ArtAI4DS: AI Art and its Empowering Role in Digital Storytelling

Teresa Fernandes[1][0009−0005−9973−9343], Valentina Nisi[1,2][0000−0002−8051−3230], Nuno Nunes[1,2][0000−0002−2498−0643], and Stuart James[3][0000−0002−2649−2133]

[1] Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal
[2] ITI/LARSyS, Lisbon, Portugal
[3] Durham University, UK

Abstract. In an era of global interconnections, storytelling is a compelling medium for fostering understanding, building connections, and facilitating cultural exchange. Throughout history, visual imagery has been used to enrich narratives. However, this has been a privilege for those with artistic skills. Artificial Intelligence, specifically Generative AI, has the potential to democratize the process, allowing individuals to bring their narratives to life visually, regardless of their artistic prowess. To address this challenge, we developed an AI-powered tool called ArtAI4DS (Art AI for Digital Storytelling), that employs generative images (i.e., from Stable Diffusion) created from story-derived keywords. ArtAI4DS emerged from a research process starting with a 'Wizard of Oz' pre-workshop, which informed the structure of a subsequent co-design workshop. Here, participants' hand-drawn images were compared with AI-generated ones, providing insights into user preferences and tool efficacy. The ArtAI4DS then went through four iterative prototypes, drawing valuable insights from various participants. The tool's refinement process balanced the intricate duality of human creativity and technological innovation, culminating in an artistic expression platform that transforms stories into vivid and captivating images. The final tool, evaluated through user interviews and AttrakDiff questionnaire, showcases its potential as an engaging platform for transforming narratives with solid user affirmation of its motivational and emotional resonance.

Keywords: Digital Storytelling · Creativity Support Tools · Generative Art · Co-Design

1 Introduction

Storytelling, an intrinsic part of human nature, has been a bridge connecting diverse generations, cultures, and civilizations from ancient campfires to modern digital platforms [50]. Particularly for exchange students, storytelling chronicles their unique journeys through diverse cultural backgrounds, personal challenges, and transformative experiences supported in Europe by the Erasmus+ program. These narratives serve as personal tales and a testament to the rich cultural exchange that strengthens the European community's bonds. However, there's a limitation: finding the perfect imagery to complement these tales.

With the dawn of the digital age, storytelling has taken on new dimensions. Modern tools enable the creation of digital stories, combining traditional narratives with multimedia elements such as images, video clips, and audio [37, 30]. Artificial Intelligence (AI) is a promising solution to this visual conundrum. With the advent of image generation, technologies like Generative Adversarial Networks (GANs)[49] and Diffusion [11] can craft imagery that aligns seamlessly with narratives [24, 34]. The use of AI has the potential to create a richer, more vivid platform for exchange students to recount their adventures. Fusing compelling narratives and powerful AI-generated imagery holds untapped potential [44]. An example is an exchange student describing a day at a historic European monument; AI can generate detailed visuals of the location, adding depth and context to the tale. Beyond just generating images, AI offers a personal touch, tailoring each visual to the unique narrative of every student's story. Users gain inspiration by directly prompting Generative AI (GenAI); the AI generative process is iteratively refined by modifying the prompt to promote deeper exploration and expression of their story's imaginary and visual outputs[18].

The synergy between AI and storytelling remains largely unexplored. According to Kexue Fu et al., the challenges faced in the GenAI co-creation include diverging from the author's anticipated outcomes. However, this deviation from expectations can stimulate participants' creativity and expand their initial narratives. The influence of GenAI on the user experience of authors presents both benefits and drawbacks, highlighting the necessity for improved design strategies in GenAI authoring co-creation processes so to take advantage of inspiration while reducing inconsistencies and biases, particularly in contexts that demand creative storytelling with GenAI support [18]. Therefore, for ArtAI4DS, we research the challenge of combining AI Visuals with storytelling, aiming to amplify and enrich the tales of exchange students. Therefore, our contribution is a platform where stories come alive, enhanced by vivid, tailor-made visuals, thus redefining the storytelling experience. The platform is initially guided by a 'Wizard of Oz' experiment to guide the iterative development process, which is then evaluated using AttrakDiff.

2 Related work

We consider works related to Digital and Visual storytelling to provide a methodological background. We then discuss specifics associated with Creativity Support Tools and image-generation techniques.

Digital Storytelling Storytelling is a timeless human tradition. With the proliferation of technology and online platforms, narratives have transcended the confines of linear progression, giving birth to diverse forms of digital storytelling. Such approaches include different types of storytelling: linear, non-linear, adaptive, collaborative, geolocation and transmedia [8, 9, 40, 41, 13, 12].

Visual Storytelling: In digital storytelling, visual elements are pivotal in captivating audiences and conveying compelling narratives. Visual storytelling leverages a wide array of visual media to engage the senses and cater to the increasingly visual nature of contemporary audiences. This includes image-centric storytelling, which places images at the forefront and is frequently employed in photography exhibitions, graphic novels, and online image galleries [14]. Video narratives combine moving images, sound, and structured storytelling, making them key formats for short films, documentaries, and online content [25]. Data visualization storytelling distills complex information into digestible stories through interactive data visualizations and infographics [28]. Interactive visual stories blur the line between audience and author, allowing users to make choices, click on elements, and explore various story branches [15]. Immersive technologies like virtual reality (VR) and augmented reality (AR) narratives enable users to step into the narrative, interact with characters, and explore environments [5, 39]. Animation, whether in 2D or 3D motion graphics, adds life to stories and is versatile in representing abstract concepts, whimsical tales, and complex narratives [55]. The choice of medium depends on the story's purpose, target audience, and the desired emotional impact. In the digital age, visual storytelling continues to evolve, expanding the realm of narrative possibilities.

