Received: 02 July 2026; **Revised:** 28 August 2026; **Accepted:** 16 October 2026; **Published:** 20 November 2026
Archs Sci. (2026) Volume 76, Issue 5 Pages 85-90, Paper ID 2026509.
https://doi.org/10.68304/as/76509

Using TF-IDF and Generative Patterns to Extract and Analyze Keywords in Ancient Korean Poetry

He Bu
Anshan Normal University, Anshan 114000, Liaoning, China
Correspondence: 13889779590@163.com

Abstract Ancient Korean poetry occupies a significant position in literary history because of its distinctive rhythmic organization, conventional poetic forms, and expressive treatment of human emotion. Nevertheless, automated interpretation and generation remain challenging because computational approaches may introduce thematic drift, fail to recognize metrical restrictions, or produce passages whose emotional tone is inconsistent with their intended subject. To address these issues, this study investigates keyword extraction and generative analysis through a combination of term frequency–inverse document frequency (TF-IDF) and deep-learning-based text generation. During theme planning, TF-IDF is supplemented by semantic expansion in a BERT embedding space to identify salient keywords and related expressions while maintaining topical relevance. A dual-constraint mechanism further combines a formal pattern dictionary with an emotion lexicon to guide the generation of verses that satisfy the specified structural requirements and exhibit a coherent emotional orientation. The generation framework employs sequence-to-sequence (Seq2Seq) learning, contextual representations produced by BERT, a bidirectional long short-term memory (BiLSTM) encoder, an attention mechanism, and beam search for candidate selection. The manuscript reports evaluations of thematic consistency, compliance with the selected metrical constraints, emotional expression, and the stylistic suitability of the resulting verses. The reported examples suggest that combining lexical, contextual, and structural information can support more coherent imitation of classical Korean poetry, although the correspondence between computational constraints and historically distinct Korean poetic traditions requires further examination.

Index Terms Ancient Korean poetry; keyword extraction; TF-IDF; BERT; sequence-to-sequence learning; metrical constraints; sentiment analysis; theme planning; text generation

I. Introduction

Rich historical, cultural, and emotional information is preserved in ancient Korean poetry, which remains an important form of literary and cultural expression. Its distinctive patterns of rhythm, imagery, and emotional communication also make it valuable for research on computational literary analysis, interpretation, and automated poem generation [1]. However, natural language processing systems continue to face substantial challenges when extracting representative keywords from classical Korean verse or producing new poems that remain consistent with specified thematic, metrical, and affective requirements. These problems concern not only the technical advancement of literary text analysis but also the digital preservation, interpretation, accessibility, and transmission of valuable cultural material. Effective computational approaches must therefore respond to the characteristics of the literary tradition rather than treating poems as ordinary, unrestricted sequences of words [2], [3].

Ancient Korean poetry comprises several historically distinct literary forms, many of which are concise, highly patterned, and rich in expressive significance. Among them, the traditional Korean sijo form, particularly prominent during the Joseon period, has attracted considerable scholarly attention because of its compact organization, thematic range, and capacity to express personal experience in culturally recognizable forms [4], [5]. Such works contribute to the documentation of social life, literary creativity, and inherited cultural values. At the same time, their characteristic rhythmic structures and conventions make automated interpretation and generation difficult. Korean-language sijo and Korean poetry composed in literary Chinese also require careful differentiation: the tonal rules commonly discussed in Chinese regulated verse cannot automatically be assumed to characterize every form of traditional Korean poetry [6].

Natural language processing has developed rapidly in recent years, especially through neural sequence-to-sequence architectures and pretrained language models such as BERT (Bidirectional Encoder Representations from Transformers). These methods provide useful foundations for computational analysis and automatic literary text generation. Nevertheless, applying them to complex poetic genres continues to present important theoretical and practical difficulties. When the gen-

erative process relies exclusively on isolated keywords, the resulting composition may drift away from its initial theme, making coherence difficult to maintain throughout successive lines. In addition, a model must balance the structural requirements selected for the target poetic form against semantic appropriateness, expressive variety, and emotional continuity. These competing objectives can become particularly difficult when formal conventions vary between poetic traditions or when examples are analyzed through translation. Conventional language-generation models may therefore produce syntactically fluent lines that are semantically incoherent, stylistically unsuitable, or insufficiently expressive [7], [8].

