Orchid: A Creative Approach for Authoring LLM-Driven Interactive Narratives

ANONYMOUS AUTHOR(S)

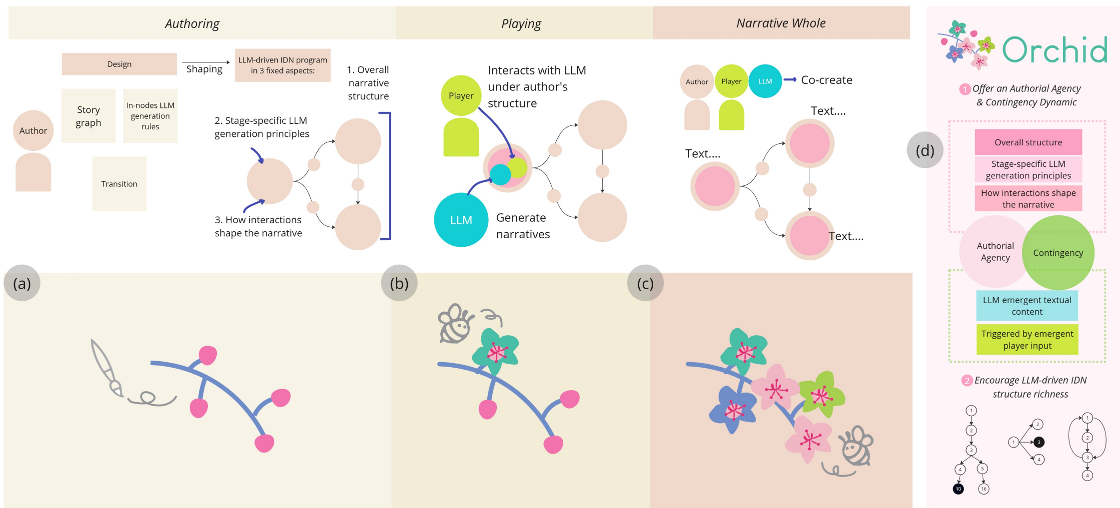

Fig. 1. The creative approach, Orchid, introducing human-AI co-creation in three stages. LLMs are represented as different buds, with player's prompts serving as the catalyst that triggers narratives to bloom from within. (a): Authoring. The author creates the structure of the generative IDN, especially provides narrative design from five perspectives. (b): The players interact with different nodes, yet the narrative is generated live through their interaction. (c): As players continually interact with the nodes, the narrative is gradually cultivated, and presented to the player as a whole. (d) demonstrates the two goals of Orchid: to provide a new dynamic between human authorial agency and machine contingency, and to encourage the creation of richer LLM-driven IDN structures.

Integrating Large Language Models (LLMs) into Interactive Digital Narratives (IDNs) enables dynamic storytelling where user interactions shape the narrative in real time, challenging traditional authoring methods. This paper presents the design study of Orchid, a creative approach for authoring LLM-driven IDNs. Orchid allows users to structure the hierarchy of narrative stages and define the rules governing LLM narrative generation and transitions between stages. The development of Orchid consists of three phases. 1) Formulating Orchid through desk research on existing IDN practices. 2) Implementation of a technology probe based on Orchid. 3) Evaluating how IDN authors use Orchid to design IDNs, verify Orchid's hypotheses, and explore user needs for future authoring tools. This study confirms that authors are open to LLM-driven IDNs but desire strong authorial agency in narrative structures, highlighted in accuracy in branching transitions and story details. Future design implications for Orchid include introducing deterministic variable handling, support for trans-media applications, and narrative consistency across branches.

CCS Concepts: • **Do Not Use This Code → Generate the Correct Terms for Your Paper**; *Generate the Correct Terms for Your Paper*; Generate the Correct Terms for Your Paper; Generate the Correct Terms for Your Paper.

