

Extending Creativity: Large Language Models and the Practice of Poetry Translation

Natalia Resende^{1,2,3} and James Hadley^{1,2}

¹Trinity College Dublin

²Trinity Centre for Literary and Cultural Translation

³ADAPT Centre

resenden@tcd.ie, hadleyj@tcd.ie

Abstract

The aim of this paper is to propose a framework to support poetry translators in making effective use of large language models (LLMs) through established prompt engineering strategies applied to both pre-translation and translation stages. The paper illustrates these strategies using poems characterised by multiple layers of syntactic, semantic, phonological, and cultural complexity, and discusses how LLMs perform in response to each prompting technique. It also engages with the longstanding claim that poetry translation is a purely human endeavour and cannot be computer-assisted, arguing instead that LLMs, rather than replacing human creativity, have the potential to extend it.

1 Introduction

For most of human history, the translation of poetry as poetry has been understood as an exclusively human endeavour, requiring creative intuition, cultural knowledge, and sensitivity to the full range of linguistic, stylistic, and cultural features a poem may involve (Lefevere, 1975). As noted by (Bassnett, 1998), much of the discourse surrounding poetry translation has historically leaned toward the mystical, treating the translator's capacity to navigate these features as a gift rather than a descriptive process. Yet the challenges involved are identifiable and, to a significant extent, systematic.

At the linguistic level, poets frequently exploit ambiguity, register, and unusual syntax in ways that stretch or deliberately violate the norms of the source language (Lefevere, 1975). At the stylistic level, many poetic forms impose formal constraints such as metrical patterns, rhyme schemes,

syllable counts, and sound patterning that restrict lexical choice while simultaneously defining the poem's aesthetic identity. At the cultural level, poems may encode historically specific references, symbolically loaded images, or culturally embedded associations that a source language reader is expected to recognise but that a translator must negotiate carefully with respect to the target readership. These three dimensions are rarely independent: a single word choice may be simultaneously constrained by metre, freighted with cultural association, and semantically ambiguous (Bassnett, 1998).

The Italian Petrarchan sonnet, for instance, consists of hendecasyllabic lines of eleven syllables, while the English Shakespearean sonnet is written in iambic pentameter, where each line contains five feet, each foot being an iamb, i.e., one unstressed syllable followed by one stressed syllable (Stageberg, 1948). Some poetic forms impose additional phonic constraints, formalising the placement of features such as alliteration (the repetition of the same initial consonant sound in closely connected words), assonance (the repetition of similar vowel sounds in nearby words), or rhyme at specific points in the poem. The Old Norse *dróttkvætt*, for example, employs a complex alliterative pattern, a feature it shares with the Welsh *cynghanedd*, which requires that the consonants surrounding the main stressed vowel before the caesura (a pause or break within a line of poetry) be repeated in the same order after it (Clunies Ross, 2010).

Poetry can also be highly culturally specific, evoking historical moments, places, people, activities, or foods for the emotional associations they carry for the source readership, and which translators must negotiate with reference to the target audience (Newman, 1958).

A further layer of complexity arises with poems that appear to address universal experiences but are in fact grounded in culturally specific notions that the poet includes on the assumption that the source reader will recognise them, and which the translator may not be able to render with equal confidence. For instance, poem number 33 from the medieval Japanese tanka collection *Ogura Hyakunin Isshu*, by Ki no Tomonori (died 905) asks:

久方の光のとけき春の日にしづ心なく花の
散るらむ

[In the everlasting | tranquil light | of a spring day
| relentlessly | the flowers fall - I wonder why]

The final unit, where the image of the flowers falling (花の散る) is introduced, may appear straightforward at first reading. However, it addresses one of the most culturally loaded images in Japanese poetry (Hume, 1995).

The flowers in question, though not explicitly identified, are almost certainly cherry blossoms. Indeed, in this form of Japanese poetry, it is understood that when the generic word 花 [flower] is invoked, cherry blossoms specifically are implied. This inference is strengthened by the verb chosen, 散る [to fall], which is specific to the scattering of blossoms and leaves, though it can be used metaphorically to indicate the death of people. Together, the two words invoke the image of the scattering cherry blossoms, used widely in Japanese aesthetics to symbolise transience in general, and the ephemeral nature of beauty.

Given all these constraints, translating poetry represents a creative exercise in which the translator must choose between prioritising fidelity to the source poem’s meanings or adopting a more creative stance aimed at recreating its effects in the target language.

It is for these reasons that machine translation systems (MT) have not traditionally been considered suitable tools for poetry translation (Bellos, 2011). Conventional MT systems select output based on the statistical probability of mappings between source and target language (Forcada, 2017), an approach structurally at odds with the nature of poetry, which actively exploits linguistic improbability, favouring the original, the novel, and the unexpected over the statistically likely (Fabb, 2015). While an MT system might produce a rendering that captures the general sense of a poem, it cannot convey the inferences arising from ambiguity, formal constraint, and culturally specific content

(Bellos, 2011).

With the emergence of large language models (LLMs), however, this picture may be changing. Unlike conventional MT systems, LLMs can be directed through targeted instructions to engage with the full range of natural language processing, including translation problems (Karpinska and Iyyer, 2023). We argue that, when prompted effectively, i.e., using appropriate prompt engineering techniques, LLMs can serve as productive assistants at both the **pre-translation** and **translation stages**, offering resources that go beyond what traditional tools such as dictionaries, glossaries, or MT systems can provide. This paper presents a set of prompt engineering strategies designed to address constraints at the linguistic, stylistic, and cultural levels, with the aim of providing translators with a practical framework for integrating LLMs into their workflow. The examples presented throughout this paper were generated using two different models, ChatGPT 5.2 and Claude Sonnet 4.6, reflecting the model-agnostic nature of the framework proposed: the prompting strategies discussed here can be applied across different LLMs, and the variation in model choice also allows for incidental observations about differences in model behaviour across tasks.

