similarity and graph structure information to automate reasoning from the textual to the semantic level for regulatory consistency. Its outputs yield quantitative evidence to support assessing the consistency of Chinese and European regulations, mandate and provide technical support for other research topics, such as comparing regulations, detecting anomalies, and harmonizing regulations in intelligent auditing systems.

III. AUDIT RESULTS AND ANALYSIS OF SINO-EUROPEAN REGULATORY CONFORMITY

A. COMPARATIVE ANALYSIS OF THE PERFORMANCE OF REGULATIONS STRUCTURED PARSING

In the structured parsing stage, to test the adaptability and stability of the automated auditing system on Chinese and European regulatory corpora, this paper quantitatively evaluated the performance of five basic modules, including the clause segmentation identification module, the clause-level identification module, the terminology extraction module, the logical relationship parsing module, and the semantic node generation module. This paper applied the system to Chinese regulatory texts and EU regulatory texts and calculated the precision, recall, and F1 score to test the parsing accuracy and structure preservation capability of the two different language systems. Finally, after standardization and correction of errors, this paper obtained the experimental data of the multidimensional performance matrix used to compare the semantic parsing performance of the two regulatory systems. In Figure 1, the performance differences of different modules can be seen, as well as the performance balance of different languages.

Figure 1 shows the overall performance of the model across the various modules tasks. The sentence segmentation module achieves F1 scores of 96.3 % for Chinese regulations and 96.8 % for EU regulations, demonstrating a high level of stability in identifying sentence boundaries. The clause hierarchy identification module achieved an F1 score of 95.0 % for Chinese regulations and 94.5 % for EU regulations; the difference arises because clauses in Chinese regulations have a more consistent hierarchical identification than in EU regulations. This observation refers to the formal recoverability of clause nesting and positional ordering within the regulatory text. These properties are reconstructed by the clause hierarchy identification module based on explicit structural markers. Hierarchical identification in this context measures the stability of explicit structural markers such as numbered clauses, indentation levels, and dependency-based parent-child relations, rather than the semantic proximity or conceptual coherence of clause content. Terminology