Manuscript submitted to ACM

Additional Key Words and Phrases: Do, Not, Us, This, Code, Put, the, Correct, Terms, for, Your, Paper

ACM Reference Format:

Anonymous Author(s). 2018. Orchid: A Creative Approach for Authoring LLM-Driven Interactive Narratives. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation emai (Conference acronym 'XX)*. ACM, New York, NY, USA, 31 pages. https://doi.org/XXXXXXX.XXXXXXX

1 Introduction

Interactive digital narratives (IDNs) are stories where users influence outcomes through their choices and actions [82]. To keep the narrative engaging, authors define possible user choices and their pre-defined responses, balancing storytelling with interactivity [81]. Large Language Models (LLMs) have improved significantly in recent years. They can simulate human-like dialogue, enabling real-time narrative creation without pre-defined outcomes. In IDNs, LLMs create a new dynamic where authors guide intelligent agents while narratives emerge from player interactions [108, 112]. However, the current application of LLMs to IDNs remains limited. On one hand, many current LLM solutions for IDNs present a low level of authorial agency. Games like AI Dungeon [3] and AI Roguelite [66] demonstrate using LLMs for real-time interaction, yet they integrate LLMs into established IDN genres without fully considering the unique context of each. As such, their narrative output can easily deviate from the author's planned narrative as interactions with the user unfold, sometimes radically changing the narrative scope. On the other hand, the potential of LLMs' real-time creativity is not fully utilized. Works like *1001 Nights* [96] integrate LLMs for semantic analysis and generating NPC responses but limit their use in producing narrative content. Furthermore, most works lack richness due to LLMs' homogenized lexical style, often resembling interactions with a corporate chatbot.

A creative approach, in this paper, refers to a description of methods for creation. In IDN, it serve as abstractions of the IDN authoring process rather than as specific systems. Introducing a creative approach could encourage the emergence of more works and enhance our understanding of how to support those creations, exemplified by work such as *Storylet* [52] and *Novella* [28]. Creative approaches are also embodied in authoring tools, such as the 'natural language programming' in Inform 7 [37] and the 'hypertext linking' in Twine [38]. Creative approach in these tools contribute to the development of distinct IDN genres, with Inform 7 facilitating parser-based interactions, while Twine supports 'Choose Your Own Adventure' (CYOA) styles. While authoring is fundamental in IDN studies [85], limited research has explored this aspect in LLM-driven IDNs, with the major focus on game creation. The current application of LLMs in IDNs reveals a need for a new creative approach for introducing a new human-AI co-creation dynamic, and encourage richer IDNs structures in this field.

In this paper, we introduce *Orchid*, a creative approach for LLM-driven IDN authoring, emphasizing LLMs' potential and raising the need to explore creative methods from an authoring perspective. In *Orchid*, authors define narratives as directed tree structures, outlining causal relationships and specifying LLM responses to player interactions at each node. Each node acts as a bud that blossoms as players progress, driving their progression to subsequent nodes. *Orchid* encourages richer authorial creativity while letting the author balance authorial control with machine contingency.

To support this approach, *Orchid* is grounded in a three-phase design research structure that emphasizes the details of the creative process, illustrated in Figure 2. The first phase employs mixed methods to explore the problem space and identifies key elements of the creative approach. *Orchid* raises two primary hypotheses: the first hypothesis suggests that LLM-based IDNs created with Orchid facilitate an open-ended collaboration between authors, machine creativity, and player interactivity to generate emergent textual content. The second hypothesis posits that Orchid supports the creation of richer narrative structures beyond traditional linear formats commonly found in many LLM-related experiences, such as in AI Dungeon. The second phase develops a technology probe as a computational realization of *Orchid* for investigating users' perceptions of *Orchid* and uncovering their needs [42]. The probe embodies *Orchid*'s core concepts into simple functions, encouraging exploratory use rather than prioritizing usability. In the third phase, we conducted a study of design activities, involving twelve IDN authors interacted with the technology probe to create complete pieces of IDNs using *Orchid*. Through the design activities, participants created novel IDN structures, and articulated key considerations regarding authorial agency and identified specific areas in which they desired enhanced creative output from LLMs. The findings from this study can inform future design implications for developing computer systems that incorporate creative approaches similar to *Orchid*.

The contribution of this paper can be summarized as follows:

(1) *Orchid*, a new IDN creative approach developed using a design-through research approach to support new IDN genres centred around LLMs.
(2) A technology probe that addresses the key ideas of *Orchid*, providing a computational system for using *Orchid* to create IDNs.
(3) An evaluation of how IDN authors interact with the technology. We derive key design insights about authors' intention in LLMs for IDN, focusing on their perception of authorial agency and contingency, and raising their specific needs for implementing *Orchid* in their real-world creative practice.

