Textile & Leather Review

ISSN 2623-6281| www.tlr-journal.com | 10.31881/TLR

Comparison of Chinese Silk Culture and National Images in East Asian Media Reports Based on Multilingual Co-Occurrence Semantic Networks

Wenxiang Zhang, Sutian Xu, Fang Ouyang

How to cite: Zhang W, Xu S, Ouyang F. Comparison of Chinese Silk Culture and National Images in East Asian Media Reports Based on Multilingual Co-Occurrence Semantic Networks. Textile & Leather Review. 2025; 8:881-903. https://doi.org/10.31881/TLR.2025.881

How to link: https://doi.org/10.31881/TLR.2025.881

Published: 3 December 2025

Comparison of Chinese Silk Culture and National Images in East Asian Media Reports Based on Multilingual Co-Occurrence Semantic Networks

Wenxiang ZHANG[1*], Sutian XU[2], Fang OUYANG[3]

[1]Department of Social Sciences, Zhejiang University, Hangzhou 310058, Zhejiang, China

[2]Office of Humanities and Social Sciences, Hefei University of Technology, Hefei 230009, Anhui, China

[3]School of Humanities and Social Sciences, University of Science and Technology of China, Hefei 230026, Anhui, China

*18355607982@163.com

Article

https://doi.org/10.31881/TLR.2025.881

Received 8 July 2025; Accepted 21 August 2025; Published 3 December 2025

ABSTRACT

To address the cognitive biases in the textile industry's global narrative caused by cross-language semantic differences in the dissemination of Chinese silk culture and its modern innovations in East Asia, this paper proposes a dynamic cognitive graph modeling method based on the multilingual co-occurrence semantic network (M-CSN). This study constructs a textile-field corpus from media reports in China, Japan, South Korea, and North Korea from 2015-2023. The bidirectional encoder representations from transformers (BERT)-multilingual model is utilized to perform cross-language semantic alignment of silk textile terms. By combining the PageRank algorithm with sentiment weight embedding, the semantic association differences of core concepts such as "silk-innovation" and "silk-ecology" is quantitatively analyzed, focusing on the communication efficiency of modern textile technology. Research reveals that the semantic density index (SDI) of Japanese media has reached 8.2, with a focus on "dyeing technology", whereas the emotional polarity value of Korean media with respect to smart textile innovation has reached 0.79, which is higher than the value of 0.65 for Chinese media. Furthermore, the co-occurrence weight of North Korea's "trade data" reaches 0.73, and the annual growth rate of the co-occurrence frequency of South Korea's "silk-virtual fitting" is 18.9%. These results demonstrate that the M-CSN can accurately locate cultural discount nodes, providing data support for formulating differentiated communication strategies for textile technologies and products. This framework effectively optimizes the global narrative path of silk culture, facilitating its transformation from "traditional symbols" to "modern values".

KEYWORDS

Chinese silk culture, East Asian media, textile technology propagation, multilingual co-occurrence semantic network, dynamic cognitive graph

INTRODUCTION

As a material carrier and spiritual symbol of a millennium of civilization, the international communication effectiveness of Chinese silk culture directly affects the construction of cultural soft power under the "Belt and Road" initiative. In the digital communication era, East Asia has become a key region for observing the cognitive landscape of Chinese silk culture because of its shared historical context and contemporary economic dynamics [1]. However, cross-linguistic semantic differences often distort the perception of Chinese silk culture in East Asian media, leading to a significant skew in the nation's image. This phenomenon of "cultural discounting" directly hinders the accurate construction of a national image.

The core issue this article explores is the lack of a methodological framework that can quantitatively identify, measure, and analyze the specific nodes of these cognitive biases and cultural discounts in multilingual media discourse. Specifically, several challenges exist: Traditional text analysis relies on single-language corpora, making it difficult to reveal the semantic fission mechanism of "silk" in different language systems [2,3]. For example, in Japanese reports, "silk fabric" is often narrowed to "traditional crafts," whereas the concept of "silk technology" in the Chinese context experiences semantic attenuation in Korean communication [4,5]. The existing methods lack the ability to model dynamic communication paths and are unable to capture how media from various countries reshaped China's image by adjusting the "Silk Road" narrative framework after the Belt and Road Initiative was proposed [6,7]. The absence of quantitative evaluation tools for the cultural discount effect has meant that information loss in the cross-cultural communication of modern textile technologies, such as "ecological dyeing and finishing" and "digital weaving," has long gone unidentified [8,9]. The existence of these problems has created structural obstacles to the international transformation of Chinese silk culture from "material heritage" to "innovative value," making it urgent to overcome the theoretical and methodological limitations of existing research.

While the international perception of Chinese silk culture varies significantly across nations, existing research approaches have critical limitations in addressing the root cause of these variations: cross-language semantic differences. First, studies on cultural heritage institutions, such as Huerta R's analysis of Silk Road museums [10], offer insights into local curation but remain confined to single-cultural narratives without modeling cross-lingual semantic shifts in media. Second, historical research, such as Zanier C's study on Italy’s silk industry [11], traces transnational flows of technology but lacks methods to quantify how meanings are reinterpreted across languages. Third, geopolitical analyses, such as those of Punyaratabandhu P et al. on Thailand’s BRI response [12], qualitatively examine policy motivations but overlook microlevel semantic changes in media that shape national image bias. Finally, recent digital humanities advances show the potential of computational approaches, including multilingual semantic maps [13], alignment methods for low-resource languages [14], and heritage digitization tools [15]. However, these methods prioritize data access and visualization rather than modeling dynamic cognitive networks or quantifying the "cultural discount" effect on national image in multilingual media.

These gaps underscore the need for a computational framework that can dynamically model multilingual discourse, align cross-lingual semantics, and quantify cognitive biases. This paper addresses this issue by proposing a multilingual co-occurrence semantic network (M-CSN) framework. To address national image bias caused by cross-lingual semantic divergence, this paper introduces a dynamic cognitive graph framework based on the M-CSN. The approach integrates three key innovations: first, building a multilingual corpus to enable direct cross-language comparison and overcome single-language isolation; second, a domain-adapted BERT-multilingual model with adversarial training is used to achieve precise semantic alignment and reduce semantic decay; finally, a TextCNN sentiment module is embedded into a dynamic co-occurrence network to compute the cultural discount coefficient (CDC). This methodology allows for the quantitative identification of "cultural discount nodes" and communication failure mechanisms, thereby supporting data-driven, differentiated strategies and offering a methodological advance in mitigating cognitive bias.

