Loremaster: Towards Better Mixed-Initiative Content Co-Creation in the Creative Industries

OLIVER WOOD, Those Beyond, United Kingdom

We present early designs and prototypes for a human-AI co-creative system designed for the creative industries, notably entertainment industries. We specifically address the surge in reproductive storytelling works such as remakes, sequels, and adaptations. Our system leverages retrieval-augmented generation (RAG) and large multimodal models (LMMs) to ensure that generative AI outputs adhere to existing complex creative lore, leveraging and fine-tuning creative direction documents for guidance. Initial applications in text and 2D image generation have yielded promising results. We work towards effectively integrating generative AI within production workflows for the creative and entertainment industries, with future work focusing on extensive testing, improving narrative and visual content accuracy within a broad range of creative contexts.

CCS Concepts: • **Applied computing → Arts and humanities**; • **Human-centered computing → Human computer interaction (HCI)**; • **Computing methodologies → Artificial intelligence**.

Additional Key Words and Phrases: Human-AI co-creation, Mixed initiative co-creation, Creative industries, Game design, Entertainment

1 INTRODUCTION

Reproductions, remakes, reboots, prequels, sequels, spin-offs, and adaptions. These forms of media - referred to in this work as reproductive storytelling - all have different attributes, but share a key feature: creating from pre-existing stories, narratives, and creative lore. In film, there has been an influx in popularity of reproductive storytelling: with 90% of top grossing box office productions worldwide between 2021-2022 being remakes, sequels, spin offs, and adaptions. [16] This influx of reproductive media is likewise noticeable in the games industry: with high profile reproductive games - such as Baldur's Gate 3, Dead Space, Resident Evil 4, Warhammer 40K: Boltgun [13], and many more being released by large and small studios alike. [3] It's safe to say that reproductive storytelling is significant in the current creative industries, particularly entertainment. This poses a new design challenge for employing generative AI in the creative industries- if we ask creatives (art directors, game designers, animators, etc) to use a system which delivers a different outcome for each generated object, how can they ensure that generative outcomes are in line with the pre-existing creative lore with which they are working? We contribute initial designs and early prototypes for a human-AI co-creative system oriented to assist creatives working on reproductive storytelling projects in industry. Leveraging retrieval-augmented generation (RAG) [9] with large language models (LLMs) and multimodal large language models (also known as LMMs or MLLMs) [20], our system introduces a unique concept: employing a creative direction document, or series of documents to guide generative content output according to pre-existing creative boundaries. We test our system on text and 2D image and gain initial user insights (n=2) for further development. Our work is oriented towards game design: however, this system can be applied to any context in which creatives employ generative AI whilst working within specific creative boundaries.

2 SURROUNDING WORK

2.1 Mixed initiative and human-AI co-creativity

This work is situated in the fields of human-AI co-creativity (HAI-CC) and mixed-initiative co-creativity (MI-CC). Although mixed initiative is not clearly defined [8], this paper refers to mixed initiative co-creativity as a collaborative

process where both human and computational participants actively contribute to solving a problem, such as generating content [19]. This means that either party can initiate changes, and there is no requirement for equal contribution from both humans and computers or humans and AI. [10] and [12] suggest that mixed initiative processes can be identified on a continuum between extremes, where either computer or human has full control. [12] propose an ontology that categorizes different types of creative relationships in human-AI co-creation, including mixed-initiative co-creation. We apply this ontology to provide perspective on our contribution. We can consider a standard content generation tool with hand-written prompts and hand-checked output as a 'computer-as-subcontractor' [12] relationship, where the AI takes ownership of only a specific part of the creation process within the boundaries defined by hand-written prompts. We position our work within the 'computer-as-teammate' [12] relationship, where the responsibility for creating and modifying generative artefacts is shared between the designer and the AI, fostering a collaborative and cooperative approach. This can also be seen as a 'shift' for content generation on a mixed-initiative continuum [1]: allocating greater contribution and ownership to AI systems to achieve designer-focused results.