Several limitations remain despite previous attempts to combine keyword-driven generation with modern deep-learning techniques. Conventional keyword extraction approaches, including TF-IDF and TextRank, can identify frequent or prominent lexical items, but they do not always capture the implicit associations and literary imagery that connect different expressions within a poem [9]. Likewise, sentiment-analysis systems developed primarily for general prose may fail to recognize emotional ambiguity, figurative language, or sentiment conveyed indirectly through poetic images. Generative systems that apply formal constraints only at the surface level may still produce verses that do not satisfy the intended rhythm, structure, or culturally specific conventions. Furthermore, imbalanced training material, inconsistent annotation, restricted vocabularies, and inadequately tuned model parameters can contribute to unstable generation quality. Accordingly, effective evaluation must address literary appropriateness alongside computational measures rather than assuming that satisfying a lexical template guarantees artistic quality [10].

This study presents a combined approach to keyword extraction and generative analysis of ancient Korean poetry, integrating TF-IDF, semantic representation learning, and controlled neural generation. The principal aim is to improve the quality and relevance of the thematic information supplied to the generative model while retaining the structural and emotional criteria defined for each generation task. More specifically, TF-IDF identifies prominent lexical elements, and BERT-based vector-space expansion is used to enrich the corresponding keyword sets with semantically related expressions. The resulting topical representations provide a more informative and context-sensitive input for subsequent generation than isolated keyword matching alone. These procedures are intended to support higher-quality topic selection without assuming that automated lexical analysis can replace expert literary interpretation.

The generation procedure considers both formal organization and emotional continuity through a dual-constraint mechanism based on a pattern dictionary and a lexicon of affective terms. Candidate expressions can therefore be screened against the structural specification used in the experiment while maintaining an intended emotional orientation. A complementary theme-planning strategy uses TextRank to identify prominent terms and controls the development of successive lines, thereby reducing unintended shifts between closely connected topics. Contextual BERT representations are combined with a BiLSTM encoder to model meaning across the input sequence, while the attention mechanism highlights relevant contextual information during decoding. Finally, beam search retains and compares plausible candidate sequences rather than selecting words exclusively through a single greedy decision. Taken together, these components establish a coordinated basis for the generation and examination of stylistically constrained poetic text.

The research approach combines poetry analysis, lexical selection, and constrained generation in a manner that may also inform computational studies of other languages and literary genres. Beyond its immediate application to classical Korean poetry, it illustrates how natural language processing techniques can contribute to the digital documentation, exploration, and presentation of cultural heritage. The combination of computational modeling with literary scholarship is especially important because interpretability, historical accuracy, and cultural appropriateness cannot be determined from algorithmic scores alone. By describing the extraction, planning, and generation stages within a common framework, this study provides a basis for subsequent research on controlled text generation, human evaluation of poetry, and the relationship between formal literary constraints and semantic creativity.

II. Poetry Generation Modeling

A. Sentiment Analysis and Theme Planning

The selection of subject matter and the control of emotional expression are central to the composition of coherent poetry and therefore constitute important challenges for automated generation. The present study introduces a theme-planning procedure that combines keywords extracted by TextRank with semantically related terms identified through the BERT embedding space. These expressions are used to guide the content of individual lines and preserve a recognizable thematic trajectory across the poem [11]. Such planning addresses the tendency of a single input keyword to dominate the entire generation process without establishing meaningful connections between consecutive verses. In parallel, the intended emotional orientation is estimated using a BERT-based sentiment classifier, and an emotion dictionary constrains candidate expressions during the decoding process. These complementary mechanisms are designed to make topic progression and affective consistency explicit objectives rather than incidental consequences of language-model prediction.