Creativity Support Tools (CSTs) with AI Creativity Support Tools (CSTs) have emerged to foster new forms of expression, aiding users across various media like graphic design, video editing, and data visualization [42, 17]. Recent breakthroughs have integrated Artificial Intelligence (AI) into CSTs, enhancing their capabilities and transforming them into collaborative co-creators [29, 20, 7]. Bala et al.'s work exemplifies this concept by presenting an AI-powered authoring tool tailored for Cultural Heritage storytelling [6]. Modern CSTs focus on specific user communities and integrate intricate interactions [27, 22].

Large Language Models (LLMs), for example, ChatGPT[4], have revolutionized the AI and CST landscape. Trained on extensive text data, these models produce human-like text, aiding artists in content creation and fostering broader engagement [3, 16]. Similarly, Image generation has witnessed rapid progress from early Variational Auto-Encoders (VAEs) to generate images conditioned by captions [36, 23]. Subsequent evolution brought in Generative Adversarial Networks (GANs) that improved image fidelity and catered to zero-shot prompts [49]. Subsequently, diffusion models introduced a probabilistic iterative refinement approach to image generation, resulting in images of exceptional quality and realism [54, 11] frequently relying on language models such as CLIP [45]. Well-known tools such as Dall-E 2 by OpenAI combine diffusion and CLIP techniques [46], which is similarly applied in Stable Diffusion, Dall-E 2, and Midjourney. Seltzer [53] comparison revealed distinct strengths and use cases for each model. While Midjourney operates through Discord, Stable Diffusion offers both online and offline usage, and DALL-E 2 provides a web-based interface with limited free access [1]. For ArtAI4DS AI-assisted illustrations, Stable Diffusion stands out as the prime choice due to its blend of versatility and quality.

Image generation has evolved exponentially, with models like Stable Diffusion showcasing the potential of marrying art with technology.

The conceptual potential of image generation [47, 45] has also been explored to relate to emotion [31]. However, CSTs, while enriching text with AI advancements and reshaping the boundaries of creative expression, need to incorporate significant strides for visual elements that we address.

Story Illustration using AI: The visualization of stories presents a unique challenge. While narratives teem with emotions and contexts, their transformation into imagery is intricate. Several works have ventured into story visualization. For instance, research has explored storyboarding based on narrative sentences [10], co-creative systems for children's visual storytelling [57], and AI algorithms to modify prompts for story illustration [33]. Meta's "Make-A-Scene" showcased the potential of controlling image elements to illustrate stories [19]. A myriad of other works, such as "Show me a story" [48] and Yturrizaga-Aguirre's proposal [56], have also made strides in the domain.

Recent advancements have gone beyond individual illustrations, aiming to generate narratives with text and images. For instance, the Long Stable Diffusion model can create a whole book, although direct text conditioning might lead to irrelevant images. Similarly, StoryGAN [32] and StoryDALL-E [35, 43] employ their respective underlying models for generating a visual sequence corresponding to textual story descriptions. Alternatively, StoryBook [2] introduces a zero-shot storybook creation, while TaleCrafter [21] enables interactive story visualization. A challenge to approaches is often character consistency, which DreamBooth [51] addressed. Focusing on specific demographics, the AIStory [26] prototype offers interactive storytelling experiences for children. Similarly, Ruskov et al. [52] introduced an approach that attempts to illustrate fairy tales using prompt engineering, detailing a four-stage process and exploring challenges faced by generation models. Despite the advancements, all approaches take a data-driven approach to storytelling. Therefore, there remains a domain gap in achieving design-centric narrative translations. As the field evolves, a deeper understanding and narrowing of the gap is pivotal, which we begin to address.

3 ArtAI4DS: Art AI for Assistive Digital Storytelling

To create ArtAI4DS, we initially designed a Wizard-of-Oz style study to understand how participants could use AI-generated imagery in their story illustration process. Therefore, the participants were asked to write a story, illustrate it, and then reflect on images generated by the expert, i.e. the wizard (Sec. 3.1). We then used the stories to guide selection of one of six State-of-the-Art automatic keyword selection methods through an online survey. We then use the study outcomes to inform the initial prototype for an iterative co-design process to refine and develop the ArtAI4DS tool (Sec. 3.2). The process of development is conceptualized in Figure 1. Finally, we outline the details of the final version of the tool, including specific implementation details in Section 3.5. For all studies, we recruit students from Instituto Superior Técnico (IST) of the University of

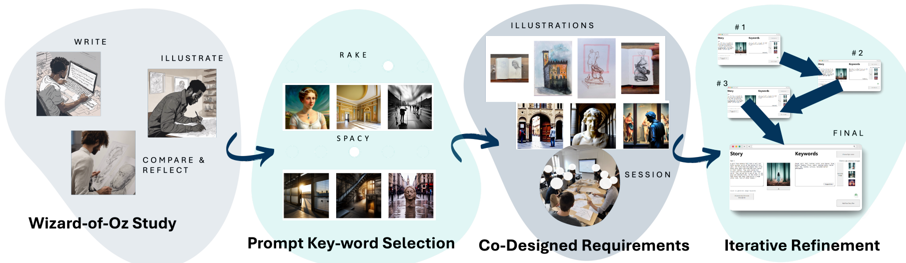


Fig. 1: ArtAI4DS Methodology: We start from a wizard-of-oz study to conceptualize the tool, then identify the prompt generation method, then expand these requirements through a co-design group session and finalize the prototype tool through an iterative refinement process.

