Towards Supporting Literary Studies Using Virtual Reality and Generative Artificial Intelligence

Jens Grubert*
Coburg University of Applied Sciences

Dieter Schmalstieg†
University of Stuttgart

Kirsten Dickhaut‡
University of Stuttgart

ABSTRACT

Literary studies critically examine fictional texts, exploring their structures, themes, stylistic features, and cultural-historical contexts. A central challenge in this field lies in bridging textual analysis with the spatial and sensory dimensions of settings described or implied in texts. Traditional methodologies often require scholars to mentally reconstruct these environments, leading to incomplete or inconsistent interpretations. Readers may be biased by their personal context or experiences, or may lack detailed knowledge of the relevant historical facts. This paper argues for the integration of virtual reality and generative artificial intelligence as supporting instruments to enhance literary research. The former enables immersive, spatially accurate reconstructions of historical environments, while the latter provides tools such as text-to-image and text-to-3D generation which let us dynamically render visual elements quoted in literary texts. Together, these technologies have the potential to significantly enhance traditional literature analysis methodologies, enabling novel approaches for contextualizing and analyzing literature in its spatial and cultural milieu.

Index Terms: literary studies, virtual reality, generative artificial intelligence

1 INTRODUCTION

Literary studies are dedicated to the critical examination of fictional texts. This typically encompasses the structure, themes, and stylistic characteristics of the text, as well as the cultural and historical contexts in which those texts were produced. Researchers of literary studies explore various literary works, ranging from novels and poetry to historical documents and correspondence, with the goal of uncovering elements going beyond the mere text, such as cultural values or societal influences. The contextualization of texts is central to this discipline. This often involves reconstructing the physical, political, and cultural environments that influenced the examined literature.

One key challenge in literary studies lies in bridging textual analysis with the spatial and sensory dimensions of the settings described or implied in the texts. Established methodologies typically use hermeneutics to create text interpretations that are culturally and historically appropriate. Central to this are stylistic, rhetorical, and narratological analyses. Those established methodologies require researchers to imagine the content studied based on abstract text (and sometimes pictorial) sources. For example, understanding the significance of space in a literary work requires researchers to mentally reconstruct settings based on limited descriptions, which can be challenging and lead to incomplete or inconsistent interpretations.

Computational tools in literary studies have begun to address some of these limitations, e.g., through the use of machine learning algorithms for co-reference resolution or sentiment analysis [15, 12]. At the same time, the application of purely quantitative approaches often favored by engineering sciences has also faced criticism. In the extreme, scholars may dismiss the use of computational tools in literary studies, such as Da [4] who states "At the same time, what is robust is obvious (in the empirical sense) and what is not obvious is not robust".

In this paper, we argue that the use of computational methods as a supporting instrument for creating context has merit. Beyond tools that merely reduce tedious work, such as summarizing texts or tabulating facts, computational tools or *explaining* could complement the humanities-driven goal of *understanding* [21].

Specifically, we think that the recent developments in virtual reality (VR) and generative artificial intelligence (GenAI) can serve this purpose. VR offers the ability to recreate historical environments, enabling researchers to study texts in immersive and, ideally, spatially accurate settings. When combined with GenAI tools such as text-to-image or text-to-3D generation [18], VR environments can dynamically provide visual, potentially interactive elements evocated in literary texts.

2 APPLICATION OF VR AND GENAI TO STUDYING LITERATURE ABOUT VERSAILLES

In this paper, we propose researching such tools in the specific context of literary studies about Versailles (e.g., a novel by Scudery [5]). We argue that the literature associated with the Palace of Versailles can benefit from understanding the architectural grandeur, court life, and historical significance of Versailles in early modern France to fully grasp the thematic and spatial nuances of the texts, as well as the network of texts related to the garden. The concept we present relies on complementarity by visualizing what has been historically lost. This approach enables new, previously unimaginable questions to arise, such as whether Apollo truly represented Louis XIV in the Grotto of Thetis, illuminated by the sunlight at sunset. To this end, we propose using 1) an immersive environment of the Versailles Garden and 2) to enhance literature studies and open up new methodological approaches in that virtual garden through complementary GenAI tools.