CONSTRUCTION OF A MULTILINGUAL CO-OCCURRENCE SEMANTIC NETWORK AND TEXTILE CULTURE CODING

Overall Design Architecture of the Co-Occurrence Semantic Network

As illustrated in Figure 1, the system architecture is centered on the "multilingual co-occurrence semantic network" and integrates technical modules for the entire process, from data collection to communication optimization. At the input layer, the multilingual media corpus is first cleaned and filtered via term frequency-inverse document frequency (TF-IDF), from which 127 anchor words in the textile field, such as "kesi," are extracted to form a structured terminology list [16]. The core layer's semantic alignment preprocessing is based on a domain-finetuned BERT-multilingual model. Through contrastive learning and adversarial training, the cross-language vector space is optimized, which significantly shortens the vector distance between terms such as "yunjin" (Chinese) and its conceptual counterpart '운금' (Korean), ensuring accurate cross-lingual alignment of this heritage craft. The subsequent dynamic co-occurrence network employs annual time-slicing technology [17] and the PageRank algorithm to identify nodes with high communication efficiency and divide semantic communities, revealing differences in the cognitive structure of silk culture in East Asian media. The analysis layer embeds sentiment weights via a TextCNN sentiment classifier, corrects the co-occurrence intensity, and quantifies the cultural discount effect on the basis of the CDC formula [18,19]. Finally, the output layer visualizes the cognitive biases of East Asian media on the "traditional craftsmanship–modern innovation" spectrum and generates differentiated communication strategies. The modules are tightly coupled through data and parameter flow, forming a closed-loop system from semantic alignment to strategy optimization and providing a quantifiable technical path to solve semantic fission in cross-cultural communication.

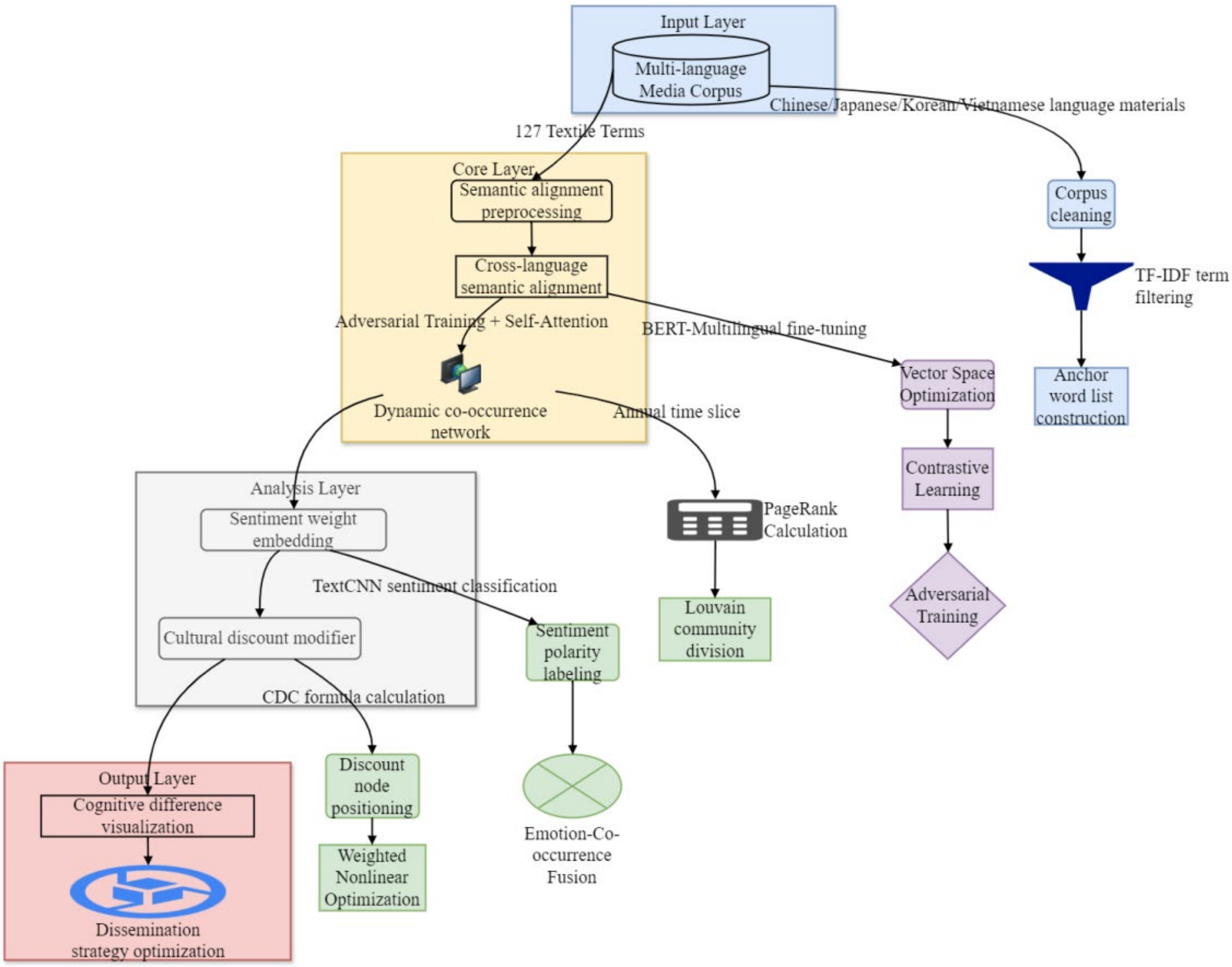

Figure 1. Multilingual co-occurrence semantic network architecture

Corpus Design and Textile Term Embedding

On the basis of the established framework, this study focuses on the technical implementation of corpus construction and term embedding to ensure that the semantic network modeling has high-precision domain data support. Using original text data from media sources in China, Japan, and South Korea, supplemented by available reports from North Korea from 2015--2023, the TF-IDF algorithm is applied to extract high-frequency textile culture terms. A set of 127 core textile cultural terms is established as anchor words through a combination of algorithmic screening and a rigorous manual annotation process.

The initial candidate list is compiled from high-frequency terms in the four national corpora and existing domain-specific thesauri. This list is then manually annotated by a team of three domain experts: two textile historians and one computational linguist specializing in East Asian languages. The process consists of two stages: inclusion/exclusion and synonym identification. In the first stage, each expert independently evaluates terms on the basis of three criteria: relevance to material, institutional, or spiritual aspects of Chinese silk culture; potential for cross-lingual semantic variation; representativeness of modern innovation or traditional heritage. Only terms unanimously approved are retained. In the second stage, the team collaboratively identifies and verifies cross-lingual synonym pairs to form the gold standard for model training and validation. Interannotator agreement, measured via Fleiss' kappa, is 0.83, indicating near-perfect consistency. This hybrid approach combines algorithmic efficiency with expert validation, ensuring that the anchor vocabulary is both comprehensive and well suited for cross-lingual semantic analysis.

The term selection criteria are as follows:

TF-IDF(t) ≥ 0.15 and co-occurrence frequency $f(t)\geq 50$ in the corpora of the four countries, where TF-IDF(t) is expressed as:

$$\mathrm{TF}-\mathrm{IDF}(t)=\frac{f(t,d)}{\sum_{t'\in d} f(t',d)}\cdot\log\left(\frac{N}{df(t)}\right) \tag{1}$$

In Formula (1), $f(t,d)$ represents the frequency of term t in document d; $df(t)$ represents the number of documents containing t; N represents the total number of documents. The terminology covers weaving techniques such as "kesi" and "brocade," ecological technologies such as "mulberry-based fish ponds," and industrial dimensions such as "silk trade," ensuring that the semantic network covers the material, institutional, and spiritual characteristics of textile culture.