The BERT-based sentiment classifier is trained or fine-tuned on an annotated sentiment corpus to distinguish among the emotional categories specified for the task [12]. In the preprocessing stage, poetic excerpts and their corresponding labels are converted into structured input records containing token identifiers (`token_ids`), category labels, sequence lengths (`seq_len`), and attention masks that distinguish valid tokens from padding positions. These records are subsequently organized into batches for iterative training. A cross-entropy objective measures disagreement between predicted

sentiment probabilities and the annotated labels, and an adaptive moment estimation (Adam) optimizer updates the trainable parameters. The classifier is assessed using its output probabilities, with the predicted category selected according to the highest class score. This workflow follows established approaches to BERT fine-tuning for supervised classification [13]. Because poetic sentiment may be indirect or ambiguous, category assignments should be interpreted in relation to the quality of the training labels and the representativeness of the selected corpus. The classifier provides an operational measure of emotional orientation rather than a definitive interpretation of literary meaning.

B. Emotional and Metrical Constraints

Traditional Korean poetry exhibits distinctive structural and rhythmic features, although the relevant conventions depend on the period, language of composition, and particular genre. One important historical example is Choe Chiwon, also romanized as Ch'oe Ch'i-won, a scholar and poet of Unified Silla who studied in Tang China and passed an imperial examination in 874. His literary reputation was established partly through poetry composed in literary Chinese, demonstrating the close historical interaction between Korean literary culture and classical Chinese verse. The poem commonly known in English as "In the Rain on an Autumn Night" conveys loneliness, distance, and emotional isolation through images of lamplight, rainfall, limited human companionship, and the autumn wind. These images illustrate the importance of literary context for understanding affective expression in historically transmitted verse [14]. Such examples also caution against applying one uniform metrical system to every ancient Korean genre.

The Korean sijo tradition developed in a different linguistic and formal setting from Korean hansi, or poetry written in literary Chinese. Sijo was composed and transmitted in Korean and is characterized by recognizable rhythmic and stanzaic conventions, whereas Chinese-script verse was closely associated with classical literary education and official culture. The manuscript's generation procedure operationalizes poetic structure through an explicitly selected formal template and combines that template with an emotion constraint. Under this procedure, each proposed line is divided into rhythmic units, and the model ranks candidate expressions according to their predicted probabilities. Candidates that fail to satisfy the configured formal categories or the required emotional orientation are removed before the remaining expressions are used to complete a verse. Figure 1 presents examples of the dictionaries used for this filtering process. In the experimental coding scheme, category 1 denotes the level-tone label and category 2 the oblique-tone label, while the corresponding emotional categories distinguish positive and negative expressions. These tonal labels describe the selected computational constraint scheme; they should not be interpreted as a comprehensive account of traditional Korean-language sijo prosody.

C. Poetry Generation Model

The poetry-generation component uses a sequence-to-sequence (Seq2Seq) architecture guided by the lexical and emotional constraints described above. The dictionary examples in Figure 1 indicate how candidate terms can be assigned to formal and affective categories before decoding. In the representation stage, a pretrained BERT language model supplies context-dependent embeddings that capture semantic distinctions among poetic terms and phrases. These representations may be further adapted to the literary vocabulary present in the training corpus and subsequently passed to the encoder [12]. The generative architecture combines these contextual representations with a bidirectional long short-term memory network and attention-guided decoding to relate each newly produced unit to relevant portions of the preceding text.