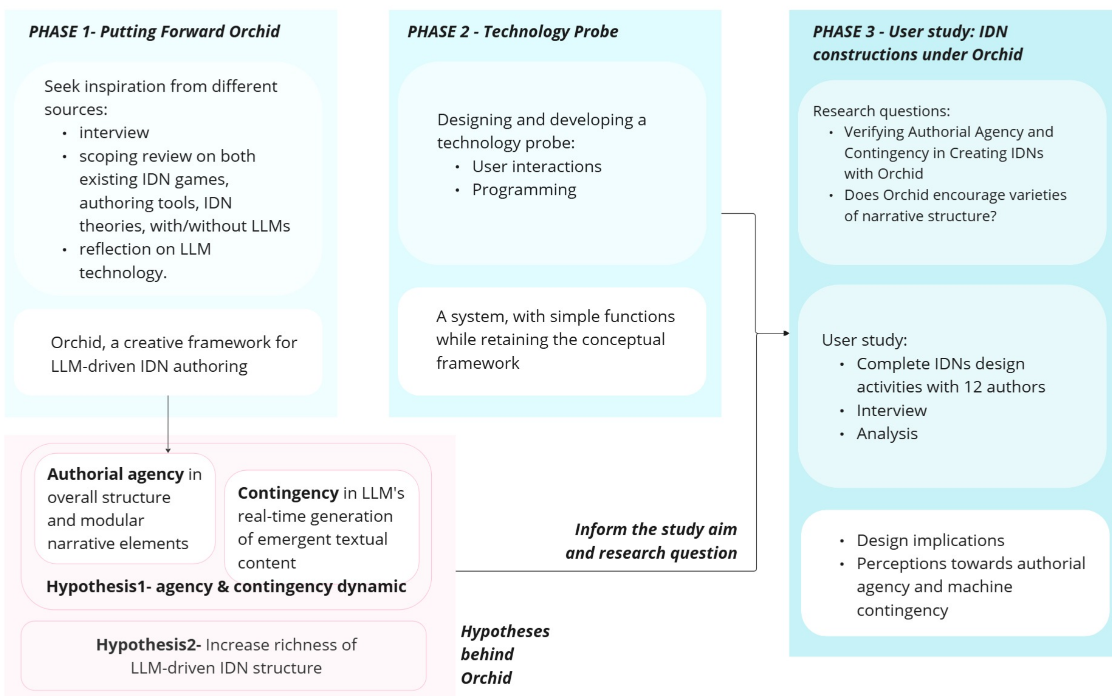

Fig. 2. Research activities of this paper, and their corresponding outputs.

2 Related Work

2.1 LLMs and Automatic Narrative Generation

Advances in computational text generation models allowed intelligent systems to automatically generate stories with minimal human input, replacing the need for fully scripted text [5, 78]. Before the introduction of Artificial Intelligence (AI), studies primarily focused on creating descriptions and dialogue [62] and outlining key events [56]. Methods like Procedural Narrative Generation enabled narratives to evolve through real-time user responses [17, 68, 69], often using parameterized node-trees and clause grammars [71]. However, this approach requires intensive coding, revealing a gap in the ease of authoring dynamic narratives. The development of AI advances this field by enhancing decision-making processes, improving realism in simulations [77, 79, 109, 110], and facilitating dynamic interactions [63, 73].

Large Language Models (LLMs) are notable for quickly generating creative content in natural language [32]. Commercial applications such as Talefy [99], Dream Gen [23], and Storyscape [91], mainly used LLMs in interactive ways for narrative creation, creating a dynamic of human-AI co-creation [61, 115]. LLMs' ability to support story creation and text completion, streamlining the transition from narrative ideas to output, faciliate the process from initial ideation to the final narrative outcome [6, 75]. In addition to the focus on algorithms and applications, the creative support for authoring process is stressed in previous research [115]. Researchers have proposed intuitive approach, which includes using mind maps to guide Chat-GPT's generation [8, 80] and employing metaphors to support the creative process, such as using sketching lines to automatically generate narrative arcs [20] and using magnets and dust to visually aid LLM-driven word-building [21]. While these work focus on "how to produce a narrative", interactive narratives require additional considerations into "how to interact with a narrative" [31].

While LLMs have proven popular in narrative generation, it is crucial to reflect on the authority humans maintain in the context of "accelerationism" [13] and technology-assisted content creation, particularly in balancing the participation of both human authors and AI. Previous studies have put forward methods addressing this, for example, in one approach, the authors can initiate a narrative and LLM decomposes it into logical representations for character action plans [46, 49]. *College Ruled* employs causality weighting to predict and select plot fragments generated by LLMs, allowing for finer authorial control throughout the storytelling process [103]. Wang et al. proposed creative workflow for IDN that enables writers define high-level plot outlines that are later transformed into concrete character action sequences via an LLM-based narrative planning process [108]. These insights indicate that through effective collaboration, AI can not only support the human creative process but also enhance the complexity and depth of narratives, leading to richer storytelling experiences. However, gaps still remain in the singular interactivity those methods can offer, questioning the applicability to complex player interactions.