Lisbon and, where possible, Erasmus+ students. We provide the nationality of students throughout the sections to provide context.

3.1 How people tell stories with accompanying art: A Wizard of Oz Study

Methodology: To ground the study, we applied a Wizard-of-Oz style study, in which we created an environment where participants believed they were engaging with a Generative Art AI; however, behind the scenes, human intervention shaped the experience. The study comprised four stages: *i)* narrative crafting and writing, *ii)* manual illustration, *iii)* AI-driven image generation, and *iv)* reflective comparison and analysis. Our AI-driven image generation exploratory path is bifurcated into two strategies: the focused Keyword-driven Illustration and the Complete Phrase Inputs approach. To delve deeper into our participants' experiences and perceptions, we employed a set of guiding questions, ensuring our dialogues remained insightful and purposeful. The questions were as follows:

– Do these generated images and your illustrations represent the story well?
– Is there something missing from the generated images or something you would change?
– What are the pros and cons of your sketches vs. the generated images?
– Would generating these images change something in the way you told your story?

Study Participants: This study consisted of three participants: P1, an Italian Agronomy student with insights into regional diversity; P2, a Brazilian and Italian Philosophy student contributing philosophical insights; and P3, a Greek and Austrian Fine Arts graduate currently residing in Florence, who added a creative dimension to the project.
Study Insights: Based on semi-structured interviews, several key themes emerged. P1 stressed the importance of accuracy in depicting regional details and valued

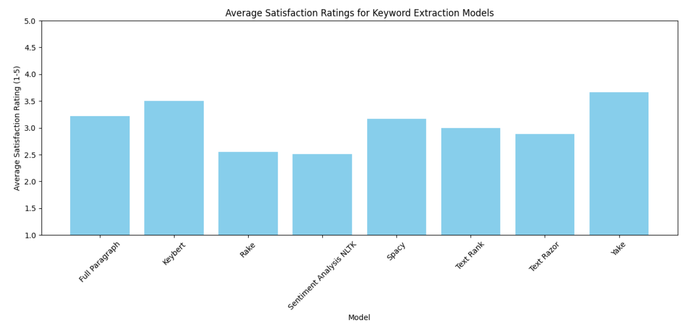
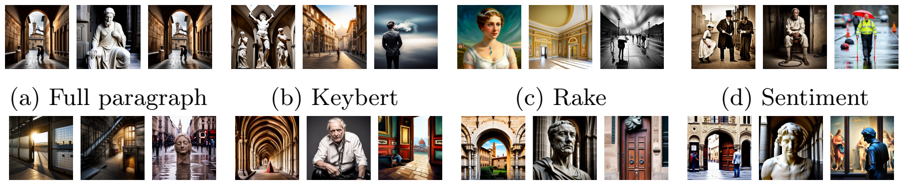


Fig. 2: Average Keyword Extraction Model Satisfaction Ratings

user-generated content for its emotional connection. P2 emphasized the need for accuracy and emotional depth in images, while P3, with a fine arts background, highlighted the gap between AI-generated images and the intended emotional depth, emphasizing the importance of artistic interpretation and symbolism. These insights have been instrumental in refining our approach and aligning it more closely with user expectations and the nuances of storytelling. An observation from the expert, which we also took to the co-design workshop, was the need for a repeatable method of creating prompts from users' stories to generate images. It was identified that using only the paragraph of text generated highly unrelated imagery, and therefore, the expert needed to derive a prompt (i.e., keywords) from the text to get an appropriate image.

3.2 Prompt Generation

To address the issues with selecting keywords for the prompt, we employ a survey of different techniques for extracting keywords from a given story paragraph. Keyword extraction in our setting aims to capture the essence of a narrative.
Keyword Extraction Methods Survey: We employed a variety of approaches to understand the prompts needed. These ranged from straightforward techniques like inputting the full paragraph to more nuanced models like Keybert, which leverages the BERT architecture. We also explored Rake, an unsupervised keyword extraction model; Sentiment Analysis from NLTK, which infuses emotional resonance into keyword extraction; Spacy, a robust natural language processing library; Text Rank, an algorithmic approach inspired by Google's PageRank; and Text Razor, another unsupervised model; and Yake, likewise an unsupervised method. We conducted illustrative experiments using the stories from our previous Wizard-of-Oz study. These stories underwent keyword extraction via all the different models and the generation of the corresponding image. Participants were then asked to select the preferred image based on its relevance to the text.
Study Participants: University of Lisbon students evaluated the images they produced. Five were selected, four of whom had an electrical engineering back-



(a) Full paragraph (b) Keybert (c) Rake (d) Sentiment

(e) Spacy (f) Text Rank (g) Text Razor (h) Yake

Fig. 3: P3's story illustrated with various techniques.

ground, and one had an industrial and management engineering background. The age range was 22/23 years old, including one female and four males, four Portuguese and one Mozambican. In this study, we did not prioritize selecting Erasmus+ students as the study evaluated methods, not methodology.
Analysis: The feedback collected was distilled into average satisfaction ratings (Figure 2). From Figure 2, while there is a similar performance between methods, it can be seen the superior performance of the Yake model is followed closely by Keybert. To contextualize these results, we also show examples for a given output of the different models in fig. 3. Guided by these insights, we adopted the Yake model for the co-Design Workshop's keyword extraction process.