In the following sections we present the strategies suitable for each of the pre-translation and translation stages.

2 Prompt Engineering for Poetry Translation

Unlike traditional machine translation models, where the translated output is fixed and can only be modified through human intervention in a post-editing process, LLMs offer greater flexibility. When it comes to translating poetry, LLMs allow users to prompt the system with specific strategic instructions. In the case of poetry translation, LLMs can be used to focus on or to reproduce particular features of the source poem, providing examples of output, working iteratively, and requesting multiple translation options for any given input. For example, a translator can prompt an LLM to translate a poem while preserving its poetic form or, alternatively, to render it as prose. Another possibility is to instruct the model to transform the poem into another form of text while incorporating specific linguistic devices. The widespread availability of LLMs since late 2022 thus makes

it meaningful, for the first time in history, to ask substantive questions about the role that technology may play in the translation of poetry.

In the specific case of LLMs and poetry translation, however, the discussion must be refined to reflect both the range of challenges involved in the pre-translation and translation stages. For each stage, prompt engineering strategies already established in the literature can be adopted by the translator to address its specific demands, as summarised in Table 1. The following section presents these strategies in turn, illustrating how each can be used to support the translator’s decision-making and help navigate the linguistic, stylistic, and cultural challenges identified above.

2.1 Pre-translation stage

At the pre-translation stage, the translator has identified the poem they would like to translate and must now develop a thorough understanding of its features and a strategy for rendering them in the target language. This phase involves identifying the formal features pertinent to the poetic form in general and the poem in particular, anticipating elements that require elaborate interpretive or creative decision-making, and formulating a set of priorities for negotiating constraints in the target language. The challenge becomes even more complex when culturally embedded meanings are at stake. As discussed in the introduction, the image of 花 in classical Japanese poetry, reinforced by the verb 散る (to fall/scatter), which may metaphorically evoke transience and mortality, exemplifies this kind of constraint. Traditionally, its resolution depends on the translator’s cultural knowledge and interpretive reasoning. Several prompt engineering strategies can support and structure this reasoning process; we present here well-studied approaches with demonstrated performance improvements compared to other methods in similar contexts. Each of the strategies discussed below externalises a different dimension of the translator’s preparatory work: cultural inference, systematic task decomposition, and goal specification.

2.1.1 Chain-of-Thought Prompting (CoT)

CoT prompting consists of a sequence of intermediate natural-language reasoning steps that guide the model toward a final output. This strategy has demonstrated evidence of improved model performance on a range of arithmetic, common-

sense, and symbolic reasoning tasks (Wei et al., 2022). Applied to poetry translation, this strategy allows the translator to guide the model through a staged process: first eliciting an explanation of the cultural significance of a given image, then generating possible symbolic interpretations, and finally producing a translation that preserves the source poem’s ambiguity rather than resolving it. In this way, culturally embedded constraints become operationalisable through prompt-mediated reasoning, transforming the process into an explicit procedural component of the translation workflow. A concrete example of this staged prompting applied to the Japanese tanka using ChatGPT 5.2 (thinking mode) discussed in the introduction is provided below.

Step 1: Cultural Elicitation

In classical Japanese poetry, what symbolic meanings are associated with 花 in the context of spring?

Step 2: Interpretive Expansion

Explain how the verb 散る (to fall/scatter) interacts with this symbolism. Does it evoke themes beyond a literal description?

Step 3: Constraint Specification

List possible interpretations that preserve both seasonal imagery and connotations of transience or mortality.

Step 4: Translation under Constraint

Now produce an English translation that preserves the symbolic ambiguity rather than resolving it explicitly.

This staged prompting guides the model to engage with the poem’s cultural and symbolic layers before producing a translation. Applied to the Japanese tanka, the model returned the following translation:

*In the gentle light
of a calm spring day,
why do the flowers fall
with no peaceful heart?*

At each intermediate step, the model produced analytical output that goes beyond surface-level glossing. In response to Step 1, it identified the symbolic associations of 花 across five distinct dimensions, correctly foregrounding culturally compressed meanings, such as *mujō* and *mono no aware*, that are precisely the layers most resistant to direct translation. In response to Step 2, it identified how 散る does not merely extend the symbolism of 花 but dramatises it, shifting the poem from a state of beauty to the enactment of its loss which is a distinction with direct consequences for word choice in any target-language rendering. Step 3 then produced a structured set of interpretive framings operating simultaneously at the literal and symbolic levels, providing an explicit scaffold that a translator could use to evaluate which connotations to foreground and which to sacrifice. What is analytically significant is not merely that the model generated relevant cultural information, but that the chain-of-thought structure caused it to make its interpretive reasoning explicit and ordered before any target-language commitments were made, a staged process with clear potential to facilitate translator decision-making. The full model output is provided in Appendix 3.

The CoT strategy is well suited to poems with dense cultural or symbolic content, but it can be extended to address other constraint types identified in the introduction. A translator using the CoT sequence can move through formal constraints (metre, rhyme, lineation, sound patterning), linguistic constraints (ambiguity, polysemy, syntax, register, wordplay), and cultural constraints (allusions, historical references, symbolic associations), before prompting a translation that accounts for all identified layers simultaneously.