2.2 IDNs and Authoring

Interactive Digital Narratives(IDNs) combines storytelling with user interactivity through digital media, allowing users to shape the narrative they finally receive [35, 81, 82]. Creative support and authoring methods are important topics for IDN researchers [30]. These are reflected in two aspects: developing authoring tools that provide a complete workspace for creating interactive narratives, actively used within the IDN communities [85], and proposing IDN creative approaches being open to any computational method to realize it in IDN. A creative approach reflects on previous IDN examples and authoring methods to help authors analyze and understand how to create a specific type of IDN [1, 11]. For example, Novella [28, 29] offers a game-centric IDN creation model focused on story objects and narrative states tracked through variables and flow graphs. Storylet [52] employs modular chunks to define scenes, allowing authors to set transition conditions and variables, while players choose their paths. Instead of aiming at a system, our study focuses on a creative approach aimed at enhancing the diversity and complexity of LLM-driven IDNs.

Several typologies of interactive digital narratives (IDNs) have been identified, each with distinct authoring methods. *Script-based branching* is an IDN genre offering high authorial agency, allowing for rich literary expression and enabling users to shape their reading experience [12, 35, 57]. This is exemplified by hypertext fiction like Michael Joyce's Afternoon, a Story[44] and games such as Disco Elysium[116] and Baldur's Gate[58], applies authoring tools with story graphs in their creation process. The tools features story graph includes Inklewriter[43], Articy [39], and Storyspace [93], allowing authors to create blueprints and write detailed content within nodes. Tools like Twine [38] and ChoiceScript [40] visualize relationships through built-in syntax, while Undum/Raconteur [104] and StoryNexus [90] enable branching based on game state variables. However, the benefits of story graphs in outlining narrative paths remain under-explored in LLM-driven IDNs.

Procedural Narrative generation is a genre that narrative evolves through dynamic interactions and real-time user responses [17], similar to a Dungeon and Dragons game master who provides a predefined structure while adapting to player improvisations [68]. The mapping between user interactivity and narrative elements is typically applied in this IDN genre, granting players significant narrative decision-making power. This is exemplified in games such as Dwarf Fortress [2], No Man's Sky [36], Road 96 [22] and Elite Dangerous [26], where entities and events are progressively generated to form a narrative whole. This presents a missed opportunity in current LLM applications, as mapping player actions to narrative parameters could enhance the integration of interactivity and narrative influence.

Emergent Narrative is an IDN genre that the output of it may not be a complete story but rather game logs and NPC behaviors, highlighting how spontaneous player actions can lead to unpredictable outcomes, as seen in sandbox games like Minecraft and open-world games like The Elder Scrolls V: Skyrim [94]. The authoring tools for emergent narrative should emphasize complex rules to facilitate player participation, encompassing story setup and high-level plot development [18, 67, 88], as seen in tools like FEARNOT! [53], which define goals, conditions, and action-emotional reactions. Contingency in *emergent narrative* highlights the core advantages of LLMs. While emergent narratives depend on complex rules for diverse combinations, LLMs act as random 'black boxes,' providing contingency without requiring intricate rule designs.

2.3 LLM driven IDNs and game

Traditional IDN requires a scripting of the most possible dialogue transitions which is often limiting for writers. This scalability problem drives the need for language generation technology to facilitate this process [98], surpassing hard-coded limitations [64]. This is addressed in work including using LLMs in making plans for the event [76], guiding NPC's behaviors without human intervention [19, 106], generating game levels governing complex functional constraints and spatial relationships [9, 41, 54, 101]. NPC dialogue generation is one of the major applications of leveraging LLMs in games [4, 7, 16, 47, 59, 74, 102, 114**?**]. Sun et al. utilized social chatbots to turn fictional game characters into "live" social entities within player communities [97]. Kumaran et al. employed an LLM-driven NPC dialogue approach in narrative-centered games, using Ink for dynamic scripts that allow real-time adjustments of branching dialogue paths based on player choices [55]. Ngaw et al. implemented LLMs to generate text for character representations, object descriptions, and character conversations [74]. NPC generation is well-suited for conversation-based games, but LLMs have untapped potential in managing diverse perspectives and overarching narratives.