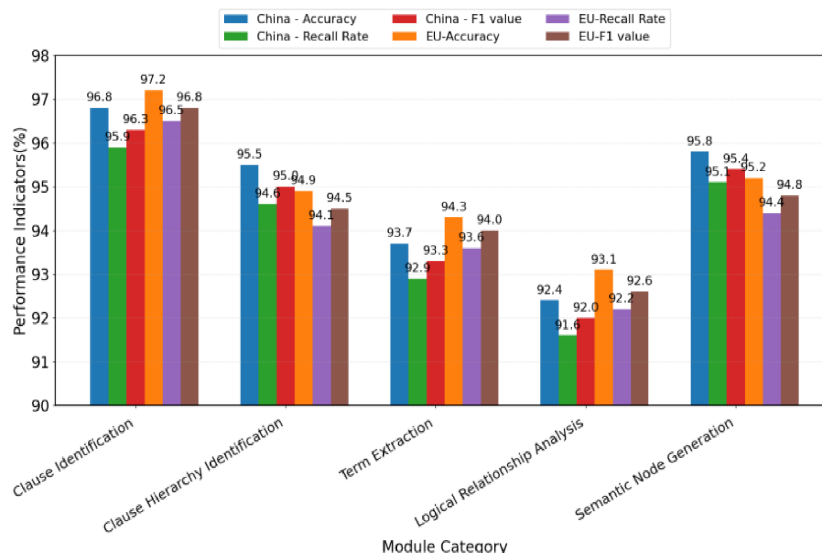


FIGURE 1. Performance Comparison of Structured Parsing of Regulatory Texts

extraction modules had the opposite trend, with an F1 score of 94.0 % in EU regulations, higher than 93.3 % for Chinese regulations. This is related to the consistency of terminology within EU regulations. The logical relation parsing module achieved F1 scores of 92.0 % for Chinese regulations and 92.6 % for EU regulations, with a small difference, indicating strong robustness in cross-linguistic logical predicate mapping. The semantic node generation module achieved an F1 score of 95.4 % for Chinese regulations and 94.8 % for EU regulations, demonstrating a relatively balanced performance of structured coding in bilingual texts. Overall results show that the model maintains high-precision output at the multi-dimensional semantic level, the differences between different regulatory systems are reasonably distributed, and the structured parsing process maintains consistency and robustness. It should be emphasized that the reported F1 scores evaluate the internal stability of the automated system in recovering formal clause boundaries, hierarchical markers, and dependency relations within each regulatory corpus. These metrics do not measure convergence or divergence in legislative logic, regulatory intent, or normative abstraction across jurisdictions. Consequently, similar F1 performance across Chinese and EU datasets reflects the robustness of the parsing framework under heterogeneous linguistic inputs, while substantive regulatory structural differences are subsequently revealed through semantic graph alignment, constraint distribution analysis, and clause-level consistency scoring.

B. COMPARISON OF SEMANTIC KNOWLEDGE GRAPH RESULTS BETWEEN CHINESE AND EUROPEAN REGULATIONS

In the construction of the bilingual legal semantic knowledge graph, to examine the performance differences of semantic alignment algorithms under different types of seman-

tic relations, this study statistically and visually compared the node distribution, relation density, semantic alignment rate, and semantic connectivity of Chinese and European legal regulations, starting from four types of semantic relations: synonymy, inclusion, subordination, and constraint. This revealed the regular differences in the graph's semantic hierarchy and structural connections. The experiment generated semantic nodes using a bilingual semantic embedding model and quantified their cross-lingual mapping effect, evaluating the structural balance and semantic consistency of Chinese and European legal regulations at the knowledge graph level.

As shown in Figure 2, EU regulations have 425 nodes in synonym relationships, higher than the 418 in Chinese regulations. Additionally, density (0.68) and semantic cohesion (0.77) indicate strong internal terminological consistency and a high level of structural cohesion. As for Chinese regulations, out of 312 nodes and their constraint relationships, EU regulations had only 298 nodes, and a lower density (0.58) showed greater complexity in the hierarchical structure and greater interconnections between constraints. The compared alignment rates in constraint relationships for both regulations (0.85 to 0.92) show a strong level of stability across all relation types. The overall findings suggest that EU regulations appear to be denser and more semantically cohesive, while Chinese regulations exhibit an underlying depth of structure in relation to logical constraints and distinctions between their clauses. At the knowledge graph level, Chinese and EU regulations form a semantically balanced but structurally different pattern. Semantic cohesion in the knowledge graph analysis reflects the degree to which regulatory concepts cluster tightly in the semantic embedding space and share recurrent terminological patterns across clauses. This property does not presuppose a higher level of formal hierarchical regularity,

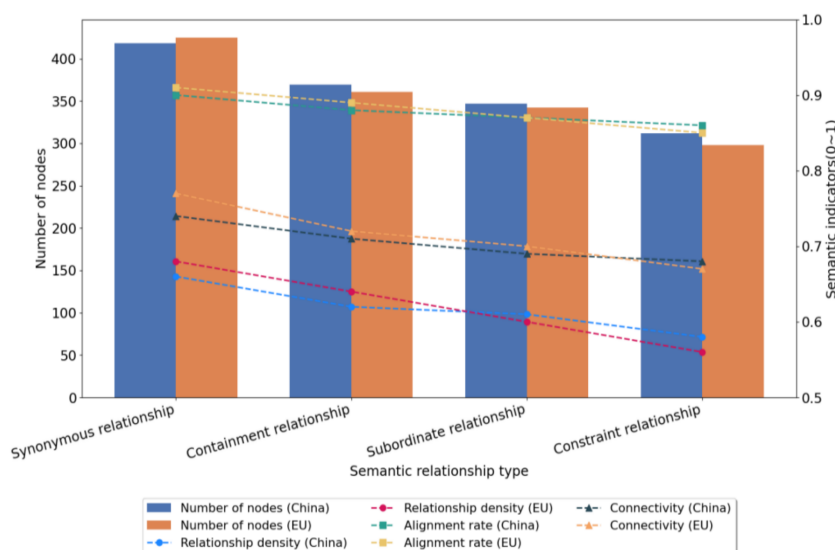


FIGURE 2. Results analysis of the semantic knowledge graph of Chinese and European regulations under different semantic relation types