2.2 Mixed initiative and co-creative systems with game design applications

There are several MI-CC and HAI-CC systems that have been designed specifically for the purpose of creating game content. As an overview, [11] utilises a neuro-evolutionary approach to generate content that matches the user's aesthetic preferences, taking into account the user's feedback and evolving content accordingly. [5] introduces a mixed-initiative editor for the generation of sprite imagery which allows for user iterations and alterations, enabling the user to actively participate in the creation of sprite imagery. [1], [2] present a mixed-initiative co-creative tool for level design, where the user receives suggestions based on their creative decisions in the design process. Turning to text-based content and narrative, there have been significant recent advancements in the development of LLM-based character agents [14], [15], [16]. These agents utilize LLMs to create dynamic, context-aware conversational NPCs within games, generally based on persona creation via prompting. By leveraging these agents, designers and developers can avoid extensive hard-coding efforts and create NPCs that can adapt to the player's actions and enhance user interaction [6]. Additionally, there are systems that may not be explicitly designed for use in games but can still be applied in game design contexts. For instance, these systems can be used for planning narrative streams, generating scripts, and creating live text outputs for use-cases such as generative NPCs. [7] is a co-creative storytelling tool which guides the user through storytelling, providing storytelling goals and options for plot events. [4] is a mixed-initiative storytelling tool which supports a group of players to incrementally weave a new story through gameplay.

3 EARLY USER INSIGHTS

We have had the opportunity to gather insights from two different creative professionals in the entertainment industry. User A is in the games industry, in a creative direction role and has extensive experience working with AI-augmented workflows. User B is in the animation industry in a senior role working under directors, and has no previous experience using AI in their work. Both users have experience in creating reproductive content. We sort our findings into two primary points, with additional findings out of scope for this paper.

3.1 Challenges managing complex lore

User A highlighted that intellectual property (IP) holders often face challenges when it comes to keeping track of complex lore. They offered the insight that this is a recurring problem within the creative industries that demands significant creative effort, expressing interest in using the tool to identify and effectively manage overlap in lore and

storytelling outside of purely checking generative content. On a similar note, User B shared their personal experience as a creative who works for a variety of studios and IP holders. Often, User B explains, they are not allocated time to acclimate to or learn the lore of the stories which they are creating content for. User B suggested that having a lore-checking tool to check their non-generative work would also be helpful, avoiding incorrect submissions to directors and saving a significant amount time doing revisions.

3.2 Alignment of language and visualisation

User A noted the potential for discrepancies in our methodology between text and image, emphasising the presence of creative divergence from page to image - even if both text and image are considered valid, they may contrast. User A points out that an image-to-image approach would be best for them as someone who works primarily with visual direction. User B raised the point that there is a difference in creative freedom between concept artists and technical artists - "If the director asks for something, I need to do exactly that, or I'll get in trouble [...] but words have different meanings." User B, who speaks English, their working language, as a second language, felt that a tool like this could be useful for ensuring their non-generative work aligns with director expectations, as well as for working with clients who aren't always able to effectively express their vision to a team.

4 METHOD

This section outlines the methodology we employ to verify generative content against pre-existing creative lore. This system is designed to be embedded in a wider generative pipeline for co-creative work, and not immediately user facing.

4.1 Design document

We run initial tests using a design document outlining the lore of a small-scale fantasy world. This starting document is similar to a 'wiki' or lore encyclopaedia - serving as a comprehensive guide to the pre-existing creative lore with detailed descriptions of locations, character profiles, relationships, political systems, etc. To facilitate clear communication in the co-creative process, this document has additional features as follows:

- `Core Rules`: These establish the foundational principles which govern the world. For example, we may take for granted the understanding that in general, traditional fantasy worlds do not usually contain emails or sports cars. However, ground rules such as this should be outlined explicitly to avoid inconsistencies [18].
- `Keywords`: Descriptive keywords are used to convey the aesthetic nature of image generation subjects, such as locations. For example: forest, ancient, bioluminescent, tree house, mossy, nature, verdant.
- `Exclusions`: In contrast to the keywords, exclusions note what should never be included in a piece of content. Examples of exclusions could be species of animals, behaviours, or technology of a certain time period: for example, no computers in a high fantasy world.