For games available in mainstream market, *AI Dungeon* [3] is a representative LLM-driven IDN example. While the AI responds to any command, the narrative shows disruptions at certain points, resulting in a lack of continuity akin to traditional narratives, exemplified in player comments "...can't keep track of what you're doing" [1]. In AI Dungeon [3], the LLM acts as a narrator, reaching the unlimited outcome based on the player's input [113]. In AI Roguelite, various LLMs generate dynamic quests and dialogues [66]. In these examples, while the AI responds to any command, the narrative shows disruptions at certain points, resulting in a lack of continuity akin to traditional narratives. Another example, *1001 Nights* [96] uses LLMs to encourage players to input their own stories and generate weapons for battles with the NPC. However, this design of using LLMs to analyze player input rather than tell stories, limits the machine's imagination and interactivity. The existing application of LLMs in IDNs reveals three shortcomings: a low level of authorial agency, insufficient use of LLMs' real-time narrative creativity, and a lack of richness in the structure of LLM-driven IDNs which often feature homogenized NPC conversations that resemble typical chatbot interactions.

2.4 Authorial Agency and Machine Contingency in LLMs

Contingency refers to the randomness and unpredictability in narrative outcomes, a concept that represents the possible occurrence of future events or conditions characterized by accidents or emergencies [105]. Contingency plays a crucial role in shaping emergent experience in IDNs, where the unfolding story is influenced by player choices and interactions, as an important concept in emergent narrative [18, 53]. The unpredictability in emergent narratives creates a fragile yet engaging experience [10], and combines pre-authored structures with user freedom, allowing meaningful player interaction within a story [67, 86, 87]. An analogy of contingency can be the digital version of game masters(GMs) to enable limitless narrative possibilities based on player improvisation [72].

Introducing LLMs adds a machine element in IDNs, further amplifying unexpectedness and resulting in narratives that can differ significantly with each generation while posing another level of challenges in authorial agency. The black box nature of LLMs, resulting in emergent outputs that are out of author's expectation *i.e. agents killing each other for resources* [19].

This unpredictability is also compounded by the unmanageable variety of prompts that players can provide, especially in natural language [47]. While LLMs can generate an unlimited range of outcomes, without proper engineering, they may produce similar results with minor variations, lacking surprise and richness due to their bias towards positive endings. [50, 100]. The computational approach and authoring tools has been put forward with a focus on maintaining authorial agency, such as Treynor's idea in providing finer control along the branches instead setting all the parameters from the starting [103]. Authoring IDN work with LLM needs an open-ended mindset while keeping the authoring work minimal [74]. The attention to agency in author-LLMs relationship can be associated to a broader debate in Human-AI Co-creation, as similar research was done in drawing [60] and language art [48].

On the other hand, "agency" refers to exercising or manifesting the capacity to act [84]. In IDNs, the author's agency emphasizes the author's ability to shape the storyline [33]. The balance between the game interactivity and the narrative that authors want to express is challenging due to factors such as digital interactions that involve user operations like viewing and navigating and expressive narrative organizes experiences through internal logic, significant sequences, and a quest for closure [70]. With agency shared among players, authors, and the digital system, players have the freedom to make choices and alter narrative mechanics, and the system can govern some levels of the rules, becomes both authors of the final narrative [33, 34, 51].

The relationship between authorial agency and contingency is an important topic, both in IDN authoring, and mixed-initiative machine creativity in interactive systems

[1] https://store.steampowered.com/agecheck/app/1519310/

3 Design of Orchid

In this section, we address the inherent problems mentioned in **??** from the authoring perspective by designing a creative approach called *Orchid*. The identified gaps include: 1) a low level of authorial agency, 2) underutilization of LLMs' real-time creativity, and 3) a lack of richness in the structure of LLM-driven IDNs. *Orchid*, visualized in Figure 1 and described in subsection 3.3, defines an IDN as a LLM-powered directed tree, where each node corresponds to a specific stage in the narrative. Figure 3 demonstrates the process of designing *Orchid*. This creative approach is constructed through interviewing authors, reflection on existing IDN practices mentioned in the related work section and comparing the design opportunities with the LLMs technology.