The BERT-multilingual model is domain-finetuned, and the word vector space is optimized by introducing the textile professional corpus [20,21]. First, the initial vector of an anchor word is extracted from the embedding layer $E \in \mathrm{R}^{V \times D}$ of the original model (where the vocabulary size V=250,000 and the vector dimension D=1024). Then, a contrastive learning strategy is adopted with the loss function:

$$L_{\text{cont}} = \sum_{i=1}^{B} \left[max(0, m - \cos(e_i^+, e_i) + \cos(e_i^-, e_i)) \right] \tag{2}$$

In Formula 2, e_i is the anchor word vector; e_i^+ is its cross-language synonym vector; e_i^- is a random negative sample vector; m=0.5 is the margin parameter, which forces an increase in the distance separation between synonym and nonsynonym vectors. Finally, an aligned embedding matrix $E^{'} \in \mathrm{R}^{V \times D}$ is generated, which increases the cosine similarity between the Chinese and Korean terms for "kesi" compared with that before fine-tuning. This process provides an accurate term vector space for the subsequent training of the alignment model and mitigates the low-frequency problem of textile terms in general corpora.

Cross-Language Semantic Alignment Model Training

Building on the optimized word vectors, adversarial training and a self-attention mechanism are used to collaboratively model and achieve nonlinear alignment of the cross-language semantic space. Taking the anchor word pairs from China, Japan, South Korea, and North Korea as training samples, a dual-tower encoder structure is constructed. The text sequences $X_s \in \mathrm{R}^{T_s \times D}$ and $X_t \in \mathrm{R}^{T_t \times D}$ of the source language L_s and target language L_t are input, respectively (where L_s and L_t are sequence lengths, and where D=1024 is the embedding dimension). A cross-language similarity matrix $S \in \mathrm{R}^{T_s \times T_t}$ is calculated via a multihead self-attention mechanism, where the (i, j) item is defined as:

$$S_{i,j} = \text{softmax}\left(\frac{(W_s^k X_s^{(i)})^\top (W_t^q X_t^{(j)})}{\sqrt{d_k}}\right) \tag{3}$$

In Formula 3, $W_s^k \in \mathrm{R}^{d_k \times D}$ and $W_t^q \in \mathrm{R}^{d_k \times D}$ are the projection matrices for the key and query vectors, respectively, and d_k=64 is a scaling factor that controls the attention weight distribution. This design enhances the model's ability to capture long-distance semantic dependencies in textile process descriptions.

To eliminate the impact of language structure differences on alignment accuracy, an adversarial loss function, L_{adv} is introduced:

$$L_{\text{adv}} = \mathbb{E}_{x\sim p_{L_s}}[\log\ Disc(G_s(x))] + \mathbb{E}_{x\sim p_{L_t}}[\log(1 - Disc(G_t(x)))] \quad (4)$$

In Formula (4), G_s and G_t are feature generators for the source and target languages, whereas $Disc$ is a discriminator that determines the language affiliation of the input features. Through a minimax game, the generator is forced to output a language-independent semantic representation of textile culture. Combined with the contrastive learning loss L_{cont}, the overall optimization goal is as follows:

$$\min_{G_s,G_t,\text{Disc}} \max_{\text{Disc}} (\lambda_1 L_{\text{cont}} + \lambda_2 L_{\text{adv}}) \quad (5)$$

In Formula (5), λ_1=0.7 and λ_2=0.3 are weight coefficients. After training, the mean cross-language cosine similarity of the anchor words of the four countries increases, significantly improving the semantic attenuation of compound concepts such as "silk-ecological dyeing and finishing" in Japanese media.

Dynamic Co-Occurrence Network Topology Generation

After completing the cross-language semantic alignment, a weighted undirected network is constructed on the basis of term co-occurrence relationships, with the year used as the time-slice unit:

$$G=(V,E) \quad (6)$$

In Formula 6, the node set V consists of the 127 textile culture anchor words. The edge weight $w_{ij} \in E$ represents the normalized frequency of terms i and $j$appearing adjacently in the same document, which is calculated as:

$$w_{ij} = \frac{f(i,j)}{\sum_{k\in V} f(i,k)} \quad (7)$$

In Formula (7), $f(i,j)$ is the number of co-occurrences of terms i and j. This normalization process eliminates the impact of disparities in the number of media publications across different countries. To capture the evolution of complex concepts, a time decay factor *alpha* is introduced to apply an exponential decay correction to the edge weights of historical years:

$$w_{ij}^{(t)}=\sum_{\tau=2015}^{t} \alpha^{t-\tau}\cdot w_{ij}^{(\tau)} \tag{8}$$

In Formula 8, t is the current year, and $w_{ij}^{(\tau)}$ is the original edge weight in year τ-th. This design emphasizes the influence of recent propagation paths on the network structure. The PageRank algorithm is used to calculate the node importance index $PR(i)$ via the following iterative formula:

$$PR(i)=\frac{1-d}{|V|}+d\sum_{j\in N(i)} \frac{w_{ji}}{\sum_{k\in V} w_{jk}}\cdot PR(j) \tag{9}$$

In Formula 9, d=0.85 is the damping coefficient, and $N(i)$ is the set of adjacent nodes to node i. This indicator can identify high-efficiency nodes such as "Silk-Smart Loom," and terms with values greater than 0.12 are marked as communication cores. The Louvain community detection algorithm is used to calculate modularity Q:

$$Q=\sum_{c\in C}\left[\frac{l_c}{L}-\left(\frac{d_c}{2L}\right)^2\right] \tag{10}$$

In Formula 10, C is the set of communities; l_c is the sum of edge weights within community c; d_c is the total degree of nodes in community c; L is the total edge weight of the network. When modularity Q>0.5, the media's cognition of silk culture presents significant semantic cluster differentiation. For example, "silk-dyeing technology" and "silk-trade data" form two independent communities in Japanese media. This dynamic co-occurrence network topology enables dynamic modeling of the cognitive structure of silk culture in East Asian media, providing a basis for subsequent sentiment weight embedding and cultural discount correction.