2.1.2 Scaffolded Prompting (SP)

SP(Suzgun and Kalai, 2024) guides the model to deconstruct a complex task into smaller, ordered subtasks completed in sequence. While the CoT strategy externalises cultural and interpretive reasoning, scaffolded prompting externalises the analytical task structure itself, ensuring the model addresses each dimension of the poem systematically before any translation decisions are made. This makes it well suited to addressing formal and linguistic constraints. A scaffolding sequence targeting these constraint types using ChatGPT 5.2 is illustrated below.

I want to translate François Villon’s *Ballade des dames du temps jadis* into English. Before translating it, complete the following stages in order:

1. Identify sound, rhythm, lineation, and syntactic compression.
2. Identify lexical ambiguities and symbolic images.
3. Note cultural references and allusions.
4. List the main translation problems likely to arise.

Applied to Villon’s *Ballade des dames du temps jadis*, this prompt produced a detailed formal and linguistic analysis. The model identified how the poem’s nasal vowels and liquid consonants contribute to its sound texture, mapped its rhythmic structure and octosyllabic lineation, and noted the syntactic compression by which entire stories are condensed into single lines. It also returned a comprehensive inventory of lexical ambiguities and symbolic images, a catalogue of cultural references and allusions, and a prioritised list of translation problems likely to arise. Such output provides insights at all levels that can be useful to the translator as, equipped with this scaffolded analysis, the translator is in a position to make informed strategic decisions before engaging with the target language. For instance, whether to domesticate or foreignise culturally specific references, and how to handle lexical ambiguities where meaning cannot be fully preserved. It is worth noting that the model itself flagged these as open questions rather than resolving them, suggesting an appropriate degree of deference to translator agency. The full output is provided in Appendix 3.

2.1.3 Goal-Oriented Prompting (GOP)

GOP (Li et al., 2024) anchors the pre-translation analysis in a specified communicative purpose, or *skopos*, before any analytical or translational work begins. Since translation decisions are shaped by the intended function and audience of the target text (Suojanen et al., 2014), specifying this context at the outset allows the model to prioritise the constraints specified in the prompt. The same poem can require different analytical emphases depending on whether the target text is intended for academic study, publication in a poetry journal, classroom use, live performance, or a bilingual critical edition, among other possibilities. An example il-

lustrating a skopos specification is provided below.

This poem will be translated for publication in a bilingual academic edition. In the pre-translation stage, identify the features that must be preserved for scholarly value and the features that may be adapted without undermining interpretation.

This prompt, submitted to Claude Sonnet 4.6 model returned a structured pre-translation feature map categorising the poem’s elements into those that must be preserved and those that may be adapted, accompanied by prose justifications for each decision. The analytical quality of the output is high: on the preservation side, the model correctly identified the refrain’s argumentative rather than merely decorative function, flagged the untranslatability of *antan* as an interpretive crux requiring annotation rather than normalisation, and treated the ordering of exempla as rhetorical argument rather than decorative listing. On the adaptation side, it acknowledged the well-documented impossibility of reproducing the *ababbcbc* rhyme scheme without lexical distortion, invoking Swinburne’s version as a cautionary case and distinguished between archaic morphology as a formal feature and temporal distance as the effect to be reproduced, a distinction with direct practical consequences for translator strategy. The goal-oriented framing thus prompted the model to reason about translatability at a level of granularity that goes beyond general cultural commentary, producing output that is directly actionable for translation. The complete output is provided in Appendix 3.

Taken together, these three strategies equip the translator with a structured pre-translation map of the poem’s constraints across linguistic, stylistic, and cultural dimensions. These tools externalise reasoning processes that have traditionally been the translator’s implicit and internal domain. They transform the preparatory phase of poetry translation into an explicit, iterative dialogue between the translator and the model. The output of this stage feeds directly into the translation process itself, which we address in the following section.

2.2 Translation

While the pre-translation strategies discussed above help the translator develop a comprehensive understanding of the source poem’s constraints be-

Strategy	Stage	Function / Constraint
CoT	Pre-translation	Cultural, semantic
SP		Linguistic, stylistic
GOP		Strategic intent
Zero-shot	Translation	Unconstrained output
Few-shot		Formal, stylistic

Table 1: Prompt engineering strategies for poetry translation by stage and constraint type.

fore the translation itself, LLMs can also be applied at the translation stage to assist in producing the target text. It is worth acknowledging that not all translators will wish to use LLMs for this purpose: some may prefer to reserve the translation act itself as a space for their own creative agency, drawing on the pre-translation analysis as scaffolding while producing the translation independently. For those who choose to use LLMs at this stage, however, prompt engineering strategies can be used to obtain a full translation, or at minimum a working draft suitable for post-editing.

The quality of the translation output is shaped by how the model is prompted (Peng et al., 2023). A minimal instruction, such as *translate the following poem into English*, constitutes what is known as zero-shot prompting (Brown et al., 2020), in which no examples or demonstrations of the desired output are provided. These prompts can produce acceptable results, but they usually generate the most predictable kind of text, and such predictable translation will not necessarily produce outputs that align with the translator’s intentions, especially in the context of poetry, whose value lies in its creative, unexpected, and formally constrained use of language. Zero-shot prompting is therefore unlikely to reproduce the formal and stylistic features of the source poem. For instance, when we prompted ChatGPT 5.2 (thinking mode) with *Please, translate this poem into English*, the model returned the following:

*Tell me where, in what land,
Is Flora, the lovely Roman;
Archipiada, nor Thai’s,
Who was her own first cousin;
Echo, who answers when sound is carried
Over river or over pond,
Whose beauty was more than human?
But where are the snows of yesteryear?
Where is the most wise Héloïse,
For whom Peter Abelard was castrated*

*And then became a monk at Saint-Denis?
For love of her he suffered that trial.
Likewise, where is the queen
Who ordered that Buridan
Be thrown in a sack into the Seine?
But where are the snows of yesteryear?
The queen Blanche, white as a lily,
Who sang with a siren's voice;
Bertha of the big foot, Beatrice, Alice;
Harembourg, who held Maine,
And Joan, the good woman of Lorraine,
Whom the English burned at Rouen—
Where are they, O sovereign Virgin?...
But where are the snows of yesteryear?*

I can also make it more literal or more poetic, depending on what you need.