4.2 Text

Testing directly on generative AI features which may feature in game design contexts, we test our text module by using actions output by large language model (LLM) based generative agents [14] as input. These agents are loaded with a character profile and subsequently generate dialogue and actions, acting as NPCs in a game-based scenario. First, we use system to attention prompting [17] to summarise agent outputs with clarity and brevity. We then run a prompt chain of including RAG queries, verifying lore accuracy from high level details (core rules), to low level (character

profile), to wide-view (broader world lore). For each query we output a binary score, with 0 being inaccurate and 1 being accurate, and a natural language explanation for the model decision. These binary scores can be used to discard or otherwise allocate inaccurate outputs.

4.3 2D Image

In our visual module, we leverage developments in LMMs [20] capable of processing images. Here, we analyse outputs from a text-to-image model by prompting our chosen LMM (GPT-4-Vision) to describe the input image in detail. We then run a prompt chain and RAG query, similar in structure to our approach for text, to check this description against location descriptions, keywords, and exclusions in our direction document. To compare results against human creative decision making, we initially selected examples which are most, or least, lore accurate based on descriptions in the design document. We generate two different types of image descriptions, including a list of most salient adjectives for the image (forest, mossy, mystical) and a detailed description of the image contents. We found that querying based on high-level adjective descriptions first catches any significantly divergent imagery - for example, if the sky in a landscape is the wrong colour. We then find that semi-divergent imagery - where for example, the buildings are the wrong shape - is best caught using a detailed description.

5 CONCLUSION AND FUTURE WORK

In light of the growing trend of reproductive storytelling across creative industries and the evolving role of human-AI co-creativity, we work towards integrating generative AI within established creative and entertainment industry workflows. This research has shown promise, with positive initial user feedback and initial success in checking text-based lore accuracy. Our limitations are the early development stage of this work, the scope of our experiments, and inherent biases of LLMs and LMMs which may impact marginalized users. This work can be developed by experimenting with narrative generation and by running extensive tests on a variety of complex creative worlds. Specific to visual content, there is significant opportunity for further development in our methodology by shifting to an image-to-image approach, rather than image-to-text, which could enable a more detailed analysis of style and detail. Overall, this system can be further developed into a co-creative suite for mixed-initiative art direction, suitable for application with pre-existing creative worlds and complex lore - for example, content creation for reproductive media. It offers the potential to operate within a comprehensive set of tools that designers and creatives can use to generate content and ensure it aligns with complex creative lore.

REFERENCES

[1] Alberto Alvarez, Steve Dahlskog, Jose Font, Johan Holmberg, Chelsi Nolasco, and Axel Österman. 2018. Fostering creativity in the mixed-initiative evolutionary dungeon designer. In *Proceedings of the 13th International Conference on the Foundations of Digital Games* (Malmö, Sweden) *(FDG '18)*. Association for Computing Machinery, New York, NY, USA, Article 50, 8 pages. https://doi.org/10.1145/3235765.3235815

[2] Alberto Alvarez, Steve Dahlskog, Jose Font, and Julian Togelius. 2019. Empowering Quality Diversity in Dungeon Design with Interactive Constrained MAP-Elites. In *2019 IEEE Conference on Games (CoG)*. 1–8. https://doi.org/10.1109/CIG.2019.8848022

[3] Logan Brown. 2024. Rethinking Remakes: Value and Culture in Video Game Temporalization. *Games and Culture* 19, 3 (2024), 337–356. https://doi.org/10.1177/15554120231163655 arXiv:https://doi.org/10.1177/15554120231163655

[4] Mijael Ricardo Bueno Perez and Rafael Bidarra. 2022. Mixed-initiative story co-creation with TaleMaker. In *Proceedings of the 17th International Conference on the Foundations of Digital Games* (Athens, Greece) *(FDG '22)*. Association for Computing Machinery, New York, NY, USA, Article 33, 13 pages. https://doi.org/10.1145/3555858.3555876

[5] Flávio Coutinho and Luiz Chaimowicz. 2024. Pixel art character generation as an image-to-image translation problem using GANs. *Graphical Models* 132 (2024), 101213. https://doi.org/10.1016/j.gmod.2024.101213

[6] Lajos Matyas Csepregi. 2021. The effect of context-aware llm-based npc conversations on player engagement in role-playing video games. *Unpublished manuscript* (2021).