Korean Vocabulary for Emotions/Feelings		
Love	사랑	salang
Worried	걱정	geogjeong
Angry	성난	seongnan
Sad	슬픈	seulpeun
Embarassed	당황	danghwang
Afraid	겁	geob
Kind	종류	jonglyu
Happy	행복한	haengboghan
Impolite	무례한	mulyehan
Painful	민감한	mingamhan
Bored	지루한	jiluhan
Tired	피곤한	pigonhan
Sensitive	아픈	apeun
Beautiful	아름다운	aleumdaun
Delighted	물론,	mullon,
Amazed	놀란	nollan
Hate	미움	mium

Figure 1: Examples of Formal Pattern and Emotion Dictionaries

To represent contextual information in the input poetry sequence, the encoder employs a bidirectional long short-term memory (BiLSTM) network that processes textual context in both directions. The attention mechanism compares encoder hidden states, denoted by H_i, with decoder states, denoted by S_i, to estimate the relevance of different source positions. The resulting attention weights, represented in the manuscript by A_W, are used to aggregate encoder information into a context representation C_i. This representation is combined with the decoder state to predict a sequence of output tokens $(y_0, y_1, \ldots, y_t)$. Rather than relying exclusively on the highest-probability token at each step, beam search maintains multiple candidate continuations and selects complete sequences according to the chosen decoding criterion. This approach can improve the coherence of candidate poems while retaining the flexibility needed to satisfy structural and emo-

ARCHIVES DES SCIENCES
A Multidisciplinary Journal

tional restrictions. The mathematical expressions employed in the original formulation are retained below:

$$A_W = \frac{\exp\left(\text{score}\left(H_i, S_i\right)\right)}{\sum_{i=1}^{t} \exp\left(\text{score}\left(H_i, S_i\right)\right)}, \tag{1}$$

$$C_i = \sum_{i=1}^{t} A_W \cdot H_i, \tag{2}$$

$$S_i = g\left(y_{i-1}, S_{i-1}, c_i\right), \tag{3}$$

$$y_i = W_0 S_i + b_0. \tag{4}$$

In these expressions, $g(\cdot)$ denotes the nonlinear decoder update operation, y_i refers to the predicted output at position i, C_i represents the attention-derived contextual information available to the decoder, and H_i represents a hidden state of the encoder. The notation reflects the model's reported computational procedure. In an implementation, the indexing of attention weights, context vectors, and decoder states must be consistent across all time steps; the mathematical definitions should therefore be checked against the source code before the method is reproduced or independently benchmarked.

III. Techniques for Information Extraction and Concealment

A. Information Concealment Algorithm

For the information-concealment component, the secret message is first represented as a bitstream B using a specified character encoding, such as Unicode with an appropriate byte representation, and any intended encryption is applied as a separate operation. The trained poetry-generation model then receives thematic keywords and user-specified structural constraints to generate a candidate poem Y. At each decoding position, the model produces a probability distribution over candidate words or characters. A restricted candidate set is constructed by retaining high-probability expressions that satisfy the required metrical and emotional criteria. The selected candidate set is subsequently assigned variable-length bit codes through a Huffman coding tree, allowing portions of the secret bitstream to determine which candidate is chosen. When the remaining secret bits cannot be accommodated in a single poem, the generation process may be repeated with the same theme and constraints until the message has been embedded [15]. This description distinguishes text encoding, encryption, probability-based candidate selection, and steganographic embedding, which perform different technical functions and should not be treated as interchangeable operations.

B. Secret Information Extraction

The extraction procedure reverses the principal operations of the information-embedding process [16]. First, the receiver identifies the form and relevant structural characteristics of the poem containing the concealed message. The sender and receiver must share the same generation model, theme specification, associated keyword lists, dictionaries, decoding rules, and coding convention. Using this common configuration, the receiver reconstructs the ordered candidate set available at each encoded position and determines the corresponding Huffman code from the observed word or character. The recovered codes are concatenated in their original sequence to reconstruct separate bitstream portions, denoted B_1 and B_2 when the message spans multiple segments. These portions are then combined to obtain the complete bitstream B. Finally, the receiver reverses any encryption and character-encoding operations applied during embedding to recover the original message. Successful reconstruction requires consistent candidate ordering and deterministic handling of vocabulary constraints at both ends of the communication process; otherwise, the recovered bitstream may not match the embedded information.