3.3 Co-designed Initial Requirements

The group workshop was an arena where collaboration and innovation thrived. At the center of our methodology was extracting the rich tapestry of Erasmus's stories from our participants. This narrative foundation was then used to compare between human-crafted illustrations and their AI-generated counterparts, similar to the Wizard-of-Oz study.
Co-Design Group Workshop Methodology: We generated the images and showed them to the participants, allowing to draw comparisons between them. Then, there was a group discussions about their illustrations and the AIs. The difference between this workshop and the pre-workshop was that it was conducted in a group setting; we used the selected Yake model to generate images.
Co-Design Group Workshop Participants: The co-design workshop participants were four current Erasmus students and two who had previously done Erasmus. The ages ranged from 21 to 25, with 3 Italians, 2 Portuguese, and one French. Of these, 3 were females and 3 were males.
Co-Design Group Workshop Analysis The co-design workshop facilitated dynamic group discussions, unveiling critical insights into the interaction between human creativity and AI capabilities in image generation for storytelling. Key themes emerged, such as i) the significance of specificity in storytelling, ii) the subjectivity of interpretations, and iii) the importance of embedding emotions in AI visuals. Participants also recognized iv) the value of manual keyword selection and v) the diversity of artistic preferences among users. Ethical considerations regarding AI, especially in decision-making processes, were highlighted, with a

call for vi) transparency and equity. The workshop's insights form a crucial foundation for developing a prototype that will reflect these findings and enhance the user experience of the AI Storytelling tool. For example, to ensure the manual keyword selection, we let users modify the prompts freely, and considering the diversity of artistic preferences and subjectivity of interpretations, we offer style selection and creative suggestions to modify users's prompts.

3.4 Iterative Prototyping

From the co-designed requirements (Sec. 3.3), we develop the initial prototype and then go through iterative developments with a small group of feedback at each step, adding additional participants to increase the diversity of feedback.
First prototype iteration: The first iteration aimed to create a functional prototype and then user insights. It featured a simplified user interface and basic image generation capabilities, focusing on user-friendly input and output. Our prototype had to be functional so that users could interact with the AI technology. Two Erasmus students, P1 and P2, were selected for formative evaluation. Their feedback was vital to ensuring the tool's alignment with the Erasmus student community's needs and expectations.
Analysis & Improvements: The results from the first formative evaluation were as follows: P1 expressed concerns about aesthetics, style consistency, and recognizing specific keywords, while P2 found the tool intriguing but suggested user-friendliness improvements such as visual feedback and a download feature. Both participants emphasized the tool's potential. Based on user feedback, several refinements were made to enhance usability and functionality, including i) optimizing button layout, ii) providing visual feedback with a spinner, iii) dynamic prompt generation, iv) historical insights, v) download capability, vi) suggestions to modify prompts, vii) improved visual clarity, viii) interface streamlining, ix) style selection, x) improved clarity with placeholder text, and xi) code refinements by fixing bugs.
Second prototype iteration: The second iteration of the storytelling illustration tool represented a significant evolution from the prototype. It incorporated all the refinements mentioned above, resulting in a more polished and user-friendly interface. In this phase, two participants, P3 and P4, provided feedback. P3, an undergraduate student without Erasmus experience, offered a unique perspective, while P4, with an artistic and design background, contributed insights regarding creative expression and design sensibilities.
Analysis & Improvements: In this evaluation phase, P3 highlighted issues with the segmented approach to storytelling, emphasizing the need for a more accurate relationship between text and visuals. P3 also advocated for clearer image keywords. P4 praised the tool's speed and originality but suggested improvements in user guidance, aesthetics, and button design standardization. Both participants appreciated the style feature. In response to the feedback, refinements were made to enhance user understanding and clarity: i) improve button text descriptions, ii) ensure consistency in button sizing, iii) introduce user guidance, and iv) refine the prompt generation algorithm. These changes aimed to streamline the

storytelling process, enhance user experience, and maintain a visually balanced and harmonious design.

Third prototype iteration: The third prototype iteration aimed to build upon the second prototype's improvements, introducing all the aforementioned refinements. A group of four participants, including P5 with an artistic background, P6 from a design research and technology integration background, P7 with Erasmus experience, and P8 without Erasmus experience, provided valuable feedback.

Analysis & Improvements: In this evaluation stage, P5 suggested improving the user interface, P6 found the tool helpful in educational contexts and suggested UI simplifications, P7 favored certain image styles and pointed out website formatting issues, and P8 praised the tool's intuitiveness and suggested UI enhancements. In response to the feedback, several refinements were made, including the i) introduction of error logging, a ii) restructured modal for past images, iii) repositioned functionality, iv) streamlined story creation, and v) uniform fonts throughout the tool. These changes enhanced the tool's performance, user experience, and visual consistency.

3.5 Final Prototype and Architecture

The tool includes several sections, such as the "Story Section" on the left, where users can write narratives and use the "Automatically Generate Description" option with the Yake extraction model to extract keywords to the *Keywords section*. The *Keywords section* displays extracted words from the story and enables user modifications while providing thematic suggestions. Users can expand their story using the *Story Progression* feature and see a visual representation of the current story part in the "Central Image Display". The "Previous Images" section on the right lets users to view and gain inspiration from previously generated images and associated stories. The "Choose Style" button offers an option for customizing the artistic style of the images. The low fidelity tool can be seen in Figure 4. Throughout the development phase, we focus only on functional design and leave a high-fidelity tool for future work.