2.1 Immersive Versailles Garden

While immersive experiences about Versailles exist[1], they are typically prescripted and do not serve the needs of literary studies. Instead, we propose an interactive virtual environment that integrates with the workflow of researchers studying historical texts about Versailles. Specifically, we propose to create a virtual replica of the Versailles Garden using a combination of existing data sources and procedurally generated 3D assets. We plan to start by generating a spatially correct base layout from geospatial data sources such as height maps[2]. This will likely result in a geographically correct, but visual low-fidelity 3D environment. Hence, we propose to enhance this base environment with a variety of digital assets.

*e-mail: jens.grubert@hs-coburg.de
†e-mail: schmaldr@visus.uni-stuttgart.de
‡e-mail: kirsten.dickhaut@ilw.uni-stuttgart.de

[1] https://en.chateauversailles.fr/news/life-estate/experience-versailles, last access, 16.12.2024

[2] https://geoservices.ign.fr/, last access 16.12.2024

Given knowledge of the species of vegetation present in historical French gardens, plant models can be obtained from commercial or open-source asset libraries (e.g. Objaverse [6]) or generated with specialist tools such as SpeedTree[3]. Artistic pieces, such as statues, which are fully preserved or at least documented through paintings or photographs, can be reconstructed using GenAI tools such as Trellis3D [22] with a fraction of the effort required to create a detailed 3D asset even a few years ago. The display of 3D content generated with these novel methods, for example in the form of 3D Gaussian splatting models, is rapidly making its way into professional game engines[4].

The resulting virtual environment can be freely navigated in VR, using conventional VR navigation techniques such as teleporting or walking-in-place, or through the use of omnidirectional treadmills. The latter has the advantage that realistic perception of spatial scales is heightened by improved immersion and embodiment. For example, following Madame de Scudery's trajectory through the imperial gardens at Versailles can be experienced in full scale, potentially physically walking the entire path of several kilometers extend described in the novel.

Interaction with the literature can be facilitated through the use of appropriate text displays (cf. [20]) and potentially gaze-supported reading aids [16], which allow a new way to experience the literature. Note-taking can be facilitated by using handwriting or symbolic text entry [7]. We also see the potential in temporarily rearranging textual artifacts, either in the virtual garden (e.g., using automatic layouts [3]) or by temporarily transitioning out of the spatial context of the garden into a text-centric layout environment facilitating sensemaking [19]. It seems advisable to both keep the geospatial reference of the garden even in a text-centric environment (e.g., through a mini-map) and to support the user in efficiently transitioning between the two environments.

Through such a bespoke virtual environment, our goal is to allow researchers to connect textual references to their corresponding physical forms. For example, descriptions of a particular fountain, grotto, or architectural feature in the Garden of Versailles can be visualized and explored interactively. Moreover, the difference between fiction, poetics of marvelous, panegyrics, and historical verisimilitude are in such referential texts can be easier to grasp and understand via VR.

Given the base immersive virtual environment, we foresee the use of the following instruments for querying and contextualizing literary works within the immersive environment itself to facilitate in situ text analysis.

1) Speech-based and text-based interaction with the literature and the environment can be supported using a combination of automatic speech recognition ([13] and large language models (LLM) capable of real-time responses, such as Llama-omni [8]. The relevant literature can be made available to the system using retrieval augmented generation [17]. 2) Teleportation to selected places within the Versailles Garden can be achieved on the basis of the results of interaction with the LLM. For example, a user could query which fountains are mentioned in a specific text and select a fountain shown in the LLM answer to navigate to and continue the inquiry at that specific virtual fountain. 3) Integration of text-to-image and text-to-3D tools: Similarly to the recognition of places, based on a text passage, various objects can be identified. Those objects, along with a description, could be used as prompt input to generate images or 3D objects to further add visual context to the literature study. Further, multimodal input techniques can be used for interaction with the generated scene elements [2].

[3] https://www.speedtree.com/, last access 16.12.2024

[4] e.g., https://github.com/xverse-engine/XV3DGS-UEPlugin, last access 16.12.2024

2.2 Challenges

The development of a VR environment enriched with Gen AI that is suitable for performing a cognitively advanced task such as literary analysis is extremely ambitious. Limited prior work exists to guide us, and we expect to face several challenges. One is the suitability of a VR headset for use over longer periods of time [1]. Recent VR headsets such as the Apple Vision Pro make a leap forward in image quality, but still fall short of expectations in terms of device weight and comfort. Even with a high native display resolution, reading text presented through the optical system of a VR headset is likely still more strenuous than reading text.