IV. Results

The experimental examples use a set of thematic expressions drawn from classical Korean poetry, including "flower," "wind," "autumn," "sky," and "spring." The algorithm identifies candidate expressions related to the supplied keywords and retrieves additional lexical items from the associated database. The original experimental description specifies an "iambic pentameter" configuration; because this designation derives from English verse, its exact relation to the formal criteria used for Korean poetry requires clarification in any reproducible implementation. Table 1 gives representative candidate words and their reported quantities for four of the displayed keyword categories. The remaining evaluation tables, Tables 2, 3, and 4, record the reported adaptation or fitness scores for several combinations of evolutionary iterations, initial population sizes, and the parameter P. The scores are preserved as supplied in the original manuscript, without assuming additional evaluation procedures that have not been documented.

Table 1: Representative Candidate Words Identified for Selected Keywords

Keyword	Number of Candidates	Examples of Candidate Words
Flower	320	Rui, Mei, Cai, Luo, Ye
Wind	375	Clear, waves, snow, rain, flute
Day	150	White, night, light, heron, green
Spring	270	Summer, warmth, dew, emerald, fragrance

Table 2: Reported Poetry Adaptation Scores for the First Parameter Group

Iterations	Initial Population	P	Flower	Wind	Day	Spring
100	20	0.1	3	2	2	4
150	20	0.1	5	4	5	3
200	30	0.3	4	4	4	4
300	40	0.5	4	3	3	4

The numerical results in Tables 2–4 provide a basis for examining how the reported poetry-adaptation scores vary with evolutionary iterations, initial population size, and parameter P, which is described as the variation or mutation rate. In several configurations, higher scores coincide with

ARCHIVES DES SCIENCES
A Multidisciplinary Journal

larger populations or more iterations; however, the rows also change more than one parameter at a time. Consequently, these tables do not isolate the independent effect of any single parameter, and a universal monotonic relationship cannot be established from the available values alone. For example, Table 2 shows that the Flower score rises from 3 to 5 as the iterations increase from 100 to 150 while the population and P remain fixed at 20 and 0.1, respectively. By contrast, the same table includes subsequent changes in both population and P, preventing a controlled comparison with that earlier pair. Table 3 reports Flower scores ranging from 3 to 4 across later iterations, whereas Table 4 contains Wind scores ranging from 2 to 5 under different parameter combinations. Such variation suggests that different thematic keywords may respond differently to the selected evolutionary settings. The observed outcomes therefore support a descriptive comparison of candidate configurations rather than definitive causal conclusions about mutation rate or population size. In addition to these numeric measures, poems generated from keywords such as "flower," "wind," "sky," and "spring" were subjected to manual screening for readability and stylistic suitability. The general idea of computational poetry-style analysis is consistent with earlier research [1], while manual literary evaluation remains necessary because a high adaptation score does not by itself establish cultural authenticity. The generated examples were further considered in relation to studies of poetry generation and evaluation [10], [17], [18]. Table 5 records a specific generation configuration used for an illustrative output. Its reported fitness value is retained exactly, although the original manuscript does not fully specify the associated scoring rubric or its uncertainty.

Table 3: Reported Poetry Adaptation Scores for the Second Parameter Group

Iterations	Initial Population	P	Flower	Wind	Day	Spring
400	20	0.1	4	5	3	4
500	20	0.2	4	4	4	3
800	30	0.3	3	4	5	4
1000	40	0.5	4	4	3	4

Table 4: Reported Poetry Adaptation Scores for the Third Parameter Group

Iterations	Initial Population	P	Flower	Wind	Day	Spring
100	30	0.1	3	4	2	4
200	40	0.1	4	5	4	2
300	50	0.4	4	2	3	4
400	60	0.4	4	5	3	2

An illustrative English rendering of a generated classical-style poem, produced under the parameter configuration reported in Table 5, is provided below. This rendering conveys the imagery and thematic sequence of the example; it should not, by itself, be treated as evidence of compliance with the prosodic conventions of a particular historical Korean verse form.