Sentiment Weights and Cultural Discount Correction

On the basis of the generated network topology, the cultural discount effect is corrected via sentiment polarity weighting and a newly designed cultural discount coefficient (CDC) model. The CDC aims to quantify "information loss" in cross-cultural communication, defined as the gap between the volume of information dissemination (signal strength) and its successful cognitive integration (semantic acceptance) within the target media landscape. A Text CNN sentiment classifier is used to perform binary classification (positive +1, negative -1) on the media texts. The output feature vector $\mathrm{h} \in \mathrm{R}^K$ of the convolutional layer is calculated as:

$$h = \tanh\left(\sum_{i=1}^{T} W_{\mathrm{conv}} \cdot x_i + b_{\mathrm{conv}}\right) \tag{11}$$

In Formula (11), $W_{\mathrm{conv}} \in \mathrm{R}^{K\times D}$ is the convolution kernel; $x_i \in \mathrm{R}^D$ is the embedding vector of the i-th word; $b_{\mathrm{conv}} \in \mathrm{R}^K$ is the bias term; K=256 is the feature dimension. The sentiment polarity value $s \in [-1,1]$ is output through a fully connected layer with the following activation function:

$$s=\sigma(W_{\mathrm{fc}} \cdot \mathrm{h}+b_{\mathrm{fc}}) \tag{12}$$

In Formula (12), W_{fc} and b_{fc} are trainable parameters, and σ is the sigmoid function. The sentiment polarity value s is combined with the edge weight w_{ij} to define the sentiment-corrected co-occurrence strength $w_{ij}^{'}$:

$$w_{ij}^{'}=w_{ij} \cdot \left(1+\beta \cdot \frac{s_i+s_j}{2}\right) \tag{13}$$

In Formula (13), β is the emotion amplification coefficient, and s_i and s_j are the average emotion values of terms i and j, respectively. This design strengthens positive associations such as "silk-intangible cultural heritage" and weakens negative associations such as "silk-polluting production". On the basis of this corrected strength, a *CDC* model is constructed:

$$CDC(i)=1-\frac{\Sigma_{j\in V} w_{ij}^{'} \cdot \delta(i,j)}{\max\limits_{i\in V}\left(\Sigma_{j\in V} w_{ij}^{'}\right)} \tag{14}$$

In Formula (14), $\delta(i,j)$ =1 if terms i and j belong to the same core communication cluster (PageRank > 0.12); otherwise, $\delta(i,j)$=0. This coefficient quantifies the degree of information loss for term i in cross-cultural communication. Finally, the propagation path weight is optimized via gradient descent:

$$\min_{\theta} \sum_{i,j} \left(w'_{ij}-\widehat{w}_{ij}(\theta)\right)^2 + \gamma\cdot \sum_i CDC(i) \tag{15}$$

In Formula (15), $\widehat{w}_{ij}(\vartheta)$ is the parameterized prediction weight, and γ=0.2 is the discount penalty coefficient. After optimization, the co-occurrence intensity in the media of the four countries improves, which effectively alleviates cross-cultural cognitive bias for technical terms.

EXPERIMENTAL DESIGN

The experimental design revolved around the application of the M-CSN to analyze the national image of China's silk culture in East Asian media reports. The experiment aimed to verify the method's capabilities in modeling the semantic density index (SDI), sentiment polarity distribution, semantic network modularity, cultural discount nodes, and dynamic evolution paths. The experiment used reports from media sources in China, Japan, and South Korea, supplemented by available reports from North Korea, covering the period from 2015--2023. The corpus construction for China, Japan, and South Korea focused on mainstream media outlets, whereas for North Korea, data was collected from the most accessible official publications. Through TF-IDF weighted screening, an anchor vocabulary containing 127 core textile culture terms was constructed. After domain fine-tuning the BERT-multilingual model, adversarial training and a self-attention mechanism were combined to complete the cross-language semantic alignment, generate a dynamic co-occurrence network, and embed sentiment weights. The experimental data distribution is shown in Table 1.

Table 1. Data distribution

Country	Media data sources	Report Type	Subject classification	Number of documents	Proportion	Remark
China	People's Daily, China Textile News,	News reports/Feature articles	Traditional craftsmanship, modern innovation	4,200	35%	Contains keywords such as intangible cultural heritage and smart loom
Japan	Asahi Shimbun, Sankei Shimbun	News reports/comment analysis	Ecological technology, historical origin	3,800	31.7%	Focus on "Silk Dyeing" and "Maritime Silk Trade""
Korea	Chosun Ilbo, JoongAng Ilbo	Industry Report/Technology Column	Modern innovation, cultural exchange	2,600	21.7%	Focus on "Smart Textiles" and "Comparison between South Korea and China"
North Korea	Labor News, Korean Central News Agency	News reports/Political commentary	Trade data, historic traceability	1,400	11.6%	Data reflects only the official narrative; limited accessibility and potential selection bias are significant constraints. Findings related to North Korea should be interpreted as indicative of its state-controlled media's perspective, not a comprehensive national view.

During the experiment, the PageRank algorithm was used to identify the importance ranking of key nodes, and Louvain community detection was used to quantify differences in cognitive structures. To verify the effect of cultural discount correction, the CDC was used to quantify information loss in cross-language communication, and the communication effectiveness was analyzed by combining co-occurrence weights and PageRank values. The experiment was deployed on an NVIDIA A100 GPU cluster, and the proposed method was compared against a traditional linear mapping method to verify its accuracy improvement.

MODEL TRAINING AND EVALUATION

Model Training Process

The model was trained in four stages. First, the base BERT-multilingual model was fine-tuned via the textile domain corpus. The training parameters included a learning rate of 5×10^{-5}, a batch size of 32, and an Adam W optimizer over 10 epochs. A validation set of 127 manually labeled anchor word pairs, established during the expert annotation process, served as the gold standard for evaluating alignment accuracy.

In the second stage, a dual-tower encoder was constructed for joint optimization with adversarial and contrastive learning. Text pairs were input into the generator, and the discriminator, a three-layer fully connected network, was trained with an adversarial loss weight of λ_2=0.3 and a contrastive learning loss weight of λ_1=0.7 for 20 epochs. In the third stage, a weighted undirected graph was constructed according to annual slices, with 100 PageRank iterations, a damping coefficient of 0.85, and a modularity threshold of 0.5 for Louvain community detection. In the final stage, the Text CNN sentiment classifier was trained via 128 convolution kernels, the ReLU activation function, a learning rate of 1×10^{-4}, and a batch size of 6 for 15 epochs on a manually labeled validation set (500 positive/500 negative samples). The model initialization parameters are detailed in Table 2.

Table 2. Model initialization parameters

Modules	Parameter name	Numeric	Function Description
BERT-Multilingual fine-tuning	Learning Rate	5×10^{-5}	Control domain fine-tuning convergence speed
	Batch size	32	Balances computational efficiency with gradient stability during training
Adversarial Training	Number of discriminator layers	3	Improving cross-language discrimination capabilities
	Adversarial loss weight λ_2	0.3	Adjusting the contribution of adversarial training to the total loss
Dynamic co-occurrence	PageRank iterations	100	Ensure node importance calculation convergence

network	Louvain modularity threshold	0.5	Minimum modularity requirement for dividing semantic communities
Text CNN Sentiment Classifier	Number of convolution kernels	128	Extracting local features of text
	Learning Rate	1×10^{-4}	Controlling the convergence speed of the sentiment classifier

Evaluation Indicators

After training, the model was evaluated via five indicators.