Technical Details For image generation, we use the Stability AI API with model version *stable-diffusion-xl-beta-v2-2-2* running within a Python Flask API endpoint with a static site using javascript to query the API. The code is available at https://github.com/teresaacsf/ai-illustrator/ for future comparability.

4 Evaluation

With the final prototype of ArtAI4DS we evaluate the experience of using the combined writing and illustration process. To do this we further recruited a set of participants who completed the AttrakDif survey.

4.1 Participants

We conducted a final study with seven participants, four new participants, and three prior participants, providing continuity to the process. The new partic-

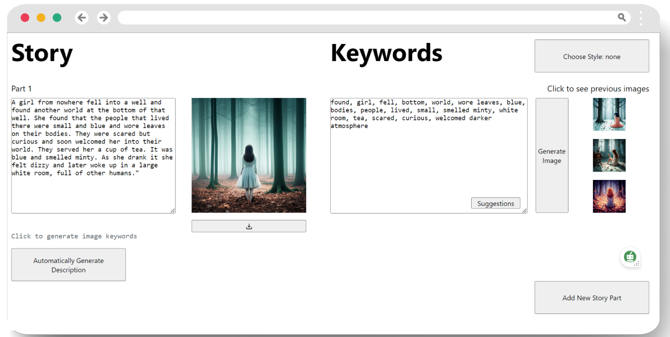


Fig. 4: The ArtAI4DS final low fidelity prototype derived from requirements and iterative prototyping: i) left part-based narrative, ii) middle the corresponding keyword extraction, iii) right the generated images.

ipants included participant P9, a student in management engineering without Erasmus experience, who contributed fresh insights. P10 and P11, who had Erasmus experiences, enriched the tool's development with a focus on accommodating users engaged in international mobility. P12, an Italian student in space engineering with prior Erasmus experience, provided valuable feedback. Their diverse experiences and perspectives ensured that the final version of the tool would effectively cater to a wide range of users. We also interviewed participants P2, P4, and P6 from early studies who were known to give rich feedback information. To help understand where participants took part across the tool development, a full table of participants can be seen in Table 1.

4.2 Analysis & Discussion

We review the feedback from participants using both qualitative and quantitative approaches. For qualitative feedback, P9 highlighted the need for more diverse keywords while appreciating the *Suggestions* button. P10 found the tool more original than traditional image searches but noted occasional mismatches with

Participant participated in	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12
Wizard of Oz	✓	✓	✓									
1st Iteration	✓	✓										
2nd Iteration			✓	✓								
3rd Iteration					✓	✓	✓	✓				
Evaluation		✓		✓		✓			✓	✓	✓	✓

Table 1: Summary of the studies the participants took part in from Wizard of Oz through iterative development to evaluation.

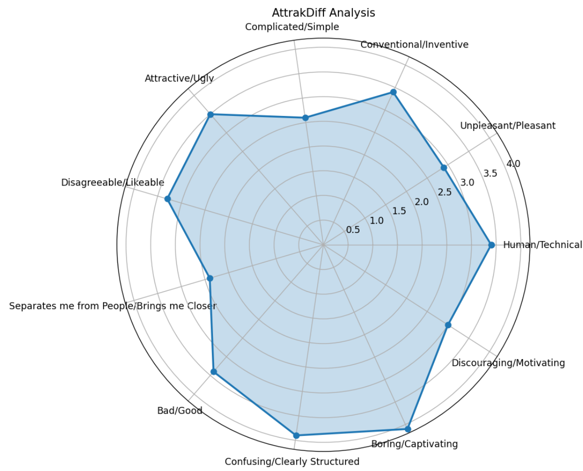


Fig. 5: Attrakdiff results of the ArtAI4DS AI Storytelling tool

image expectations and issues with realism. P11 praised the tool's ability to change keywords for diverse image options and its usefulness in quick visual storytelling. P6 and P4 emphasized the tool's utility in illustrating concepts and suggested features for changing past images. P4 found the final iteration more intuitive, highlighting its storytelling aspect, keyword, and style customization. P12 focused on usability improvements, suggesting clearer initial instructions.

Specifically, on reflection on the tool, the participants achieved the objective of creating more appropriate imagery P10 specifically noted, "It's more original than Google images.". However, P10 also commented, "It's not as realistic as Google' which would be interesting to explore and understand if this is an issue of the artistic nature of Image Generation methods or if the images generated could be tuned simply by greater global control (i.e. Realistic / Artistic option). An interesting extension to this work could also be in collaborative storytelling as noted by P6 "collaboratively, like in a classroom, where each student can write a part and modify images" allowing the exploration of stories. This could also address the personal nature of imagery [6] when describing autobiographical stories, and subsequent imagery does not reflect the author's memory. However, we note that ArtAI4DS is the first prototype in the strand of research, and as commented by P12 "The tool should be good enough to get the meaning of the story without the need for the user to modify something", which would be the goal for future research further integrating Language Processing techniques.