The output demonstrates that the model focuses on translating the meaning of the words in a straightforward manner, while the rhyme scheme, among other poetic features, is not reproduced. However, at the end of the poem, the model tells the user it can make the translation more literal or more "poetic", though it is not clear what the model means by that term. In this scenario, the translator can continue prompting the model, accepting its suggestions or using a one-shot strategy in which the user provides one example of output so that the model can learn from it, or a few-shot strategy, which involves providing two or more examples to the model. For instance, for François Villon's poem *Ballade des dames du temps jadis* the few-shot prompting strategy can be designed as follows:

Main prompt:

François Villon's poem *Ballade des dames du temps jadis* contains the following rhyme scheme ababbcbC ababbcbC ababbcbC bcbC. Translate this poem into [language x] and reproduce the source text's rhyme scheme. Examples of how the rhyme scheme works in this poem are provided below:

A: *pays, Thaïs, Heloïs, Sainct-Denys*
B: *Romaine, germaine, maine, humaine, moyne, essoyné, royne, Seine, sereine, Mayne, Lorraine, souveraine*
C: *estan, d'antan, Buridan, d'antan, Rouen,*

d'antan, cest an, d'antan

This prompt was designed to explicitly show to the model how the rhyming scheme works so that the model can learn from this information and generalise across similar tasks. The literature shows that few-shot prompting, i.e., when the model is provided with one or more examples of the desired output format before being asked to produce its own, yields better results across a range of generation tasks (Brown et al., 2020). In the context of poetry translation, this means demonstrating the target output, for instance a translation that preserves a specific rhyme scheme or metrical pattern, before asking the model to apply the same approach to the source poem. Rather than leaving the model to infer what a good translation looks like, the few-shot approach provides explicit evidence of the translator's priorities and expectations.

In our previous work (Resende and Hadley, 2024), we compared zero-shot and few-shot prompting for the specific task of reproducing rhyme schemes in the translation of sonnets. Consistent with findings from other translation tasks (Karpinska and Iyyer, 2023), the few-shot approach produced better results, increasing the rate of rhyme scheme reproduction by 14% for ChatGPT models. However, it is worth noting that, as evidenced by our results, a model's performance in reproducing rhyme schemes may vary from model to model and between language pairs. It is also worth noting that over time, models tend to improve performance on a variety of tasks, including reproducing rhyme schemes and other tasks related to translation. In this section we experiment with a zero-shot approach applied to François Villon's poem *Ballade des dames du temps jadis* using Claude Sonnet 4.6, in order to analyse what kind of translation is obtained with the simplest prompting strategy.

The prompt used was:

*Please, translate this poem into English:
[poem]*

The model returned the poem translated into English along with an explanation of its strategy. The complete model output can be seen in Appendix 3. What is interesting in this output is that, despite not producing the rhyme scheme of the source poem, the model itself "detects" the possibility of deliv-

ering a poem reproducing a rhyme scheme and asks the user "Would you like an alternative version with a stricter rhyme scheme, or a more literal/scholarly rendering?". Such behaviour has not been observed in ChatGPT models and the Gemini model tested in (Resende and Hadley, 2024), which reflects a certain level of improvement in handling tasks of such complexity. When we reply "yes" to this question, the model returns an analysis of the rhyming of the source, clarifying the rhymes at each level of the scheme as follows:

Let me take a proper attempt at holding ABAB-BCBC throughout, with the envoi BCBC, and only three rhyme sounds.

I'll use:

A: -ain/-ane (plain, remain, Seine, Lorraine...)

B: -is/-ese (this, bliss, Héloïse...)

C: -ear/-ere (year, here, yesteryear...)

The output is:

Ballad of the Ladies of Bygone Days

after François Villon - strict scheme

*Tell me in what country, what domain [A],
Dwells Flora, Rome's great beauty, after this [B]?
Archipiades, and Thaïs plain [A],
Her cousin and her counterpart in bliss [B]?
And Echo, whose voice haunts the precipice [B]
Of every pool and river, cold and clear [C],
Whose loveliness surpassed all artifice? [B]
But where are the snows of yesteryear? [C]*

*Where is wise Héloïse, her name [A]
Still bright, for whom was shorn of all amiss [B]
Pierre Abélard — made monk, unmanned, and lame [A]
In love — who bore so dark a wound as this? [B]
And where is that queen whose cold emphasis [B]
Sent Buridan in a sack to disappear [C]
Beneath the Seine's indifferent genesis? [B]
But where are the snows of yesteryear? [C]*

*Queen Blanche, white lily of a great domain, [A]
Who sang with a voice no mortal throat could miss, [B]
Bertha Broadfoot, Beatrix, Alice plain, [A]
Arembour who held Maine in iron fess — [B]
And Joan, the good maid, burned to the abyss [B]
By English hands at Rouen — year by year, [C]
Where are they, sovereign Virgin? Answer this. [B]
But where are the snows of yesteryear? [C]*