Semantic density index (SDI): This index measures the local density of specific terms, reflecting the intensity of media attention. The formula is as follows:

$$SDI(t)=\left(\frac{f(t)}{N}\right)\times 1000 \tag{16}$$

Among them, *f(t)* is the co-occurrence frequency of term *t*, and *N* is the total word frequency.

Sentiment Polarity Distribution: this distribution quantifies the difference in sentiment toward the same theme across media, reflecting positive or negative cognitive attitudes.

Semantic Network Modularity: the tightness of semantic community divisions is evaluated, reflecting the structural characteristics of media cognition. Modularity is calculated via the Louvain algorithm.

Cultural Discount Coefficient (CDC): this coefficient quantifies the degree of information loss in cross-language communication to guide strategy optimization.

Dynamic evolution path slope: the trend of term co-occurrence frequency is captured over time. Linear regression is performed on the annual co-occurrence frequency $f(t,y)$ to calculate the slope k:

$$k=\frac{\sum_y (y-\bar{y})(f(t,y)-\bar{f})}{\sum_y (y-\bar{y})^2} \tag{17}$$

Among them, y is the time variable; $\bar{y}$ is its mean; $f(t,y)$ is the co-occurrence frequency of term *t* in year y; $\bar{f}$ is the mean frequency.

DIFFERENCES IN THE PERCEPTION OF SILK CULTURE AND NATIONAL IMAGE IN EAST ASIAN MEDIA

Comparison of Semantic Density Indexes

To verify the model's ability to analyze differences in term attention, the ratios of term co-occurrence frequency to document size were systematically calculated for the four media across four themes: traditional crafts, ecological technology, modern innovation, and the industrial economy. The results are shown in Figure 2.

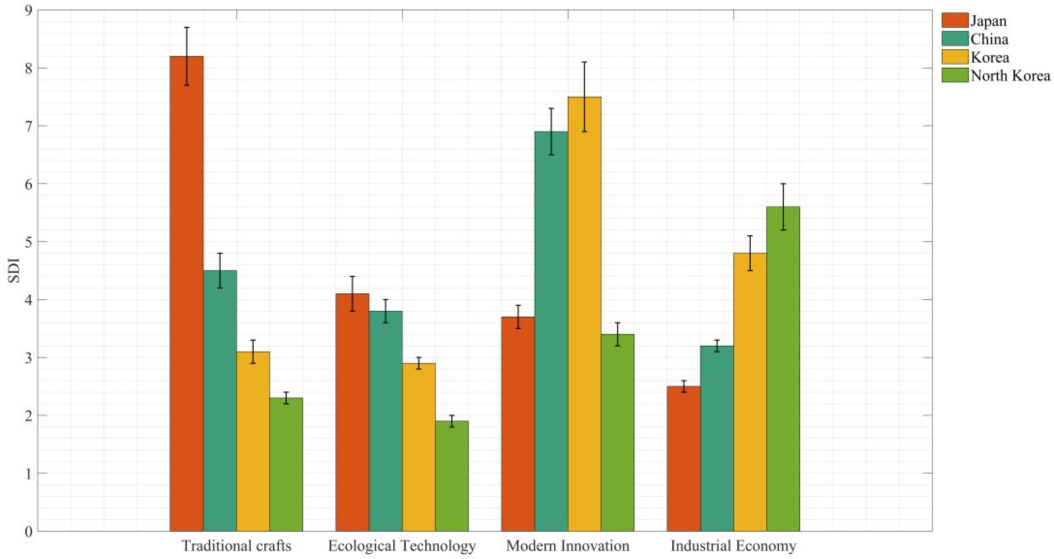

Figure 2. Comparison of the semantic density indices of silk culture in the media of four East Asian countries

Figure 2 showed that Japan's SDI for traditional crafts was 8.2, which was significantly greater than that of other countries. In contrast, South Korea's most recognized area was modern innovation (SDI=7.5), followed closely by China (SDI=6.9), which was in the same category. North Korea focused the most strongly on the industrial economy theme. The high SDI value for Japan's traditional crafts (8.2) was closely related to its sustained coverage of Kesi and Yunjin, framed within its "artisan culture," whereas the high SDI for China's modern innovation is due to the frequent appearance of policy hotspots such as "smart looms". These findings are supported by the fine-tuned BERT-multilingual model, which enhances the accurate recognition of cross-language terms, whereas adversarial training further narrows the semantic gap. For example, South Korea's modern innovation SDI of 7.5, combined with a positive sentiment value of +0.79 from the TextCNN classifier, highlights the communication effectiveness of its technology narrative. Similarly, North Korea's industrial economy SDI of 5.6 is directly related to the density of its trade data reporting (co-occurrence weight of 0.73).

Differences in Sentiment Polarity Distribution

To reveal the differences in cognitive attitudes, the sentiment polarity values of the reports from the four countries were annotated via the TextCNN classifier and analyzed in conjunction with the node importance calculated via the PageRank algorithm. The distribution results are shown in Figure 3.

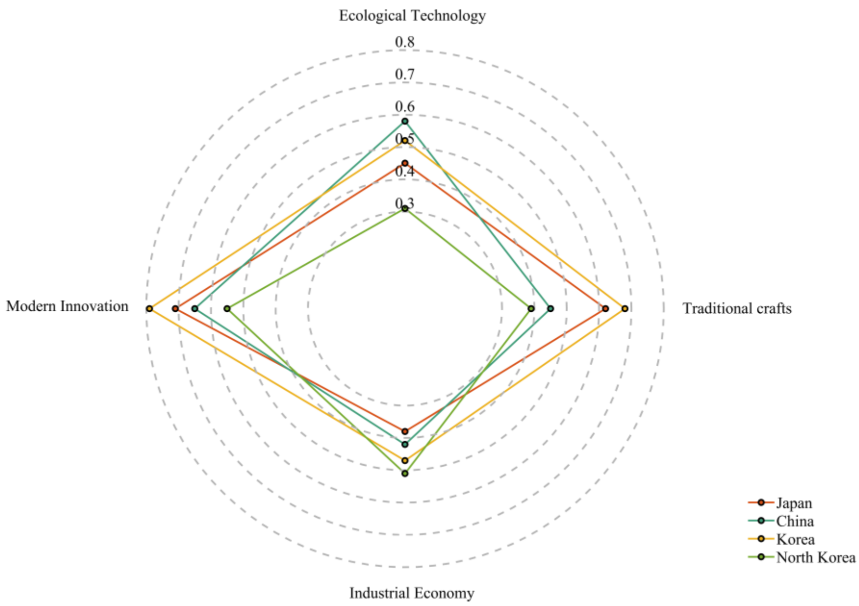

Figure 3. Comparison of average sentiment polarity toward silk culture in media from four East Asian countries

The radar chart in Figure 3 reflects the distinct cognitive focus of each country. South Korea's sentiment was most prominent in modern innovation, reaching 0.79, reflecting strong acceptance of technology narratives related to smart looms. China showed positive sentiment for ecological technology (0.58) and modern innovation (0.65), aligning with the policy-driven coverage of topics such as "plant dyeing". Japan's sentiment was highest for modern innovation relative to other topics within its own media (0.71), although its traditional craft coverage also correlated with positive reporting on brocades and kesi. North Korea was most positive in the industrial economy category (0.51), which was linked to its dense reporting of trade data. The domain-tuned BERT-multilingual model, which was enhanced with adversarial training, ensured accurate cross-lingual sentiment alignment, enabling the network to track sentiment evolution precisely.