For quantitative evaluation of the ArtAI4DS prototype tool, we employed the AttrakDiff questionnaire to gauge participants' subjective experiences. The questionnaire assessed the hedonic and pragmatic qualities of the user experience. Participants provided ratings on scales, which included pragmatic quality (PQ), hedonic quality-identification (HQ-I), hedonic quality-stimulation (HQ-S), and overall attractiveness (ATT). In terms of strengths, the tool demonstrated a balance between human-centric and technical aspects (3.0), promoted social interaction (2.6), had the potential to motivate users (2.0), and scored well in overall attractiveness (3.5). It was highly captivating (4.1) and received positive feedback for being good (3.4) and likable (3.3). Areas for improvement include simplifying aspects for broader accessibility and ensuring the tool is not overly complicated (2.4). While visually appealing, addressing specific aspects could further enhance its attractiveness (3.5). The results are displayed in Figure 5.

The final prototype of the AI Storytelling tool achieved a balanced blend of creativity and technology, providing an approachable and user-friendly experience. It encouraged social interaction and motivated users to engage in creative storytelling, eliciting a positive emotional connection. Users found it attractive, likable, and captivating. However, user feedback identified potential areas for improvement, such as initial user guidance, keyword suggestions, and user interface. Addressing these recommendations would enhance the tool's usability and overall value. Unlike prior Creativity Support Tools, which focus on integrating existing imagery [6] or multimedia [42, 17], the tool through the provisional survey indicates the positive impact it can have on story authoring.

5 Conclusions and Future Work

Here, we presented the ArtAI4DS, the first Creativity Support Tool for Illustrated Digital Storytelling, which employed a user-centric co-design process. While the low-fi prototype demonstrates the challenges, especially concerning the gap between story writing and prompts for AI generative art, it captivated the participants with a new approach to story illustration. Drawing from our findings and related work in GenAI and storytelling, we highlight areas for improvement, particularly in the iteration process of generating images with Generative AI. However, users often lament inconsistencies in results produced by the AI, and reverting to a previous step for modifications is challenging. Kexue Fu et al. [18] proposed in-image editing, which can incorporate satisfactory elements or refinement of areas via prompt image parts. Another possibility for future work is using Blockchain to track of changes and revert to an exact satisfying process moment. By placing GenAI results on-chain, we can maintain an immutable record of each creative step, allowing authors to revert to any previous point in the creative process [38]. This approach not only enhances the reliability and satisfaction of the workflow but also ensures transparency and traceability. Future work will explore advanced blockchain features to support and streamline the creative process further and expand the application of this methodology to other domains where iterative creative processes are critical.

Acknowledgments. This study was funded by LARSyS (Project UIDB/50009/2020. In addition, this work benefited from the BIG ERAChair EU-funded project through Grant agreement ID: 952226 and the Blockchain.PT PRR Agenda PC644918095-00000033.

References

1. Dall-e 2. https://labs.openai.com/
2. et al., H.J.: Zero-shot generation of coherent storybook from plain text story using diffusion models (2023), https://arxiv.org/abs/2302.03900
3. et al., L.: Opening up chatgpt: Tracking openness, transparency, and accountability in instruction-tuned text generators (2023), https://arxiv.org/abs/2307.05532
4. et al., Y.L.: Summary of chatgpt-related research and perspective towards the future of large language models (2023), https://arxiv.org/abs/2304.01852
5. Antony, V.N., Huang, C.M.: Id.8: Co-creating visual stories with generative ai (2023)
6. Bala, P., James, S., Del Bue, A., Nisi, V.: Writing with (digital) scissors: Designing text editing tool for assisted storytelling using crowd-generated content. In: Interactive Storytelling: 15th International Conference on Interactive Digital Storytelling, ICIDS 2022, Santa Cruz, CA, USA, December 4–7, 2022, Proceedings. p. 139–158. Springer-Verlag, Berlin, Heidelberg (2022). https://doi.org/10.1007/978-3-031-22298-6_9
7. Brown, T., et al.: Language models are few-shot learners (2020), https://arxiv.org/abs/2005.14165
8. Casillo, M., Colace, F., De Santo, M., Lemma, S., Lombardi, M., Pietrosanto, A.: An ontological approach to digital storytelling. In: Proceedings of the The 3rd Multidisciplinary International Social Networks Conference on SocialInformatics 2016, Data Science 2016. MISNC, SI, DS 2016, Association for Computing Machinery, New York, NY, USA (2016). https://doi.org/10.1145/2955129.2955147, https://doi.org/10.1145/2955129.2955147
9. Cesário, V., Acedo, A., Nunes, N., Nisi, V.: Promoting social inclusion around cultural heritage through collaborative digital storytelling. In: International Conference on ArtsIT, Interactivity and Game Creation. pp. 248–260. Springer (2021)
10. Chen, S., Liu, B., Fu, J., Song, R., Jin, Q., Lin, P., Qi, X., Wang, C., Zhou, J.: Neural storyboard artist: Visualizing stories with coherent image sequences. In: Proceedings of the 27th ACM International Conference on Multimedia. p. 2236–2244. MM '19, Association for Computing Machinery, New York, NY, USA (2019). https://doi.org/10.1145/3343031.3350571, https://doi.org/10.1145/3343031.3350571
11. Dhariwal, P., Nichol, A.: Diffusion models beat gans on image synthesis (2021), https://arxiv.org/abs/2105.05233
12. Dionisio, M., Nisi, V.: Leveraging transmedia storytelling to engage tourists in the understanding of the destination's local heritage. Multimedia Tools and Applications **80**(26–27), 34813–34841 (Jun 2021). https://doi.org/10.1007/s11042-021-10949-2, http://dx.doi.org/10.1007/s11042-021-10949-2
13. Dionisio, M., Nisi, V., Nunes, N., Bala, P.: Transmedia Storytelling for Exposing Natural Capital and Promoting Ecotourism, p. 351–362. Springer International Publishing (2016). https://doi.org/10.1007/978-3-319-48279-8_31, http://dx.doi.org/10.1007/978-3-319-48279-$8_3$1
14. El-Desouky, D.F.: Visual storytelling in advertising: A study of visual storytelling as a marketing approach for creating effective ads. International Journal of Humanities, Social Sciences and Education (2020), https://api.semanticscholar.org/CorpusID:243616056
15. Ferreira, M., Nunes, N.J., Nisi, V.: Interacting with climate change: A survey of hci and design projects and their use of transmedia story-