> Spring Breeze and a Shower of Blossomss the spring breeze blows, the willows dance,nd a shower of blossoms scatters, filling the bridge with fragrance.he cry of a blue bird drifts from a distant tree,hile a stream flows onward, vanishing into the clouds. [Translated from Korean]

The example was subsequently subjected to manual rhetorical revision informed by the study's intended metrical specifications, lexical relationships, sequencing of poetic images, and sentiment-oriented analysis. These editorial adjustments were intended to improve readability and emotional continuity while retaining the theme of the initially generated text. Because the examples are displayed in English translation, the relation between their visible line structure and the original Korean-language formal constraints cannot be independently assessed from the rendered versions alone.

> A Quiet Poem for an Autumn Nightoonlight illuminates the window and brushes against the old wall;herever the autumn breeze stirs, yellow leaves scatter.ho stands by the river, holding a cup of wine?nly the cold river water keeps company through the long night. [Translated from Korean]

Table 5: Selected Parameters for an Illustrative Poetry Generation

Iterations	Initial Population	P	Keyword	Fitness (F)
1000	40	0.2	Flower	5

V. Conclusion

This study examines an integrated approach to keyword extraction and controlled poetic text generation that combines TF-IDF, semantic expansion through BERT, and a neural sequence-to-sequence architecture. The proposed procedure addresses several challenges in the computational analysis and imitation of classical Korean poetry, including selecting thematic terms, maintaining emotional continuity, and applying explicit formal constraints during decoding. In the keyword extraction stage, TF-IDF provides a lexical relevance signal, while contextual embeddings expand the representation of associated poetic concepts. The generation stage introduces a dual-constraint mechanism that screens candidate expressions according to selected formal categories and an emotion lexicon. BERT-based input representations, a BiLSTM encoder, attention-guided decoding, and beam search are then used to maintain contextual relevance and identify plausible candidate lines. The reported results illustrate changes in adaptation scores under different evolutionary parameter configurations and provide examples of literary output following manual screening. Although several reported configurations yield comparatively high scores, the supplied tables do not establish a controlled comparison with all preceding methods, and the extent to which the generated verses conform to historically specific Korean poetic forms requires independent specialist evaluation. The approach nevertheless offers a useful basis for further investigation of controlled literary generation and computationally supported cultural analysis.

The findings contribute to the continuing discussion of how computational language technologies can support the study, preservation, and presentation of traditional literary materials. Combining probabilistic text generation with explicit thematic and affective criteria may help researchers investigate the interaction between lexical choice, poetic imagery, and formal organization. At the same time, several limitations must be considered when interpreting the manuscript. The relationship between Korean-language sijo, Chinese-script poetry produced in Korea, and imported tonal or metrical labels is not fully resolved by the current experiments. Furthermore, the composition of the training corpus, the exact definition of the fitness function, the annotation protocol, and repeated-run variability are not described in sufficient detail for complete independent reproduction. Future research should therefore include larger and clearly documented corpora, genre-specific prosodic annotation, controlled parameter studies, evaluation by qualified literary specialists, and stronger comparisons with established text-generation systems. Investigating additional language models and evolutionary search methods could also improve diversity, coherence, and reproducibility across a broader range of poetic forms and cultural contexts.

Conflict of Interest

The author declares that there are no known financial or non-financial conflicts of interest.

Funding

This work was undertaken as part of the 2026 General Project concerning the construction and applied research of digital textbooks and artificial intelligence (AI) general education textbooks. The project is entitled “Practical Research on the Integration of AI General Education, Digital Textbooks and Teaching Resource Construction in the Tourism Professional Scenario.” According to the information provided by the author, the project was approved by the Information Technology Steering Committee for Vocational College Education, Ministry of Education of the People’s Republic of China, under Project No. KT2608162, with official document reference XXJZWTW [2026] No. 38.

Data Availability

The datasets used and analyzed during the present study are available from the author upon reasonable request.