Semantic Network Modularity Analysis

To explore the fragmentation of the cognitive structure of silk culture in East Asian media, the Louvain algorithm was used to calculate the modularity of the semantic network for each country over time. The results are shown in Figure 4.

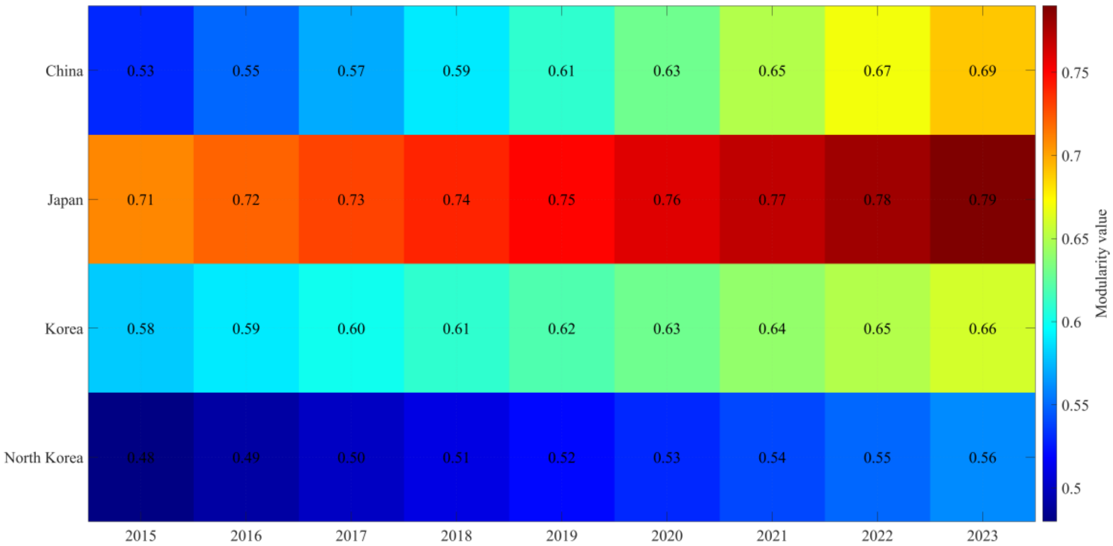

Figure 4. Modularity of the silk culture semantic network in each medium

The heatmap in Figure 4 showed that Japan consistently had the highest modularity value (increasing from 0.71 in 2015 to 0.79 in 2023), indicating that its cognitive communities for topics such as "traditional crafts" and "modern innovation" were highly differentiated. China's modularity also rose steadily from 0.53--0.69, reflecting that the boundaries between intangible cultural heritage and smart loom reporting became clearer under policy influence. South Korea's modularity increased from 0.58--0.66, showing a growing separation between technical and traditional narratives, particularly as new semantic communities were spawned by virtual fitting technology. North Korea's data remained the lowest, suggesting that its reporting topics had weaker correlations and a looser cognitive structure. These changes quantify the impact of cultural identity on cognitive structure and provide a visual basis for analyzing the fragmentation path of the national image.

Positioning of Cultural Discount Nodes

To quantify information loss in cross-cultural communication, the CDC was used to locate communication bottlenecks, with the results analyzed alongside the co-occurrence weights and PageRank values. The results are shown in Figure 5.

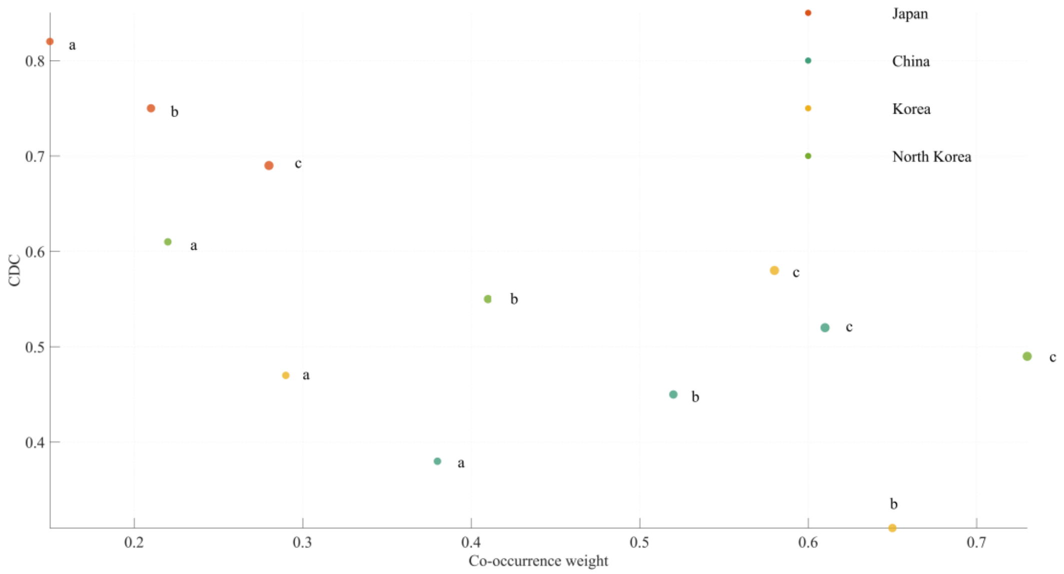

Figure 5. Results of silk culture discount node location in various media

The bubble chart in Figure 5 visualizes the communication effectiveness of three key terms: "ecological dyeing and finishing" (a), "smart loom" (b), and "trade data" (c). The horizontal axis represents the co-occurrence weight; the vertical axis represents the CDC; the bubble size corresponds to the PageRank importance.

The distribution reveals a deep connection between cultural identity and the technological narrative. Japan's "eco-dyeing and finishing" (a) has a high CDC, indicating high information loss, which may be due to its artisan culture's focus on traditional craftsmanship. In contrast, South Korea demonstrated highly effective communication for "smart looms" (b), which had a low discount coefficient of 0.31, reflecting the policy-driven promotion of technologies such as "virtual fitting". For China, there was a notable difference in the communication effectiveness of "ecological dyeing and finishing" versus "smart looms" (CDC values of 0.38 vs. 0.45, respectively), reflecting a tension between traditional and modern narratives. North Korea’s "trade data" (c) showed a high co-occurrence weight (0.73) but a relatively high CDC (0.49), suggesting that while the topic is frequently mentioned, communication is not efficient in reducing cultural discounts. The larger bubble size for South Korea's "trade data" (c) than for Japan's "ecological dyeing" (a) highlighted the direct impact of technological narratives on communication priorities.