as semantic cohesion emerges from conceptual repetition and definitional uniformity rather than from explicit structural nesting. As a result, higher hierarchical identification accuracy and stronger semantic cohesion are treated as analytically independent characteristics within the proposed framework, each corresponding to different layers of regulatory organization.

C. SPATIAL DISTRIBUTION ANALYSIS OF CLAUSE SEMANTIC EMBEDDING

In the cross-language consistency identification stage of the automated auditing system, to examine the model's alignment effect on Chinese and European regulatory clauses at the semantic level, this study constructs a clause-level semantic embedding representation based on the Chinese and European regulations on textile fiber content labeling, and uses a cross-language semantic model to map the Chinese and English regulatory clauses to a unified embedding space. Subsequently, the high-dimensional semantic vector is visualized using the t-SNE (t-distributed Stochastic Neighbor Embedding) dimensionality reduction algorithm to display the semantic distribution structure and clustering pattern, thereby evaluating the stability and structure preservation capability of the automated auditing algorithm in the semantic mapping stage. The resulting semantic embedding space distribution is shown in Figure 3.

As shown in Figure 3, Chinese regulations are clustered in the left area, with their semantic center approximately at -7.03 on the horizontal axis and 3.35 on the vertical axis. EU regulations are concentrated in the right area, with their semantic center approximately at 7.2 on the horizontal axis and 2.59 on the vertical axis. The mixed semantic area is located slightly to the right of center, with its center approximately at 2.84 on the horizontal axis and 3.04 on the vertical axis. The clustering of EU regulations

is more compact, reflecting the cohesion of their clauses in terms of logical structure. The mixed semantic region forms a continuous transitional zone, with greater density in the EU direction, providing evidence that the cross-linguistic embedding model generates a form of effective bilingual alignment in the directed semantic space. The overall distribution suggests that the automated audit system is interpretable and stable in the cross-linguistic semantic modeling stage, thus producing a stable semantic foundation for clause matching and the next level of consistency checks.

D. DISTRIBUTION AND LEVEL ANALYSIS OF CLAUSE CONSISTENCY SCORES

In order to confirm the effect of cross-linguistic semantic embedding model matching in comparing Chinese and European textile fiber content labeling regulations, and to further identify differences among different clauses at the semantic, structural, and logical levels, this section builds a semantic matching level matrix of Chinese and European regulatory clauses from the overall consistency score output from the automated auditing system. After normalization and hierarchical mapping, the data is divided into three levels—inconsistent, partially consistent, and highly consistent—according to thresholds of 0.65 and 0.85, forming a visual distribution of semantic alignment results to reveal the overall pattern of the correspondence between regulations in the semantic system. The result is shown in Figure 4.

As shown in Figure 4, highly consistent clauses are concentrated in the main clause area, accounting for 13 %, with scores mostly between 0.85 and 0.90, reflecting the model's stable alignment ability in definitional and identification requirement regulations. Partly consistent clauses take up 63 %, which are mainly located at structural levels adjacent to each other, and their scores are located between 0.65