*Prince, ask not of the season nor the year, [C]
Nor seek the week or month where all dismiss — [B]
One answer only haunts the listening ear: [C]
But where are the snows of yesteryear? [C]*

The rhyme scheme is maintained in the three first stanzas (ABABBCBC / ABABBCBC / ABABACBC), but in the final four lines, where the envoi should follow the scheme BCBC, the output instead yields CBCC. With a simple prompt asking to fix the error, the model is capable of returning the corrected envoi:

*Prince, do not ask what hour or world is this, [B]
Nor what week holds them, nor what vanished year — [C]
The only answer echoes in the abyss: [B]
But where are the snows of yesteryear? [C]*

For comparison purposes between languages, we also asked the Claude Sonnet 4.6 model to translate the same poem into Portuguese while reproducing the source's rhyme scheme. Using only a zero-shot prompt such as *Please, translate this poem into Portuguese*, the model returned a Portuguese version with the exact rhyme scheme of the source, again illustrating that performance may vary between language pairs. The output of the Portuguese translation can be seen in Appendix 3. Of course, it is the translator's responsibility to determine whether the rhyming scheme proposed by the model aligns with their objectives and aesthetic preferences. However, the experiments presented in this section illustrate that, at the current stage of generative AI development, a simple zero-shot prompt is capable of delivering a working first draft of a formally complex poem: one that reproduces, at least partially, the rhyme scheme of the source text. More strikingly, the model demonstrated an awareness of its own output not previously observed in comparable evaluations: rather than passively delivering an unconstrained translation, it detected the possibility of a more formally faithful rendering and invited the user to specify their translational goals before proceeding (see 3). This capacity for goal-eliciting dialogue, where the model itself asks the user whether they wish to prioritise formal reproduction, literal accuracy, or another criterion, represents a qualitative shift in what LLM-assisted poetry translation can offer, moving beyond single-output genera-

tion toward something closer to a collaborative negotiation of translational priorities. Furthermore, when asked to attempt a stricter reproduction of the rhyme scheme, the model returned an explicit analysis of its own performance, identifying where the scheme held and where it broke down, and providing the user with the information needed to guide subsequent refinement. This self-analytical capacity, combined with the ability to correct specific errors on instruction, suggests that even without the elaborate pre-translation scaffolding described above, LLMs can function as active interlocutors in the translation process rather than passive generation tools. Variation between language pairs, as illustrated by the comparison between the English and Portuguese outputs, further confirms that performance remains uneven and context-dependent which represents a reminder that human oversight and domain knowledge remain indispensable to the workflow.

3 Conclusions

This paper has introduced a set of prompt engineering strategies that have demonstrated measurable improvements in model performance across a range of tasks, including those specific to poetry translation. Its broader aim has been to make the case for LLMs as assistive tools for literary translators across the different stages of the poetry translation process.

At the pre-translation stage, LLMs can surface the multiple layers of complexity a poem carries such as formal, linguistic, cultural, and symbolic, returning structured output that supports the translator’s decision-making while also drawing attention to dimensions of the text that might otherwise remain implicit or go unnoticed. At the translation stage, these models can produce a working first draft from which the translator can depart, or, with appropriate prompt engineering, a more formally constrained version targeting specific features of the source text such as the rhyme scheme. Notably, with Claude Sonnet 4.6, we observed proactive behaviour in which the model independently detected the possibility of a more formally faithful rendering and invited the user to specify their goals before proceeding, suggesting that state-of-the-art models are evolving toward more complex reasoning capabilities that demand progressively less prompt engineering expertise from the translator, a behaviour not ob-

served in our previous experiments (Resende and Hadley, 2024). This is a significant development, as it lowers the barrier to entry for translators who may lack technical familiarity with prompt engineering but wish to engage with these tools. It is worth noting, however, that such capabilities vary across LLMs and language pairs, and should not be assumed to be uniform. Their relative strength in producing or rendering English poetry is at least partly attributable to the disproportionate volume of English-language training data available to them (Jackson et al., 2023). For poetry translated into less-resourced languages, it could be that the utility of these tools may be less immediately apparent, and the limitations more pronounced.

In offering this overview, we also aim to promote a conscious and informed engagement with these tools, one grounded in an understanding of both their capabilities and their constraints. More broadly, we hope to contribute to a growing conversation that seeks to challenge the assumption that literary translation, and poetry translation in particular, lies beyond the reach of computational assistance. Our argument is not that LLMs replace human creativity or judgement, but that they can extend it: by providing access to a breadth of linguistic and cultural data that no individual translator can retain simultaneously, they offer the possibility of augmenting, rather than displacing, the creative process.

What is clear is that the intersection of technology and poetry translation represents an emerging field of research, one with significant potential for questions that could not previously have been asked, let alone answered. Whether translators of poetry will ultimately come to see LLMs as valuable collaborators and begin making sustained, informed use of them as part of their practice remains an open question. But the conditions for that conversation have never been more present.