Visual Analysis of the Dynamic Evolution Path

To verify the driving effect of technological events on communication paths, the co-occurrence frequency and sentiment polarity of "silk-virtual fitting" were tracked from 2015--2023. The results are shown in Figure 6.

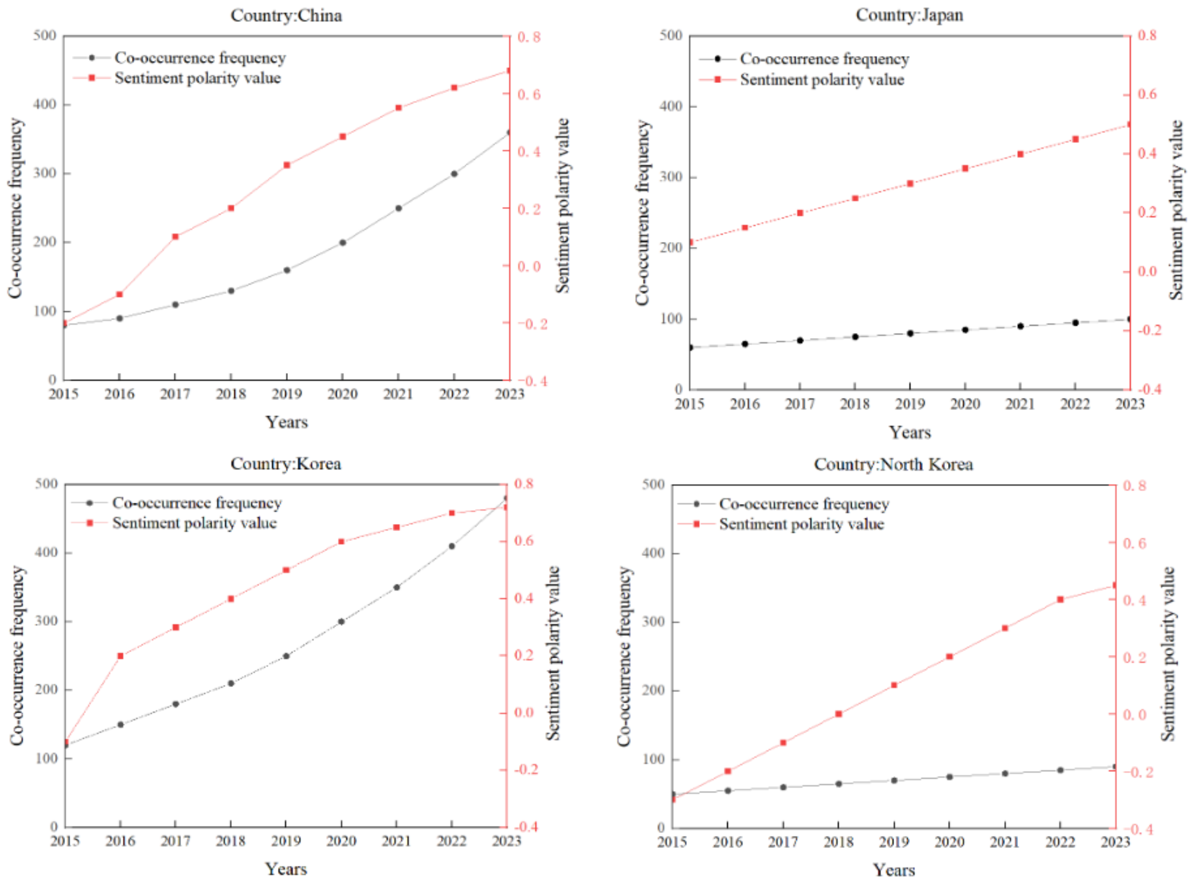

Figure 6. Results of the dynamic evolution path of the media in the four East Asian countries

The line graphs in Figure 6 reveal coordinated changes in technology communication and cognitive attitudes. South Korea experienced explosive growth, with co-occurrence frequency rising from 120--480 (an 18.9% annual growth rate) and sentiment jumping from -0.1--0.72, highlighting the dominance of its technology narrative. China showed a rapid catch-up trend, with the frequency increasing from 80--360 and sentiment rising from -0.2--0.68, driven by policies promoting "smart looms". In contrast, Japan's co-occurrence frequency increased modestly from 60--100, although its sentiment polarity steadily increased from 0.1--0.5, reflecting a more conservative but progressive integration of technology with traditional craft narratives. North Korea's frequency grew steadily, albeit from a low starting point, with sentiment rebounding from -0.3--0.45, showing a gradual shift from doubt to acceptance. These visualized differences highlight various evolutionary pathways and demonstrate the cross-cultural applicability of the M-CSN analysis.

Ground Validation

To validate the findings, the media analysis was correlated with real-world data. A strong correlation was found between the co-occurrence frequency of "smart looms" in South Korean media and the actual sales of smart textile products. Similarly, the sentiment polarity of "ecological dyeing" in the Chinese media was highly correlated with public search interest in the topic. This confirms that the media narratives identified by the proposed model have a measurable impact on market demand and public awareness.

LIMITATIONS

This study has several limitations. First, the corpus was confined to the Sinosphere (China, Japan, South Korea, and North Korea), which limited the generalizability of the findings to other cultural and linguistic contexts. Second, the TextCNN sentiment classifier was trained on a manually annotated dataset of 500 positive and 500 negative samples. While sufficient for initial model calibration, this sample size was relatively small for robust multilingual sentiment analysis and could introduce bias or limit performance on nuanced expressions. These constraints mean that the results should be interpreted as reflective of the East Asian media landscape specifically; future research should expand the corpus and use larger, more diverse labeled datasets to increase model robustness. Furthermore, the analysis acknowledges the inherent bias in media reporting, particularly from North Korea, where state-controlled narratives may not reflect public perception. This study interprets "cognitive differences" as variations in media discourse rather than national cognition. Therefore, the findings should be understood as insights into the framing of Chinese silk culture within specific media ecosystems, not as definitive measures of a country's collective understanding.

CONCLUSIONS

This paper analyzed the differences in Chinese silk culture across the media of four East Asian countries via a multilingual co-occurrence semantic network system. By proposing cross-language semantic alignment and dynamic network modeling techniques, this work achieved a quantitative evaluation and strategy optimization of the effectiveness of national image communication. Through domain fine-tuning of the BERT-multilingual model and a combined optimization of adversarial training and contrastive learning, this study constructed a high-precision semantic alignment space. The subsequently generated dynamic co-occurrence network, which was based on the PageRank algorithm, accurately captured the cognitive differentiation between Japan's focus on "traditional craftsmanship" and South Korea's focus on "modern innovation". The study further revealed the difference in communication effectiveness between Japan's "ecological dyeing and finishing" and South Korea's "smart loom" by using a TextCNN sentiment classifier and a cultural discount coefficient model, which can guide the generation of differentiated content strategies.