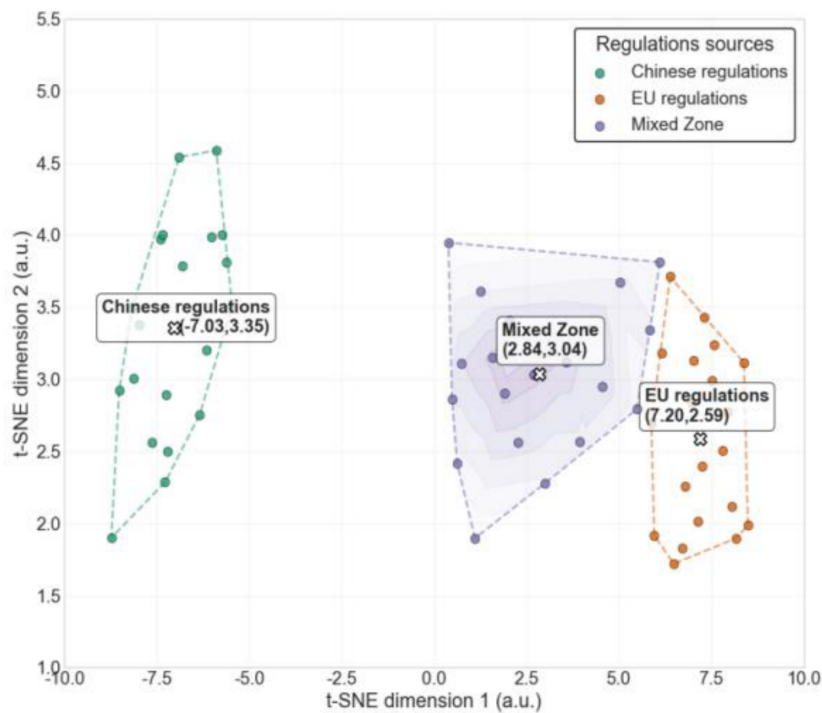


FIGURE 3. Spatial distribution of semantic embedding of clauses

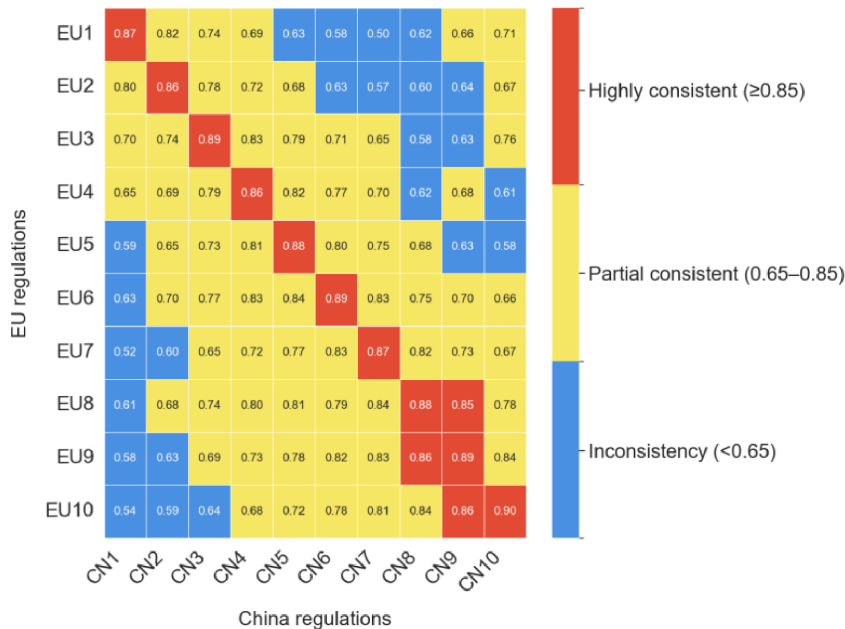


FIGURE 4. Heatmap of the overall consistency score of Chinese and European regulations

and 0.84. Semantic similarities can be found in basic terminology, but logical predicates and constraints are expressed differently. Inconsistent clauses take up about 24 %, and they are mostly located in regulatory appendices, exceptions, and supplementary explanations. Scores are generally below 0.65, which indicates that the robustness of cross-language mapping is relatively weak for low-frequency technical