References

[1] Deng, S., Wang, G., Wang, H., & Chang, F. (2023). An artificial-intelligence-driven Spanish poetry classification framework. *Big Data and Cognitive Computing, 7*(4), Article 183. https://doi.org/10.3390/bdcc7040183

[2] Lee, P. H. (Ed.). (2003). *A history of Korean literature*. Cambridge University Press. https://doi.org/10.1017/CBO9780511485954

[3] Lee, P. H. (Ed.). (2003). *The Columbia anthology of traditional Korean poetry*. Columbia University Press. https://cup.columbia.edu/book/the-columbia-anthology-of-traditional-korean-poetry/9780231111133/

[4] Rutt, R. (Ed. & Trans.). (1998). *The bamboo grove: An introduction to sijo*. University of Michigan Press. https://doi.org/10.3998/mpub.8299

[5] McCann, D. R. (2000). *Early Korean literature: Selections and introductions*. Columbia University Press. https://doi.org/10.7312/mcca11946

[6] Zhang, X., & Lapata, M. (2014). Chinese poetry generation with recurrent neural networks. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing* (pp. 670–680). Association for Computational Linguistics. https://doi.org/10.3115/v1/D14-1074

[7] Yi, X., Li, R., & Sun, M. (2016). *Generating Chinese classical poems with RNN encoder-decoder* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.1604.01537

[8] Yi, X., Sun, M., Li, R., & Li, W. (2018). Automatic poetry generation with mutual reinforcement learning. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing* (pp. 3143–3153). Association for Computational Linguistics. https://doi.org/10.18653/v1/D18-1353

[9] Mihalcea, R., & Tarau, P. (2004). TextRank: Bringing order into text. In *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing* (pp. 404–411). Association for Computational Linguistics. https://aclanthology.org/W04-3252/

[10] Gonçalo Oliveira, H. (2017). A survey on intelligent poetry generation: Languages, features, techniques, reutilisation and evaluation. In *Proceedings of the 10th International Conference on Natural Language Generation* (pp. 11–20). Association for Computational Linguistics. https://doi.org/10.18653/v1/W17-3502

[11] Wang, Z., He, W., Wu, H., Wu, H., Li, W., Wang, H., & Chen, E. (2016). Chinese poetry generation with planning based neural network. In *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers* (pp. 1051–1060). The COLING 2016 Organizing Committee. https://aclanthology.org/C16-1100/

[12] Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1* (pp. 4171–4186). Association for Computational Linguistics. https://doi.org/10.18653/v1/N19-1423

[13] Sun, C., Qiu, X., Xu, Y., & Huang, X. (2019). How to fine-tune BERT for text classification? In M. Sun, X. Huang, H. Ji, Z. Liu, & Y. Liu (Eds.), *Chinese computational linguistics* (pp. 194–206). Springer. https://doi.org/10.1007/978-3-030-32381-3_16

[14] Literature Translation Institute of Korea. (n.d.). *Ch’oe Ch’i-won*. Digital Library of Korean Literature. https://library.ltikorea.or.kr/writer/200682

[15] Yang, Z.-L., Guo, X.-Q., Chen, Z.-M., Huang, Y.-F., & Zhang, Y.-J. (2019). RNN-Stega: Linguistic steganography based on recurrent neural networks. *IEEE Transactions on Information Forensics and Security, 14*(5), 1280–1295. https://doi.org/10.1109/TIFS.2018.2871746

[16] Ziegler, Z., Deng, Y., & Rush, A. M. (2019). Neural linguistic steganography. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing* (pp. 1210–1215). Association for Computational Linguistics. https://doi.org/10.18653/v1/D19-1115

[17] Chen, Y., Gröner, H., Zarrieß, S., & Eger, S. (2024). Evaluating diversity in automatic poetry generation. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing* (pp. 19671–19692). Association for Computational Linguistics. https://doi.org/10.18653/v1/2024.emnlp-main.1097

[18] Walsh, M., Preus, A., & Antoniak, M. (2024). Sonnet or not, bot? Poetry evaluation for large models and datasets. In *Findings of the Association for Computational Linguistics: EMNLP 2024* (pp. 15568–15603). Association for Computational Linguistics. https://doi.org/10.18653/v1/2024.findings-emnlp.914