References

- Bassnett, Susan. 1998. Transplanting the seed: Poetry and translation. In *Constructing Cultures*, pages 57–75. Multilingual Matters, Bristol and Blue Ridge Summit.
- Bellos, David. 2011. *Is That a Fish in Your Ear? Translation and the Meaning of Everything*. Penguin Books, London.
- Brown, Tom, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D. Kaplan, Prafulla Dhariwal, Arvind

- Neelakantan, Pranav Shyam, Girish Sastry, and Amanda Askell. 2020. Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33:1877–1901.
- Clunies Ross, Margaret. 2010. *The Cambridge Introduction to the Old Norse-Icelandic Saga*. Cambridge University Press.
- Fabb, Nigel. 2015. *What is Poetry?: Language and Memory in the Poems of the World*. Cambridge University Press, Cambridge.
- Forcada, Mikel L. 2017. Making sense of neural machine translation. *Translation Spaces*, 6(2):291–309.
- Hume, Nancy G., editor. 1995. *Japanese Aesthetics and Culture: A Reader*. State University of New York Press. Accessed: 2026-04-22.
- Jackson, Samantha, Barend Beekhuizen, Zhao Zhao, Yi Cheng Zhao, and Rhonda N. McEwen. 2023. Llms and linguistic competency: An exploration of gpt-4 and a non-hegemonic english variety. *New-house Impact Journal*, 1(1):Article 8.
- Karpinska, Marzena and Mohit Iyyer. 2023. Large language models effectively leverage document-level context for literary translation, but critical errors persist. In Koehn, Philipp, Barry Haddow, Tom Kocmi, and Christof Monz, editors, *Proceedings of the Eighth Conference on Machine Translation*, pages 419–451, Singapore, December. Association for Computational Linguistics.
- Lefevre, André. 1975. *Translating Poetry: Seven Strategies and a Blueprint*, volume 3 of *Approaches to Translation Studies*. Van Gorcum, Assen.
- Li, Haochen, Jonathan Leung, and Zhiqi Shen. 2024. Towards goal-oriented prompt engineering for large language models: A survey.
- Newman, John B. 1958. The meanings of poetry. *Today's Speech*, 6(3):28–30.
- Peng, Keqin, Liang Ding, Qihuang Zhong, Li Shen, Xuebo Liu, Min Zhang, Yuanxin Ouyang, and Dacheng Tao. 2023. Towards making the most of ChatGPT for machine translation. In Bouamor, Houda, Juan Pino, and Kalika Bali, editors, *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 5622–5633, Singapore, December. Association for Computational Linguistics.
- Resende, Natália and James Hadley. 2024. The translator's canvas: Using llms to enhance poetry translation. In *Proceedings of the 16th Conference of the Association for Machine Translation in the Americas (Volume 1: Research Track)*.
- Stageberg, Norman C. 1948. The aesthetic of the petrarchan sonnet. *The Journal of Aesthetics and Art Criticism*, 7(2):132–137.
- Suojanen, Tytti, Kaisa Koskinen, and Tiina Tuominen. 2014. *User-centered translation*. Routledge.
- Suzgun, Mirac and Adam Tauman Kalai. 2024. Meta-prompting: Enhancing language models with task-agnostic scaffolding.
- Wei, Jason, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Brian Ichter, Fei Xia, Ed H. Chi, Quoc V. Le, and Denny Zhou. 2022. Chain of thought prompting elicits reasoning in large language models. In Oh, Alice H., Alekh Agarwal, Danielle Belgrave, and Kyunghyun Cho, editors, *Advances in Neural Information Processing Systems*.

Appendices

Appendix A

Step 1 response

In classical Japanese poetry — especially in traditions like waka and later haiku — the word 花 (hana, “flowers”) carries rich and layered symbolism in the context of spring. While it can refer to flowers in general, it most often implicitly means cherry blossoms (sakura).

Here are the main symbolic meanings:

1. Ephemerality (無常, *mujō*)

Flowers — particularly cherry blossoms — bloom beautifully but fall quickly. This makes them a central symbol of the Buddhist idea of impermanence: life is fleeting, beauty is transient, and moments should be appreciated because they pass. This idea is deeply tied to Japanese aesthetics and worldview.

2. Beauty and Peak Moment

花 represents beauty at its height: the blossoms are admired precisely at full bloom, this peak moment is intense but short-lived, and in poetry, this often reflects human experiences like youth, love, or happiness at their most vivid point.

3. Melancholy and Awareness (*mono no aware*)

Closely related to impermanence is *mono no aware* — a sensitivity to the sadness of things passing: falling petals evoke quiet sorrow, beauty is moving because it does not last. This emotional tone is central in classical works such as *The Tale of Genji*.

4. Seasonal Marker (Spring / Renewal)

In the poetic system of seasonal references (*kigo*), 花 signals springtime, renewal and new beginnings, and social gatherings like flower viewing (*hanami*). However, unlike purely joyful spring symbols, 花 often carries a bittersweet nuance.

5. Love and Human Relationships

Flowers can also symbolise romantic encounters (often fleeting), the fragility of relationships, and separation and longing: a brief bloom may mirror a brief romance.

In short: 花 in spring poetry = beauty + impermanence + emotional depth. It is not just “flowers”, but a compressed cultural symbol combining nature, time, and human feeling.

Step 2 response

散る (*chiru*, “to fall/scatter”) strongly deepens the symbolism of 花 beyond a literal description.

In classical poetry, it evokes: the moment of loss — not just that beauty is brief, but the instant it disappears; emotional climax, as the falling petals often mark the peak of *mono no aware*; resignation and acceptance, a quiet acknowledgement that decline is natural and inevitable; and parallels to human life, including fading love, aging, or death. So while 花 suggests fleeting beauty, 散る dramatises its passing, turning observation into emotional insight.