telling. In: International Conference on Interactive Digital Storytelling (2021), https://api.semanticscholar.org/CorpusID:244884482
16. Fraiwan: A review of chatgpt applications in education, marketing, software engineering, and healthcare: Benefits, drawbacks, and research directions (2023), https://arxiv.org/abs/2305.00237
17. Frich, J., MacDonald Vermeulen, L., Remy, C., Biskjaer, M.M., Dalsgaard, P.: Mapping the landscape of creativity support tools in hci. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. p. 1–18. CHI '19, Association for Computing Machinery, New York, NY, USA (2019). https://doi.org/10.1145/3290605.3300619, https://doi.org/10.1145/3290605.3300619
18. Fu, K., Wu, R., Tang, Y., Chen, Y., Liu, B., LC, R.: Being eroded, piece by piece": Enhancing engagement and story- telling in cultural heritage dissemination by exhibiting genai co-creation artifacts. In: Proceedings of DIS 2024 (2024). https://doi.org/https://doi.org/10.1145/3643834.3660711
19. Gafni, O., Polyak, A., Ashual, O., Sheynin, S., Parikh, D., Taigman, Y.: Make-a-scene: Scene-based text-to-image generation with human priors (2022), https://arxiv.org/abs/2203.13131
20. Gero, K.I., Chilton, L.B.: Metaphoria: An algorithmic companion for metaphor creation. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. p. 1–12. CHI '19, Association for Computing Machinery, New York, NY, USA (2019). https://doi.org/10.1145/3290605.3300526, https://doi.org/10.1145/3290605.3300526
21. Gong, Y., Pang, Y., Cun, X., Xia, M., Chen, H., Wang, L., Zhang, Y., Wang, X., Shan, Y., Yang, Y.: Talecrafter: Interactive story visualization with multiple characters (2023), https://arxiv.org/abs/2305.18247
22. Green, D., Hargood, C., Charles, F.: Contemporary issues in interactive storytelling authoring systems. In: Interactive Storytelling: 11th International Conference on Interactive Digital Storytelling, ICIDS 2018, Dublin, Ireland, December 5–8, 2018, Proceedings 11. pp. 501–513. Springer (2018)
23. Gregor, K., Danihelka, I., Graves, A., Rezende, D., Wierstra, D.: Draw: A recurrent neural network for image generation. In: International conference on machine learning. pp. 1462–1471. PMLR (2015)
24. Gu, R., Li, H., Su, C., Wu, W.: Innovative digital storytelling with aigc: Exploration and discussion of recent advances (2023)
25. Hagedoorn, B.: Doing history, creating memory : Representing the past in documentary and archive-based television programmes within a multi-platform landscape (2016), https://api.semanticscholar.org/CorpusID:132273590
26. Han, A., Cai, Z.: Design implications of generative ai systems for visual storytelling for young learners. In: Proceedings of the 22nd Annual ACM Interaction Design and Children Conference. pp. 470–474 (2023)
27. Han, H.L., Yu, J., Bournet, R., Ciorascu, A., Mackay, W.E., Beaudouin-Lafon, M.: Passages: Interacting with text across documents. In: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems. CHI '22, Association for Computing Machinery, New York, NY, USA (2022). https://doi.org/10.1145/3491102.3502052, https://doi.org/10.1145/3491102.3502052
28. Harder, C.: Master's project: Dashboard 2.0: A visual storytelling mechanism to inspire relationship building, participation, & collaboration for storytelling (2018), https://api.semanticscholar.org/CorpusID:217412701