[7] Max Kreminski, Melanie Dickinson, Noah Wardrip-Fruin, and Michael Mateas. 2022. Loose Ends: A Mixed-Initiative Creative Interface for Playful Storytelling. *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment* 18, 1 (Oct. 2022), 120–128. https://doi.org/10.1609/aiide.v18i1.21955

[8] Gorm Lai, Frederic Fol Leymarie, and William Latham. 2022. On Mixed-Initiative Content Creation for Video Games. *IEEE Transactions on Games* 14, 4 (2022), 543–557. https://doi.org/10.1109/TG.2022.3176215

[9] Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, Sebastian Riedel, and Douwe Kiela. 2020. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. In *Advances in Neural Information Processing Systems*, H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin (Eds.), Vol. 33. Curran Associates, Inc., 9459–9474. https://proceedings.neurips.cc/paper_files/paper/2020/file/6b493230205f780e1bc26945df7481e5-Paper.pdf

[10] Antonios Liapis, Gillian Smith, and Noor Shaker. 2016. *Mixed-initiative content creation.* Springer International Publishing, Cham, 195–214. https://doi.org/10.1007/978-3-319-42716-4_11

[11] Antonios Liapis, Georgios N. Yannakakis, and Julian Togelius. 2012. Adapting Models of Visual Aesthetics for Personalized Content Creation. *IEEE Transactions on Computational Intelligence and AI in Games* 4, 3 (2012), 213–228. https://doi.org/10.1109/TCIAIG.2012.2192438

[12] Zhiyu Lin and Mark Riedl. 2023. An Ontology of Co-Creative AI Systems. arXiv:2310.07472 [cs.AI]

[13] C. McGarry. 2023. The New Video Game Releases We Can't Wait to Play in 2023: Starfield, Baldur's Gate 3, and More. Wirecutter: Reviews for the Real World. Retrieved August 8, 2023, from https://www.nytimes.com/wirecutter/blog/new-video-game-releases-2023/.

[14] Joon Sung Park, Joseph C. O'Brien, Carrie J. Cai, Meredith Ringel Morris, Percy Liang, and Michael S. Bernstein. 2023. Generative Agents: Interactive Simulacra of Human Behavior. (2023). arXiv:2304.03442 [cs.HC]

[15] Yunfan Shao, Linyang Li, Junqi Dai, and Xipeng Qiu. 2023. Character-LLM: A Trainable Agent for Role-Playing. arXiv:2310.10158 [cs.CL]

[16] Alexander Sasha Vezhnevets, John P. Agapiou, Avia Aharon, Ron Ziv, Jayd Matyas, Edgar A. Duéñez-Guzmán, William A. Cunningham, Simon Osindero, Danny Karmon, and Joel Z. Leibo. 2023. Generative agent-based modeling with actions grounded in physical, social, or digital space using Concordia. arXiv:2312.03664 [cs.AI]

[17] Jason Weston and Sainbayar Sukhbaatar. 2023. System 2 Attention (is something you might need too). arXiv:2311.11829 [cs.CL]

[18] Oliver Wood, Rebecca Fiebrink, et al. 2023. Setting Switcher: Changing genre-settings in text-based game environments populated by generative agents. (2023).

[19] Georgios N. Yannakakis, Antonios Liapis, and Constantine Alexopoulos. 2014. Mixed-Initiative Co-Creativity. In *Proceedings of the 9th International Conference on the Foundations of Digital Games.* https://www.um.edu.mt/library/oar/handle/123456789/29459 Retrieved from https://www.um.edu.mt/library/oar/handle/123456789/29459.

[20] Shukang Yin, Chaoyou Fu, Sirui Zhao, Ke Li, Xing Sun, Tong Xu, and Enhong Chen. 2024. A Survey on Multimodal Large Language Models. arXiv:2306.13549 [cs.CV]