Step 3 response

Here are concise interpretations of 花が散る / 花の散り that preserve both spring imagery and connotations of transience and mortality:

- Falling cherry blossoms → peak spring beauty already slipping away
- Blossoms scattering in the wind → fragility of life under uncontrollable forces
- Petals drifting to the ground → quiet, natural passing (death without violence)
- A sudden shower of petals → abrupt loss or the swiftness of change
- Blossoms at the end of full bloom → the moment beauty turns into decline
- Petals covering the earth or water → aftermath of life; memory or trace of what was
- Continuous falling over days → gradual fading (aging, slow separation)
- Wind-driven scattering → external causes shaping one's fate
- Petals dissolving or disappearing → complete impermanence; return to nature
- Late blossoms falling → lingering attachment before an inevitable end

All of these keep the seasonal reference to spring (sakura) while foregrounding impermanence, loss, and mortality, rather than merely describing a physical event.

Appendix B

1. Sound, Rhythm, Lineation, Syntactic Compression

Sound: Nasal vowels and liquid consonants (l, r) create a flowing, melancholic resonance. Rhythm

& Lineation: Octosyllabic lines with regular balade structure; refrain reinforces musicality. Syntactic Compression: Entire narratives (e.g., lives of famous women) condensed into single lines, requiring dense semantic packing.

2. Lexical Ambiguities & Symbolic Images

Ambiguities in words like *temps jadis* (past time, “olden days”) and references to women’s names. Symbolic images: transience, memory, mortality, lost beauty.

3. Cultural References & Allusions

Historical and legendary women (e.g., Héloïse, Jeanne d’Arc). Medieval French cultural knowledge; some names carry moral or poetic connotations.

4. Main Translation Challenges

Reproducing sound texture in English (nasals, liquids). Maintaining octosyllabic rhythm. Conveying syntactic compression without loss of meaning. Handling culturally specific allusions: domestication vs. foreignization. Preserving lexical ambiguities and poetic resonance.

If you want, I can also suggest specific strategies for each challenge to guide your English translation.

Appendix C

This is a rich pre-translation analysis task. Let me work through the poem carefully and then present the findings as a structured visual reference.

Pre-translation feature map — Villon, Ballade des dames	
Must be preserved	May be adapted
Refrain: exact wording Mais où sont les neiges d’antan? must recur verbatim (four times) — its weight is cumulative	Rhyme scheme (ababbcbc) Faithful rhyme in English usually requires lexical distortion; prose or assonance is defensible
Stanzaic structure Eight-line stanzas (huitains) + four-line envoi — the form is part of the poem’s argument	Octosyllabic line length English iambic tetrameter is a natural equivalent, but stress patterns need not mirror French
Proper names (all fourteen) Flora, Thaïs, Echo, Héloïs, Buridan, Blanche, Berthe, Jeanne — each carries a historical gloss	Archaic morphology Roi/ne for reine, essoine, antan — archaism may be signalled by different means in English
Rhetorical mode: interrogation Every stanza is a question, never an elegy — the poem asks, it does not mourn directly	Syntax: inversion & delay Dites-moi où, n’en quel pays — rhetorical effect must be kept but exact word order may shift
Hierarchy of exempla Mythological → romantic → royal → Jeanne d’Arc (martyrdom last) — sequence is the argument	Proper name variants Archipiades (Alcibiades?), Bietrix, Aliz — spelling may be standardised with annotation
Envoi: address to Prince Conventional ballade close, but the dismissal (n’enquerez) is philosophically crucial	Historical allusions Buridan’s sack in Seine — apocryphal; may be glossed rather than paraphrased in line
Register: plain diction Villon avoids elevated rhetoric; the vernacular directness is itself the aesthetic statement	Repetition markers Sembablement, où est — anaphora should be reflected; exact connector word may differ
The neige / antan image Untranslatable compound (antan = yesteryear, not last year); any translation must flag this in notes	

Features that must be preserved

The refrain in its exact semantic content. “Mais

où sont les neiges d’antan?” is not decorative: it performs the poem’s argument each time it recurs, converting the preceding catalogue of names into a single unanswered question. The word *antan* — Old French for “the year before last”, hence “yesteryear” rather than “last year” — has no clean English equivalent. Any translation is already an interpretation here, but a scholarly edition must signal that loss in a note rather than normalising it away. The four recurrences must be identical in the translation; variation would destroy the cumulative effect.

The stanzaic form: *huitain* plus envoi. The balade is a genre with structural rules, and those rules carry meaning. The envoi in particular — the address to the “Prince”, the dismissal of the question — is not padding but the poem’s philosophical conclusion: stop asking when, because the refrain already answered you. Any translation that collapses stanzas or drops the envoi misrepresents the argument.

All fourteen proper names, in their original positions. Flora, Thaïs, Archipiades, Echo, Héloïs, Buridan, Blanche, Berthe, Bietrix, Aliz, Haramburgis, Jeanne — each is a weight-bearing node. The poem does not describe these women; it lists them, trusting the name to conjure the person. Replacing any name with a description (“the abbess Héloïse”) would shift the genre from catalogue-poem to exposition. The scholarship belongs in the footnotes.

The interrogative mode throughout. This is an elegy that refuses the elegiac statement. Villon never says “they are gone” — he asks “where are they?” That rhetorical stance is the poem’s most distinctive intellectual move, and it must be preserved as a grammatical and tonal fact of the translation.

The ordering of exempla. The movement from classical-mythological (Flora, Thaïs) through medieval romantic (Héloïse and Abélard) to royal figures (Blanche de Castille, the queen of the Buridan legend) and finally to Jeanne d’Arc — the only figure whose death is named explicitly, and named by a specific atrocity — is a deliberate crescendo. The sequence is argument, not decoration.