3.1 Interview

Centering the question of *What IDN format can better integrate with LLMs?*, we conducted a semi-structured interviews with six participants, selected through the author's personal connections. The participants not only had experience with IDNs, but also used LLM in their IDN practice. The interview was conducted through phone call, includes the following two questions: Can you give the example of the scenario when you use LLMs in IDNs creation? What are the difficulties you have faced in the current use of LLMs in the creative process?

Through the interview, we found that all the authors indicated that LLMs are primarily used for automatic narrative generation. Their common workflow involves creating an interaction blueprint, inputting drafts into an LLM (e.g., ChatGPT) to generate a complete story script, and then put the story script into engines like Twine or Unity to program to add interactions.

For real-time generation, the authors primarily use LLMs as interactive mechanisms instead of generating the main story contents.

For example, the interviewees use LLMs to search player dialogues for keywords to trigger events or to generate new game objects.

The interviewees without coding backgrounds also reported the challenges of interfacing with LLMs. In their practice of developing LLM-integrated games, they provide design blueprints to technical teams rather than handling development themselves.

We then explore current LLMs-driven IDNs in academia and in the mainstream gaming market [2]. Existing applications primarily leverage LLMs for NPC dialogue generation and most of them reported the instability issue of LLM in their game [16, 47, 74, 102, 114]. *AI Dungeon* [3] is a representative LLM-driven IDN example. While the AI responds to any command, the narrative shows disruptions at certain points, resulting in a lack of continuity akin to traditional narratives, exemplified in player comments "...can't keep track of what you're doing" [3]. Another example, *1001 Nights* [96] uses LLMs to encourage players to input their own stories and generate weapons for battles with the NPC. This design, which uses LLMs to analyze player input rather than tell stories, limits imagination and interactivity, as one Steam player noted, "it feels more like a constrained writing assignment..." [4].

Both the interviews and existing IDN examples highlight a gap in using LLMs as the primary narrative medium for real-time player interactions, specifically in three scenarios: Using LLMs to assist with script creation before implementing interactive programs, which creates a disconnect between LLMs and real-time narrative generation. Using LLMs to handle interactive mechanism but being detached with the storytelling. Low author expression and

[2] https://store.steampowered.com/
[3] https://store.steampowered.com/agecheck/app/1519310/
[4] https://store.steampowered.com/app/2782660/1001_Nights_Demo/