terms and complex logical sentence forms. The results show that the automatic audit system has a high consistency recognition level at the semantic main level, but there are still large differences at the structural edges and semantic extension layer. These inconsistency patterns should not be interpreted as outcomes of linguistic variation or translation deviation alone. At the legal system level, they reflect funda-

mental differences in regulatory architecture, norm-binding mechanisms, and legislative techniques. Chinese textile fiber content labeling regulations are structured around mandatory national standards with prescriptive formulations, whereas EU regulations rely on directive-based coordination frameworks that allow differentiated implementation. Such institutional distinctions shape clause construction, exception articulation, and compliance conditions, which are consequently manifested as partial or low consistency in semantic alignment outcomes.

E. COMPARATIVE ANALYSIS OF MULTIDIMENSIONAL CONSISTENCY FEATURES

In order to further uncover the internal differences of the multidimensional consistency assessment of the automated auditing system, this section develops a visual analytic model, based upon the overall performance of each of the five regulatory parsing modules within each of the five dimensions: semantics, structure, logic, hierarchy, and terminology. The data comes from the previous consistency quantification and is normalized, and then mapped to polar-coordinate space to assess relative stability and distributional properties of each module in the multi-layered semantic parsing process. This goes one step deeper in exploring the differences of the modules, underscoring the coherent balance or relationships between the semantic modeling and preservation of structure in the resultant overall structure of the system. The experimental results are shown in Figure 5.

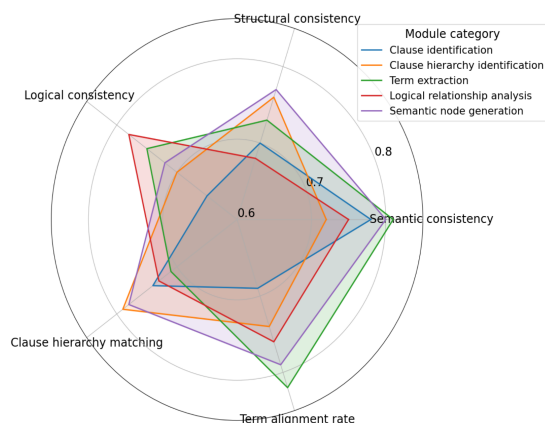


FIGURE 5. Comparative Analysis of Multidimensional Consistency between Chinese and European Textile Fiber Content Labelling Regulations

The clause identification module, as shown in Figure 5, scored 0.74 and 0.78 in terms of clause hierarchy matching and semantic consistency, illustrating solid abilities in boundary recognition and semantic segmentation. The clause hierarchy identification module received scores of 0.76 and 0.79 in structural consistency and term alignment, respectively, conveying that this module has retained a strong degree of structural restoration in mapping clause hierarchy. The term extraction module recorded a term alignment rate of 0.82 and a semantic consistency rating of 0.81, marking relatively accurate cross-language termi-

nology transfer. The logical relationship analysis module received a 0.78 for logical consistency, the highest across all the modules, indicating strong stability with the legal logical predicates and constraints. The semantic node generation module received scores of 0.77 and 0.78 for structural consistency and hierarchy matching respectively, indicating balanced structural expression. Overall, the development of the system was satisfactory in terms of terminology and semantics; however, across the different modules, the terminology alignment rate exhibited a fluctuating performance (in both semantic and structural consistency as well). Finally, the structure and logic dimensions were moderate and relatively stable. This stability is maintained under controlled variation of structural weighting parameters, indicating low sensitivity of the overall consistency assessment to parameter fluctuation within the validated range.