The plain, vernacular register. Villon does not use elevated diction. The poem’s emotional force comes from the gap between its modest language and the enormity of what it contemplates. An English translation that reaches for archaism-as-

beauty misrepresents this.

Features that may be adapted

The rhyme scheme. The ballade's ababbcbc octosyllabic pattern is close-woven in French in ways that cannot be reproduced in English without serious lexical distortion. An unrhymed or loosely assonant translation that preserves meaning is more defensible in a scholarly context than one that rhymes at the cost of sense. This is a well-documented translator's dilemma for this poem specifically (Swinburne's famous version makes precisely this sacrifice). Exact line length. French octosyllables have eight syllables by count; English stress-syllabic verse does not map onto this directly. English iambic tetrameter is a reasonable functional equivalent, but it need not be enforced mechanically. Archaic morphology as such. *Roine* for *reine*, *essoine* (legal constraint or excuse), the contracted forms — these signal temporal distance in the French. In English, different means (slightly elevated diction, deliberate archaism of a different kind) can signal the same distance without pretending to mirror the specific features of late medieval French. Syntactic inversions. The opening "Dites-moi où, n'en quel pays" foregrounds the question dramatically. The inversion should produce a comparable rhetorical effect in English, but the exact grammatical strategy need not be calqued. Spelling variants of the proper names. Archipiades (almost certainly a corruption of Alcibiades, here cast as a woman, which is itself editorially interesting), Bietrix and Aliz (variant spellings of Beatrice and Alice) may be standardised in the translation with a note, since the variant spelling conveys nothing to an English reader that the standard form would not. The Buridan allusion. The story of Jean Buridan thrown in a sack into the Seine by the queen is apocryphal and has no firm historical basis — in fact the legend may post-date Villon. In a translation for a scholarly bilingual edition, this is better handled by annotation than by any attempt to domesticate the image in line.

Appendix D

Here is a translation that tries to preserve Villon's tone — the mix of erudite melancholy, rhetorical formality, and the haunting simplicity of that refrain:

Ballad of the Ladies of Bygone Days

English: after François Villon

*Tell me where, and in what land,
Is Flora, loveliest of Rome,
Archipiades, and Thaïs,
Who was her cousin and her own?
And Echo, speaking when sound is thrown
Across a river or a pool,
Whose beauty passed what flesh has known?
But where are the snows of yesteryear?*

*Where is the wise Héloïse,
For whose love was gelded, then made monk,
Pierre Abélard at Saint-Denis?
He bore that wound for her alone.
And where is that queen, likewise gone,
Who ordered Buridan be flung
In a sack into the Seine and drowned?
But where are the snows of yesteryear?*

*Queen Blanche, white as the lily's face,
Who sang with a siren's silver voice,
Bertha Broadfoot, Beatrix, Alice,
Arembour who held Maine by choice —
And Joan, the good maid of Lorraine,
Whom Englishmen burned at Rouen's stake —
Where are they, sovereign Virgin? Say.
But where are the snows of yesteryear?*

*Prince, do not ask from week to week,
Nor ask what year or season's name —
The only answer you will reap:
But where are the snows of yesteryear?*

A few notes on choices made: "yesteryear" for *antan* is Rossetti's famous coinage and has become inseparable from this poem in English — it earns its keep. I've kept the envoi's direct address to the "Prince" (the conventional dedicatee of a ballade), and tried to preserve the rhetorical movement of each stanza: accumulation of names, then the refrain landing like a door closing. The women's names are kept recognisable without heavy annotation — Abélard is spelt in the modern form since it's the more familiar English rendering.

Would you like an alternative version with a stricter rhyme scheme, or a more literal/scholarly rendering?

Appendix E

Let me work this out carefully. Villon's scheme is ABABBCBC per stanza, envoi BCBC, with only three rhyme sounds throughout. Portuguese is much richer in rhyme than English, so this should be more tractable.

I'll use:

A: -anal-ã (romana, humana, Bretanha...)

B: -ois/-ês (Héloïse, depois, burgês...)

C: anchor on outrora (-ora: outrora, agora, demora, aurora, chora)

The refrain is the anchor: “Mas onde estão as neves de outrora?”

Balada das Damas do Tempo que Foi

após François Villon

*Dizei-me em que terra, em que paragem,
Está Flora, a bela, a mais romana,
Arquipiádes, Taís, cuja imagem
Foi sua prima, flor da raça humana,
E Eco, que fala quando a voz se engana
Sobre o lago ou o rio, aonde demora,
Cuja beleza ultrapassa o que é mundana?
Mas onde estão as neves de outrora?*

*Onde está a sábia Heloísa, depois
De quem foi castrado e feito monge
Pierre Abelardo — e o amor, através
Da dor, durou, ainda que se alongue?
E onde está a rainha que, além,
Mandou que Buridan, sem mais demora,
Fosse ao Sena num saco? Onde também?
Mas onde estão as neves de outrora?*

*A rainha Branca, lírio da Bretanha,
Que cantava com voz de sereia pura,
Berta do pé grande, Beatriz, estranha
Aramburgis que o Maine teve à sua cura,
E Joana, a boa, de alma tão segura,
Que os ingleses queimaram em má hora —
Onde estão, Virgem, Mãe, ó Virgem pura?
Mas onde estão as neves de outrora?*

*Príncipe, não pergunteis, sem demora,
Onde estão, nem em que ano ou estação —
Pois só este refrão vos ecoa:
Mas onde estão as neves de outrora?*