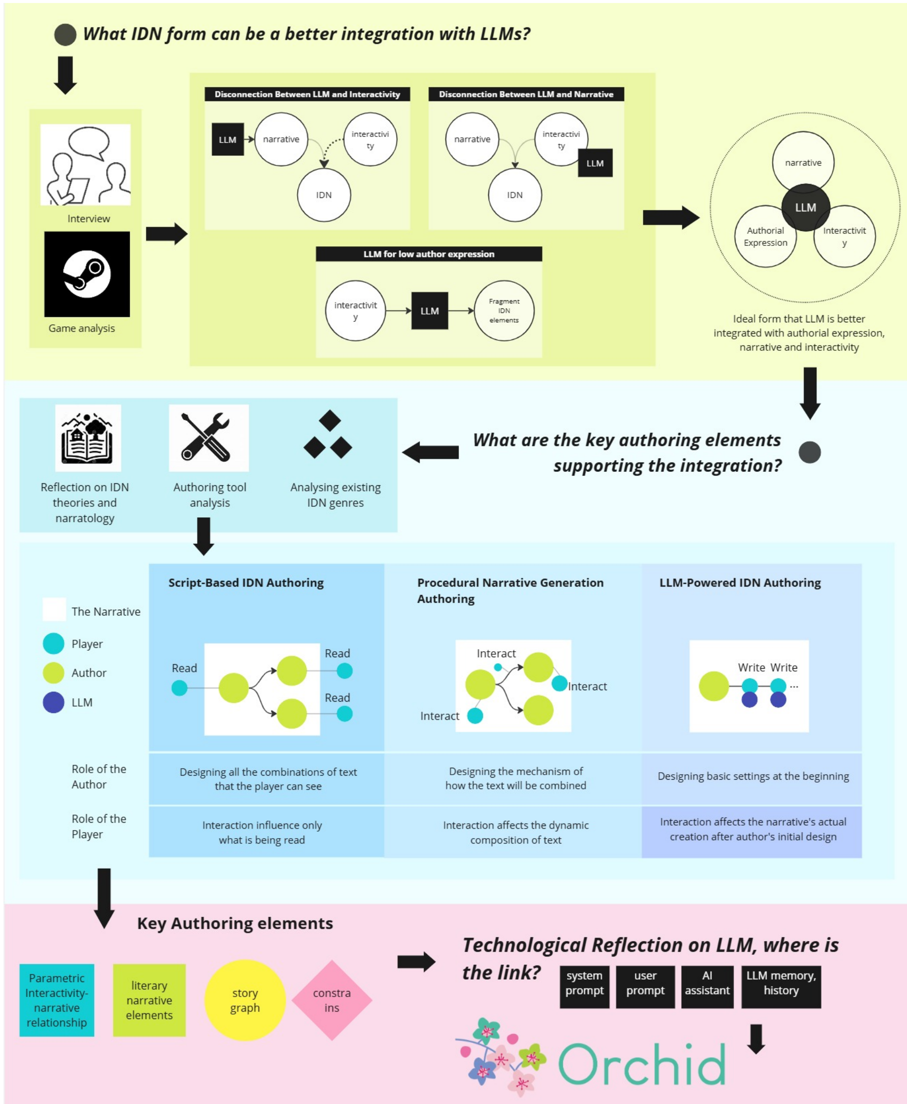

Fig. 3. Three stage overview of the process of putting forward the conceptual approach of Orchid, centering around three questions: What IDN form can be a better intergration with LLMs?(??), What are the key authoring elements supporting the integration?(subsubsection 3.3.1), and where is the link between LLM technology and the authoring elements?(??)

narrative quality, while LLM applied for real-time narrative generation. We propose the ideal characteristics of LLM-IDN integration as the opposite of these situations: 1) LLMs should be used for real-time narrative generation instead of text-completion and writing assistant. 2)The players should have real-time influence on the narrative they receive. 3)Maintaining authorial expression while still providing emergent outcomes.

3.2 Reflection on IDN Practices and LLM technology

Identifying key authoring elements through analysing IDN authoring before LLMs This stage focuses on studying creative IDN genres and their associated authoring tools that emerged before the popularity of LLMs, to react our second design question: What are the key authoring elements of supporting the proposed integration? The identified authoring elements will be described in subsection 3.3. We explored both academic literature and real-world examples to deepen our understanding of the creative processes, particularly the authoring activities behind their creation. Through this analysis, we identified three IDN genres and proceed with detailed discussion of their inspirations for *Orchid*:

Script-based branching inspires the use of **story graph** in *Orchid.* This IDN genre presents high authorial agency, enabling rich literary expression from authors while demonstrating how nonlinear structure allows users to shape their own reading experience [12, 35, 57]. It is exemplified by hypertext fiction like Michael Joyce's Afternoon, a Story [44] and Zork [111], and video games including Disco Elysium [116], Fallen London [24], The Stanley Parable [27], and Baldur's Gate [58]. Story graph interfaces in most of the branching authoring tools facilitate the creation. Graphs in Inklewriter [43] and Articy [39], and Storyspace [93] allow authors to create a blueprint and write detailed content within each node. In tools like Twine [38] and ChoiceScript [40], the story graphs visualizes the jump relationships established through the author's use of built-in syntax while writing their narrative scripts. In Undum/Raconteur [104] and StoryNexuss [90], authors can use the story graphs to define branching based on game state variables. The benefit of story graph in outlining various narrative path is currently under-explored in LLMs-driven IDNs.

Procedural Narrative generation inspires the **link between player interactivity and narrative elements**. Narratives evolve through dynamic interactions and real-time user responses in procedural narrative [17], similar to a Dungeon and Dragons game master who provides a predefined structure while adapting to player improvisations [68]. In this IDN genre, authorial agency is sacrificed for player interactivity, granting players significant narrative decision-making power. The interactivity-narrative elements mapping has been applied in games such as Dwarf Fortress [2], No Man's Sky [36], Road 96 [22] and Elite Dangerous [26], where entities and events are progressively generated to form a narrative whole. This aligns well with our vision for LLMs, as establishing a mapping between player actions and narrative parameters can create a structured integration of interactivity and narrative influence.

Contingency in *emergent narrative* reflects the core advantages of LLMs. The output of it may not be a complete story but rather game logs and NPC behaviors, and rely much on player authority, as seen in sandbox games like Minecraft and open-world games like The Elder Scrolls V: Skyrim [94], highlighting how spontaneous player actions can lead to unpredictable outcomes. The authoring tools for emergent narrative should emphasize complex rules to facilitate player participation, encompassing story setup and high-level plot development [18, 67, 88], as seen in tools like FEARNOT! [53], which define goals, conditions, and action-emotional reactions. However, while it relies on complex rules for greater combinations, LLMs serve as random "black boxes", offering contingency without the need for intricate rule designs.