29. Hwang, A.H.C.: Too late to be creative? ai-empowered tools in creative processes. In: Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems. CHI EA '22, Association for Computing Machinery, New York, NY, USA (2022). https://doi.org/10.1145/3491101.3503549, https://doi.org/10.1145/3491101.3503549
30. Lathem, S.: Learning communities and digital storytelling: New media for ancient tradition (2005), https://www.learntechlib.org/p/73892/
31. Lee, Y.K., Park, Y.H., Hahn, S.: A portrait of emotion: Empowering self-expression through ai-generated art (2023), https://arxiv.org/abs/2304.13324
32. Li, Y., Gan, Z., Shen, Y., Liu, J., Cheng, Y., Wu, Y., Carin, L., Carlson, D., Gao, J.: Storygan: A sequential conditional gan for story visualization. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 6329–6338 (2019)
33. Ma, Y., Yang, H., Liu, B., Fu, J., Liu, J.: Ai illustrator: Translating raw descriptions into images by prompt-based cross-modal generation. In: Proceedings of the 30th ACM International Conference on Multimedia. pp. 4282–4290 (2022)
34. Maerten, A.S., Soydaner, D.: From paintbrush to pixel: A review of deep neural networks in ai-generated art (2023)
35. Maharana, A., Hannan, D., Bansal, M.: Storydall-e: Adapting pretrained text-to-image transformers for story continuation. In: European Conference on Computer Vision. pp. 70–87. Springer (2022)
36. Mansimov, E., Parisotto, E., Ba, J.L., Salakhutdinov, R.: Generating images from captions with attention (2015), https://arxiv.org/abs/1511.02793
37. Moutafidou, A., Bratitsis, T.: Digital storytelling: Giving voice to socially excluded people in various contexts. In: Proceedings of the 8th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-Exclusion. p. 219–226. DSAI '18, Association for Computing Machinery, New York, NY, USA (2018). https://doi.org/10.1145/3218585.3218684, https://doi.org/10.1145/3218585.3218684
38. Nguyen, C.T., Liu, Y., Du, H., Hoang, D.T., Niyato, D., Nguyen, D.N., Mao, S.: Generative ai-enabled blockchain networks: Fundamentals, applications, and case study. IEEE Network pp. 1–1 (2024). https://doi.org/10.1109/MNET.2024.3412161
39. Nisi, V., James, S., Bala, P., Del Bue, A., Nunes, N.J.: Inclusive digital storytelling: Artificial intelligence and augmented reality to re-centre stories from the margins. In: International Conference on Interactive Digital Storytelling. pp. 117–137. Springer (2023)
40. Nisi, V., Oakley, I., Haahr, M.: Inner city locative media: design and experience of a location-aware mobile narrative for the dublin liberties neighborhood. In: Intelligent Agent. vol. 6 (2006)
41. Nisi, V., Oakley, I., Haahr, M.: Location-aware multimedia stories: turning spaces into places. Universidade Cátolica Portuguesa pp. 72–93 (2008)
42. Palani, S., Ledo, D., Fitzmaurice, G., Anderson, F.: "i don't want to feel like i'm working in a 1960s factory": The practitioner perspective on creativity support tool adoption. In: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems. CHI '22, Association for Computing Machinery, New York, NY, USA (2022). https://doi.org/10.1145/3491102.3501933, https://doi.org/10.1145/3491102.3501933
43. Pan, X., Qin, P., Li, Y., Xue, H., Chen, W.: Synthesizing coherent story with auto-regressive latent diffusion models (2022), https://arxiv.org/abs/2211.10950

44. Pearson, E.: Making a good (virtual) first impression: The use of visuals in online impression management and creating identity performances. In: IFIP International Conference on Human Choice and Computers. pp. 118–130. Springer (2010)
45. Radford, A., Kim, J.W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G., Sutskever, I.: Learning transferable visual models from natural language supervision. arXiv preprint arXiv:2103.00020 (2021), https://arxiv.org/abs/2103.00020
46. Ramesh, A., Dhariwal, P., Nichol, A., Chu, C., Chen, M.: Hierarchical text-conditional image generation with clip latents (2022), https://arxiv.org/abs/2204.06125
47. Ramesh, A., Pavlov, M., Goh, G., Gray, S., Voss, C., Radford, A., Chen, M., Sutskever, I.: Zero-shot text-to-image generation. arXiv preprint arXiv:2102.12092 (2021), https://arxiv.org/abs/2102.12092
48. Ravi, H., Wang, L., Muniz, C., Sigal, L., Metaxas, D., Kapadia, M.: Show me a story: Towards coherent neural story illustration. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 7613–7621 (2018)
49. Reed, S., Akata, Z., Yan, X., Logeswaran, L., Schiele, B., Lee, H.: Generative adversarial text to image synthesis. In: International conference on machine learning. pp. 1060–1069. PMLR (2016)
50. Robin, B.: Digital storytelling: A powerful technology tool for the 21st century classroom (2008), https://digitalstorytellingclass.pbworks.com/f/Digital+Storytelling+A+Powerful.pdf
51. Ruiz, N., Li, Y., Jampani, V., Pritch, Y., Rubinstein, M., Aberman, K.: Dreambooth: Fine tuning text-to-image diffusion models for subject-driven generation (2022), https://arxiv.org/abs/2208.12242
52. Ruskov, M.: Grimm in wonderland: Prompt engineering with midjourney to illustrate fairytales (2023), https://arxiv.org/abs/2302.08961
53. Seltzer, F.: Comparison of ai art generators (2022), tweet
54. Sohl-Dickstein, J., Weiss, E.A., Maheswaranathan, N., Ganguli, S.: Deep unsupervised learning using nonequilibrium thermodynamics (2015), https://arxiv.org/abs/1503.03585
55. Wu, Z., Zhou, X.: The application of visual image and interactive storytelling to stage performance. In: Interacción (2019), https://api.semanticscholar.org/CorpusID:196610956
56. Yturrizaga-Aguirre, A., Silva-Olivares, C., Ugarte, W.: Story visualization using image-text matching architecture for digital storytelling. In: 2022 IEEE Engineering International Research Conference (EIRCON). pp. 1–4. IEEE (2022)
57. Zhang, C., Yao, C., Wu, J., Lin, W., Liu, L., Yan, G., Ying, F.: Storydrawer: A child–ai collaborative drawing system to support children's creative visual storytelling. In: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems. CHI '22, Association for Computing Machinery, New York, NY, USA (2022). https://doi.org/10.1145/3491102.3501914, https://doi.org/10.1145/3491102.3501914