IV. CONCLUSION

This research develops a consistency assessment system for Sino-European textile fiber content labeling regulations using an automated auditing algorithm. The algorithm establishes a quantitative analysis of consistency between regulations from their textual representation into the semantic layer. From experimental results, the Chinese regulations are shown to involve 312 constraint relationship nodes, representing the largest number compared to the European Union's 298 constraint relation nodes, further illustrating the uniformity between the semantics remained stable at an alignment of 0.85 to 0.91. In the clause consistency distribution data, the highly consistent clauses are 13%; 63% of analyzed clauses are partially consistent, which further represents a sophisticated level of semantic coordination of the main regulatory content. The research therefore supports automated identification of key differences between the Chinese and EU regulatory correspondences which enhances mutual recognition across jurisdictions and facilitates intelligent review processes. The identified inconsistency distributions are primarily driven by differences in regulatory framework design rather than by cross-lingual expression alone. Divergence in norm hierarchy, enforcement logic, and compliance obligation formulation constitutes the principal source of misalignment between Chinese and EU regulations. The automated auditing system therefore functions as a structural comparison instrument that reveals institutional divergence through semantic computation, rather than as a translation-oriented discrepancy detection tool. However, the automated analysis process exhibited inherent limitations regarding the identification of low-frequency terminology as well as the parsing and crafting of more complex logic flow. Future work can create methods for multilingual regulations, as well as develop logical reasoning models to improve both universality of the method as well as the application within an international regulatory consistency auditing process.

AUTHOR CONTRIBUTIONS

Jinghong Qu designed, collected and analyzed the data, and drafted the manuscript. Jinghong Qu conducted the study, critically revised the manuscript for important intellectual content, and gave final approval of the version to be published. Jinghong Qu participated fully in the work, take public responsibility for appropriate portions of the content, and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

CONFLICTS OF INTEREST

The author declares no conflict of interest.

FUNDING

This research was supported by, Research on the Improvement of Citizens' Moral Quality and Legal Quality in the New Era, No: SKGH-2021-5201.