The quality of the 3D assets may also be the subject of concern. While the overall visual quality of assets produced with GenAI is now on a more than acceptable level, the method faces intrinsic problems with control of details, inaccuracies and occasional tendencies to hallucinate. These characteristics can be at odd with a scholar's need for factual or historical accuracy. Further, the generation process can be subject to bias in the underlying models [9] as also noted for other applications of GenAI in XR [11].

Beyond these primarily technical limitations, we will also have to explore how a visual analytics system, especially an immersive one, fits the workstyle of literary scholars and what 3D tools prove useful to their goals and practices.

Another potential challenge could be that presenting generated images and 3D models could influence the mental imagery that a reader has when reading the literature. It should be investigated if and how this would impact literacy study tasks.

Finally, methodological challenges will also be significant, because historical sources are sparse, so that fictional texts fill the gap left by historiographical texts. Consequently, it is important to combine precise text analyses with precise prompting in order to obtain adequate results.

On the notion of fictionality and fictiveness, another observation can be added. We will work through this together using an example that will be briefly explained here, along with its significance, as it involves the above mentioned fictional text: Madeleine de Scudéry's novel "La promenade de Versailles" [5] is a novel structured as a dialogue, a genre quite common in the early modern period. The dialogic element, unlike purely narrative prose, directly functions as a marker of fiction. This makes the dialogic novel clearly a fictional text, and indeed one that employs strong markers of fictionality. Conversely, the content of the characters' conversations is not fictional but highly real, as they discuss the garden of Versailles—in particular, the contemporary design, which was novel to everyone at the time, given that the park had only recently been conceived and laid out by André Le Nôtre.

To better grasp the relationship between fictionality and fictivity, we may cite the relevant reflections of Kablitz [14, p. 169, our translation], who emphasizes that the degree of fictionality in representation is scalable: "*What is represented can indeed be more or less fictive; and this possibility exists for fictional texts no less than for 'factual' ones. Thus, what is scalable is the fictive, but not the fictional.*"

This assertion is based on a fundamental observation that Kablitz explains earlier in his argument, where he outlines the premises grounded in Grice's conversational maxims [10, pp- 64-75] . This key passage of Kablitz [14, p. 166, our translation] should therefore also be quoted: "*It seems crucial to me to distinguish between the fictive as a property of what is represented and the fictional as a property of representation. This distinction already makes it clear that literature shares the fictive with a multitude of other texts (and media). Fabricated, invented, hypothetical, simulated, or fictitious elements are, as we know, ubiquitous; but fictional is only literature itself.*"

3 Conclusion and Future Work

This position paper has sketched an attempt to develop a VR system enriched with generative AI to provide visual and spatial context to the task of analyzing historical literature. We have described the imperial gardens of Versailles and Scudery's novel playing in these gardens as an example of the technical measures and design ideas for immersive analytics systems. Our assumption is that combining VR and GenAI for literary studies adds a new dimension to interpretation, allowing a better assessment of the complexity of artistic conditions in their congruence or incongruence. It thereby could provide a groundbreaking basis for literary studies. We hope to report in the near future on early findings of building and studying such a system. Moreover, we hope to discover new ways of systematic comparison between fiction as represented in literature and in VR.

Acknowledgements. We thank the Alexander von Humboldt Foundation funded by the German Federal Ministry of Education and Research.