The research results revealed that the SDI of Japanese media reached 8.2, with a focus on "dyeing technology", whereas the emotional polarity value of Korean media toward modern innovation reached 0.79, which was higher than the value of 0.65 for Chinese media. The co-occurrence weight of North Korea's "trade data" reached 0.73. Moreover, dynamic evolution path analysis revealed that the annual growth rate of the co-occurrence frequency of South Korea's "silk-virtual fit" reached 18.9%, verifying the role of the technological narrative in reshaping the cognitive structure.

The innovation of this paper lies in applying cross-language semantic alignment technology to the study of textile culture communication, establishing a quantitative framework for dynamic cognitive graphs, and proposing a CDC-based communication compensation algorithm. Future work can expand this model to multimodal data and introduce self-supervised learning to improve generalizability, further promoting the transformation of the global communication paradigm of silk culture from "traditional symbols" to "modern values".

Author Contributions

Conceptualization – Wenxiang Zhang; methodology – Wenxiang Zhang and Sutian Xu; investigation – Fang Ouyang; writing-original draft preparation – Wenxiang Zhang, Fang Ouyang. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

Funding

This research received no external funding.

Acknowledgements

Not applicable.

REFERENCES

[1] Tao T, Park HW. The Structure of the Semantic Network Regarding “East Asian Cultural Capital” on Chinese Social Media Under the Framework of Cultural Development Policy. Information. 2025; 16(8):673. doi: 10.3390/info16080673

[2] Qi L, Wang Y, Chen J, Liao M, Zhang J. Culture under Complex Perspective: A Classification for Traditional Chinese Cultural Elements Based on NLP and Complex Networks. Complexity. 2021; 2021(1):6693753. doi: 10.1155/2021/6693753

[3] Khan MI, Hanzala, Huraira A. Silk Road Revival: Language Contact and Cross-Cultural Communication in the China-Pakistan Economic Corridor. Annals of Human and Social Sciences. 2024; 5(2):181-194. doi: 10.35484/ahss.2024(5-II)17

[4] Alba E, Gaitán M, León A, Mladenić D, Brank J. Weaving words for textile museums: The development of the linked SILKNOW thesaurus. Heritage Science. 2022; 10(1):59. doi: 10.1186/s40494-022-00681-x

[5] Zhao Y. The Application of Semantic Analysis Technology in the Analysis of Chinese and Korean Literature Mutual Translation. Mobile Information Systems. 2022; 2022(1):7249397. doi: 10.1155/2022/7249397

[6] Van Noort C, Chatterje-Doody PN. Visualizing China’s Belt and Road Initiative on RT (Russia Today): From infrastructural project to human development. Eurasian Geography and Economics. 2023; 64(4):431-459. doi: 10.1080/15387216.2021.2005650

[7] Loh DMH. The ‘Chinese Dream’ and the ‘Belt and Road Initiative’: Narratives, practices, and sub-state actors. International Relations of the Asia-Pacific. 2021; 21(2):167-199. doi: 10.1093/irap/lcz018

[8] Kawlra A, Sarkar S. Hybrid creature motifs as cross-cultural transmission along the Silk Roads. In: Zhao F, Marie-Louise N, editors. Textile and Clothing along the Silk Road. Paris, France: UNESCO Publishing; 2022. Volume 1, p. 209-228. doi: 10.54678/JDZP2879

[9] Ravichandran G, Nam SH. Interwoven Legacies: Exploring the Rich Textile Traditions of Korea and Tamil Nadu (India) through Cross-Cultural Influences. International Journal. 2024; 5(5):152-164. doi: 10.61707/rp1c5q26

[10] Huerta R. Silk Road Museums: Design of Inclusive Heritage and Cross-Cultural Education. Sustainability. 2021; 13(11):6020. doi: 10.3390/su13116020

[11] Zanier C. Italy, East Asia and Silk. One hundred years of a relationship (1830–1940). Journal of Modern Italian Studies. 2021; 26(2):173-185. doi: 10.1080/1354571X.2021.1896880

[12] Punyaratabandhu P, Swaspitchayaskun J. Thailand's Perception and Strategy toward China's BRI Expansion: Hedging with Cooperating. The Chinese Economy. 2021; 54(1):69-77. doi: 10.1080/10971475.20.1809819

[13] Gagliardi I, Artese MT. Exploring and Visualizing Multilingual Cultural Heritage Data Using Multi-Layer Semantic Graphs and Transformers. Electronics. 2024; 13(18):3741. doi: 10.3390/electronics13183741

[14] Gui A, Xiao H. Multi-level multilingual semantic alignment for zero-shot cross-lingual transfer learning. Neural Networks. 2024; 173:106217. doi: 10.1016/j.neunet.2024.106217

[15] Sebastián Lozano J, Alba Pagán E, Martínez Roig E, Gaitán Salvatella M, León Muñoz A, Sevilla Peris J, et al. Open Access to Data about Silk Heritage: A Case Study in Digital Information Sustainability. Sustainability. 2023; 15(19):14340. doi: 10.3390/su15191440

[16] Zeroual I, Lakhouaja A. MulTed: A multilingual aligned and tagged parallel corpus. Applied Computing and Informatics. 2022; 18(1-2):61-73. doi: 10.1016/j.aci.2018.12.003

[17] Cao X, Chen X, Huang L, Deng L, Cai Y, Ren H. Detecting technological recombination using semantic analysis and dynamic network analysis. Scientometrics. 2024; 129(11):7385-7416. doi: 10.1007/s11192-023-04812-4

[18] Chen H, Zhang Z, Huang S, Hu J, Ni W, Liu J. TextCNN-based ensemble learning model for Japanese Text Multi-classification. Computers and Electrical Engineering. 2023; 109:108751. doi: 10.1016/j.compeleceng.2023.108751

[19] Li Q, Li X, Du Y, Fan Y, Chen X. A New Sentiment-Enhanced Word Embedding Method for Sentiment Analysis. Applied Sciences. 2022; 12(20):10236. doi: 10.3390/app122010236

[20] Fernández-Martínez F, Luna-Jiménez C, Kleinlein R, Griol D, Callejas Z, Montero JM. Fine-Tuning BERT Models for Intent Recognition Using a Frequency Cut-Off Strategy for Domain-Specific Vocabulary Extension. Applied Sciences. 2022; 12(3):1610. doi: 10.3390/app12031610

[21] Chen LC, Chang KH, Wu CH, Chen SC. A machine-based corpus optimization method for extracting domain-oriented technical words: An example of COVID-19 corpus data. Journal of Intelligent & Fuzzy Systems. 2024; 46(4):9945-9956. doi: 10.3233/JIFS-236635