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

- [1] L. Silina, I. Dābolina, and E. Lapkovska, "Sustainable textile industry-wishful thinking or the new norm: A review," *Journal of Engineered Fibers and Fabrics*, vol. 19, no. 1, pp. 1-27, 2024, doi: 10.1177/15589250231220359.
- [2] A. Cieřlik and M. Ghodsi, "The impact of regulatory divergence in non-tariff measures on the cross-border investment of multinationals," *Emerging Markets Finance and Trade*, vol. 60, no. 15, pp. 3598-3637, 2024, doi: 10.1080/1540496X.2024.2356870.
- [3] E. B. Tirkolaee, A. Goli, H. Golpıra, and E. D. R. S. Gonzalez, "Sustainable global supply chain management from an international perspective," *Sustainability*, vol. 15, no. 16, Art. no. 12154, 2023, doi: 10.3390/su151612154.
- [4] E. Francesconi and G. Governatori, "Patterns for legal compliance checking in a decidable framework of linked open data," *Artificial Intelligence and Law*, vol. 31, no. 3, pp. 445-464, 2023, doi: 10.1007/s10506-022-09317-8.
- [5] P. Jiang and X. Cai, "A survey of semantic parsing techniques," *Symmetry*, vol. 16, no. 9, pp. 1201, 2024, doi: 10.3390/sym16091201.
- [6] K. Hämmerl, J. Libovický, and A. Fraser, "Understanding Cross-Lingual Alignment-A Survey," *arXiv e-prints*, vol. 2404, no. 1, Art. no. 06228, 2024, doi: 10.18653/v1/2024.findings-acl.649.
- [7] M. Burdon, A. Huggins, N. Godfrey, R. Simcock, J. Buckley, S. Slevin, et al., "From rules as code to mindset strategies and aligned interpretive approaches," *Journal of Cross-disciplinary Research in Computational Law*, vol. 2, no. 1, pp. 1-21, 2023.
- [8] J. Hernandez, D. Golpayegani, and D. Lewis, "An open knowledge graph-based approach for mapping concepts and requirements between the eu ai act and international standards," *AI and Ethics*, vol. 5, no. 1, pp. - 4474, 2025, doi: 10.1007/s43681-025-00708-6.
- [9] J. Zhu, J. Wu, X. Luo, and J. Liu, "Semantic matching based legal information retrieval system for COVID-19 pandemic," *Artificial Intelligence and Law*, vol. 32, no. 2, pp. 397-426, 2024, doi: 10.1007/s10506-023-09354-x.
- [10] B. Wen, T. Wang, J. Xu, Y. Liu, J. Li, and S. Lin, "File Compliance Detection Using a Word2Vec-Based Semantic Similarity Framework," *Informatica*, vol. 49, no. 18, pp. 51-66, 2025, doi: 10.31449/inf.v49i18.7421.
- [11] B. Li, H. Wen, S. Wang, T. Hu, X. Liang, and X. Luo, "An Optimized Semantic Matching Method and RAG Testing Framework for Regulatory Texts," *Electronics*, vol. 14, no. 14, pp. 2856, 2025, doi: 10.3390/electronics14142856.
- [12] O. A. Cejas, M. I. Azeem, S. Abualhaija, and L. C. Briand, "Nlp-based automated compliance checking of data processing agreements against gdpr," *IEEE Transactions on Software Engineering*, vol. 49, no. 9, pp. 4282-4303, 2023, doi: 10.1109/TSE.2023.3288901.
- [13] A. Fan, S. Wang, and Y. Wang, "Legal document similarity matching based on ensemble learning," *IEEE Access*, vol. 12, pp. 33910-33922, 2024, doi: 10.1109/ACCESS.2024.3371262.
- [14] C. Sai, L. Rossi, A. Damaratskaya, K. Winter, and S. R. Ma, "Identification and visual representation of explicit legal definitions, their relations and implicit actors in regulatory documents," *Computer Law & Security Review*, vol. 58, no. 1, Art. no. 106174, 2025, doi: 10.1016/j.clsr.2025.106174.
- [15] P. Prince Tritto and H. Ponce, "Causal Artificial Intelligence in Legal Language Processing: A Systematic Review," *Entropy*, vol. 27, no. 4, pp. 351, 2025, doi: 10.3390/e27040351.
- [16] J. Peng and X. Liu, "Automated code compliance checking research based on BIM and knowledge graph," *Scientific Reports*, vol. 13, no. 1, pp. 7065, 2023, doi: 10.1038/s41598-023-34342-1.
- [17] L. Waltersdorfer and M. Sabou, "Leveraging knowledge graphs for AI system auditing and transparency," *Journal of Web Semantics*, vol. 84, no. 1, Art. no. 100849, 2025, doi: 10.1016/j.websem.2024.100849.
- [18] P. Dario, M. Harel, and R. Peled, "Random-field random surfaces," *Probability Theory and Related Fields*, vol. 186, no. 1, pp. 91-158, 2023, doi: 10.1007/s00440-022-01179-0.