References

[1] V. Biener, S. Kalamkar, N. Nouri, E. Ofek, M. Pahud, J. J. Dudley, J. Hu, P. O. Kristensson, M. Weerasinghe, K. Č. Pucihar, et al. Quantifying the effects of working in vr for one week. *IEEE TVCG*, 28(11):3810–3820, 2022. 2

[2] J. Chen, J. Grubert, and P. O. Kristensson. Analyzing multimodal interaction strategies for llm-assisted manipulation of 3d scenes. In *2025 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 2025. 2

[3] Y. F. Cheng, C. Gebhardt, and C. Holz. Interactionadapt: Interaction-driven workspace adaptation for situated virtual reality environments. In *ACM UIST '23'*, pp. 1–14, 2023. 2

[4] N. Z. Da. The computational case against computational literary studies. *Critical inquiry*, 45(3):601–639, 2019. 1

[5] M. De Scudéry. La promenade de versailles. 1669. 1, 2

[6] M. Deitke, R. Liu, M. Wallingford, H. Ngo, O. Michel, A. Kusupati, A. Fan, C. Laforte, V. Voleti, S. Y. Gadre, et al. Objaverse-xl: A universe of 10m+ 3d objects. *NIPS '24'*, 36, 2024. 2

[7] T. J. Dube and A. S. Arif. Text entry in virtual reality: A comprehensive review of the literature. In *HCII '19*, pp. 419–437. Springer, 2019. 2

[8] Q. Fang, S. Guo, Y. Zhou, Z. Ma, S. Zhang, and Y. Feng. Llama-omni: Seamless speech interaction with large language models. *arXiv preprint arXiv:2409.06666*, 2024. 2

[9] J. W. Gichoya, K. Thomas, L. A. Celi, N. Safdar, I. Banerjee, J. D. Banja, L. Seyyed-Kalantari, H. Trivedi, and S. Purkayastha. Ai pitfalls and what not to do: mitigating bias in ai. *The British Journal of Radiology*, 96(1150):20230023, 2023. 2

[10] H. P. Grice. Logic and conversation. In D. Davidson, ed., *The logic of grammar*. Dickenson Pub. Co., Encino, CA, USA, 1975. 2

[11] J. Grubert, J. Chen, and P. O. Kristensson. Generative ai for accessible and inclusive extended reality. *arXiv preprint arXiv:2410.23803*, 2024. 2

[12] H. O. Hatzel, H. Stiemer, C. Biemann, and E. Gius. Machine learning in computational literary studies. *it-Information Technology*, 65(4-5):200–217, 2023. 1

[13] N. Jeffries, E. King, M. Kudlur, G. Nicholson, J. Wang, and P. Warden. Moonshine: Speech recognition for live transcription and voice commands. *arXiv preprint arXiv:2410.15608*, 2024. 2

[14] A. Kablitz. *Kunst des Möglichen. Theorie der Literatur*. Freiburg, Germany, 2013. 2

[15] E. Kim and R. Klinger. A survey on sentiment and emotion analysis for computational literary studies. *arXiv preprint arXiv:1808.03137*, 2018. 1

[16] G. Lee, J. Healey, and D. Manocha. Vrdoc: Gaze-based interactions for vr reading experience. In *IEEE ISMAR '2022'*, pp. 787–796. IEEE, 2022. 2

[17] P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, H. Küttler, M. Lewis, W.-t. Yih, T. Rocktäschel, S. Riedel, and D. Kiela. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks, Apr. 2021. arXiv:2005.11401 [cs]. 2

[18] C. Li, C. Zhang, J. Cho, A. Waghwase, L.-H. Lee, F. Rameau, Y. Yang, S.-H. Bae, and C. S. Hong. Generative ai meets 3d: A survey on text-to-3d in aigc era. *arXiv preprint arXiv:2305.06131*, 2023. 1

[19] L. Lisle, K. Davidson, L. Pavanatto, I. A. Tahmid, C. North, and D. A. Bowman. Spaces to think: A comparison of small, large, and immersive displays for the sensemaking process. In *IEEE ISMAR '23'*, pp. 1084–1093. IEEE, 2023. 2

[20] R. Rzayev, P. Ugnivenko, S. Graf, V. Schwind, and N. Henze. Reading in vr: The effect of text presentation type and location. In *ACM CHI '21'*, pp. 1–10, 2021. 2

[21] C. P. Snow and R. Smoluchowski. The two cultures and the scientific revolution, 1961. 1

[22] J. Xiang, Z. Lv, S. Xu, Y. Deng, R. Wang, B. Zhang, D. Chen, X. Tong, and J. Yang. Structured 3D Latents for Scalable and Versatile 3D Generation, Dec. 2024. arXiv:2412.01506 [cs]. doi: 10.48550/arXiv.2412.01506 2