- [19] T. Brugger, M. Stürmer, and J. Niklaus, "MultiLegalSBD: A Multilingual Legal Sentence Boundary Detection Dataset," *arXiv e-prints*, vol. 2305, no. 1, Art. no. 01211, 2023, doi: 10.1145/3594536.3595132.
- [20] B. Li, Y. Fan, Y. Sataer, Z. Gao, and Y. Gui, "Improving semantic dependency parsing with higher-order information encoded by graph neural networks," *Applied Sciences*, vol. 12, no. 8, pp. 4089, 2022, doi: 10.3390/app12084089.
- [21] Y. Fan, B. Li, Y. Sataer, M. Gao, C. Shi, S. Cao, et al., "Hierarchical clause annotation: Building a clause-level corpus for semantic parsing with complex sentences," *Applied Sciences*, vol. 13, no. 16, pp. 9412, 2023, doi: 10.3390/app13169412.
- [22] J. Guo, Y. Cai, Y. Fan, F. Sun, R. Zhang, and X. Cheng, "Semantic models for the first-stage retrieval: A comprehensive review," *ACM Transactions on Information Systems (TOIS)*, vol. 40, no. 4, pp. 1-42, 2022, doi: 10.1145/3486250.
- [23] G. Bourahouat, M. Abourezq, and N. Daoudi, "Word embedding as a semantic feature extraction technique in arabic natural language processing: an overview," *Int. Arab J. Inf. Technol*, vol. 21, no. 2, pp. 313-325, 2024, doi: 10.34028/iajit/21/2/13.
- [24] S. Castano, A. Ferrara, E. Furioli, S. Montanelli, S. Picascia, D. Riva, et al., "Enforcing legal information extraction through context-aware techniques: The ASKE approach," *Computer Law & Security Review*, vol. 52, no. 1, Art. no. 105903, 2024, doi: 10.1016/j.clsr.2023.105903.
- [25] S. Castano, M. Falduti, A. Ferrara, Montanelli, and S., "A knowledge-centered framework for exploration and retrieval of legal documents," *Information Systems*, vol. 106, no. 1, Art. no. 101842, 2022, doi: 10.1016/j.is.2021.101842.
- [26] L. Vogt, T. Kuhn, and R. Hoehndorf, "Semantic units: Organizing knowledge graphs into semantically meaningful units of representation," *Journal of Biomedical Semantics*, vol. 15, no. 1, pp. 7, 2024, doi: 10.1186/s13326-024-00310-5.
- [27] H. Zhong, D. Yang, S. Shi, L. Wei, and Y. Wang, "From data to insights: The application and challenges of knowledge graphs in intelligent audit," *Journal of Cloud Computing*, vol. 13, no. 1, pp. 114, 2024, doi: 10.1186/s13677-024-00674-0.
- [28] M. Y. Jaradeh, K. Singh, M. Stocker, A. Both, and S. Auer, "Information extraction pipelines for knowledge graphs," *Knowledge and Information Systems*, vol. 65, no. 5, pp. 1989-2016, 2023, doi: 10.1007/s10115-022-01826-x.
- [29] G. Liu, C. Jin, L. Shi, C. Yang, J. Shuai, and J. Ying, "Enhancing Cross-Lingual Entity Alignment in Knowledge Graphs through Structure Similarity Rearrangement," *Sensors*, vol. 23, no. 16, pp. 7096, 2023, doi: 10.3390/s23167096.
- [30] Y. Liu, S. Chen, and Y. Yang, "Semantic alignment: A measure to quantify the degree of semantic equivalence for English-Chinese translation equivalents based on distributional semantics," *Behavior Research Methods*, vol. 57, no. 1, pp. 51, 2025, doi: 10.3758/s13428-024-02527-9.
- [31] Y. Yang and R. H. Baayen, "Comparing the semantic structures of lexicon of Mandarin and English," *Language and Cognition*, vol. 17, no. 1, pp. e10, 2025, doi: 10.1017/langcog.2024.47.
- [32] X. Zhang, W. Zhang, and H. Wang, "Cross-language entity alignment based on dual-relation graph and neighbor entity screening," *Electronics*, vol. 12, no. 5, Art. no. 1211, <https://doi.org/10.3390/electronics12051211>, 2023.

- [33] P. Galuščáková, D. W. Oard, and S. Nair, "Cross-language Retrieval," *Information Retrieval: Advanced Topics and Techniques*, vol. (1), pp. 321-357, 2024, doi: 10.1145/3674127.3674136.
- [34] X. Li and K. Zhang, "Heterogeneous Graph Neural Network with Multi-View Contrastive Learning for Cross-Lingual Text Classification," *Applied Sciences* (2076-3417), vol. 15, no. 7, pp. 3454, 2025, doi: 10.3390/app15073454.
- [35] L. Bai, N. Li, G. Li, Z. Zhang, and L. Zhu, "Embedding-based entity alignment of cross-lingual temporal knowledge graphs," *Neural Networks*, vol. 172, no. 1, Art. no. 106143, 2024, doi: 10.1016/j.neunet.2024.106143.