Across various IDN authoring types, we also identified the opportunities from shared components in different tools. First is **the variety of input modalities**, from previous work has provided specific authoring tools to expand the authoring of types of input [15, 25, 45, 65, 89]. This is absent in current LLMs for interactive narrative creation, which primarily rely on natural language interaction. Second is **Story breaking** [80], such as the modular definitions of characters and various story components found in Articy and Versu. Regarding the user interface, this story breaking is open associated with card sorting based interaction [8, 69, 95], allowing the allocation of different modules into different narrative stages intuitively. The third is **fast testing and history tracking**. Real-time testing interface is an important feature that allows authors to simulate player inputs. Especially for node-based interactive narrative tools, this feature enable authors to test from specific nodes, providing insights into various narrative possibilities.

Technological Reflection on LLMs

We also examine LLM API interfaces and their effect to the interactive narratives, while also exploring how non-LLM creative elements can be enhanced through computational methods. In the API interface of LLMs, such as ChatGPT, the engineering components consist of the user prompt—the query sent by the user to the AI, the system prompt—defining how the AI agent should behave, and agent memory—the history of generated outputs [5].

For player interactivity, When authors hand over their LLM-driven IDN to players, the player's input becomes the user prompt. The user prompt is connected to the player interactivity. By formatting the user prompt, authors can map user interactions to changes of user prompt. Advanced interfaces like Langchain allow the exposure of input variables to influence the system prompt[6]. When applied to player interactivity, this enables authors to define how player input shapes the story in natural language.

For the literary narrating content, as LLMs serve as the narrator in the entire process, while the system prompt acts as guidance for this narrator on how to narrate. system prompt links as it instruct the LLM on what content to present and how to present it. On the other hand, RAG (Retrieval Augmented Generation) can be applied to build knowledge bases for narrative generation. Its main function is to first retrieve relevant information from an external knowledge base and then integrate that information with user prompts [7]. Authors can use RAG to embed specific narrative elements, such as world-settings and character background stories, guiding the LLMs in generating rich content based on the embedded information.

For the story graph, additional computational methods are required to manage transitions, which involve real-time updates to the LLM's system prompt and settings. This can be achieved using a node and graph-based state machine mechanism that provides computational support for managing transitions between different nodes and determining whether the transition conditions are met. We believe this should be based on the generated story, which serves as the LLM's memory.

3.3 Description of Orchid

A Conceptual approach of LLM powered IDNs authoring

3.3.1 Key authoring elements. After analyzing the previous process and comparing it with the current LLM applications, we summarized the following authoring elements that are insufficient or missing in the current LLM interactive narratives, referred to as AE.

(1) **AE1: Story graph.** The key authoring activities that are missing in current LLMs-driven IDNs. Introducing story graph can enable authors to define the narrative throughout the entire blueprint. This may extend authorial agency more than systems like AI Dungeon, which only allow initial settings at the beginning.
(2) **AE2: Constrains.** The key authoring activities that are lacking in current LLM-driven interactive digital narratives (IDNs) include the introduction of a story graph, which allows authors to shape the narrative throughout the entire blueprint. This approach may enhance authorial agency more than systems like AI Dungeon, which only permit initial settings at the beginning.

[5]https://platform.openai.com/docs/api-reference
[6]https://python.langchain.com/api_reference/core/prompts/langchain_core.prompts.prompt.PromptTemplate.html
[7]https://python.langchain.com/docs/concepts/rag/

(3) **AE3: Parametric Interactivity-Narrative relationship.** Authors need to consider how the interactive behaviors, processed by the computing system resulting in the prompt for LLMs.
(4) **AE4: Input modalities beyond natural language.** Following the above item, the input modalities interacting with LLM can be further abstracted instead of letting the player to type anything.
(5) **AE5: Modular Narrative Components.** For narrative involving LLMs, considering the real-time generation capabilities, it may be more suitable for authors to decompose different narrative components, rather than pre-writing a large amount of text in a script-based manner. This includes the contents such as world settings, characters, events, 叙事角度和风格体裁 and especially for LLM, the instruction of how LLM should narrate

Orchid provides a novel lens for integrating LLM's narrative power, authorial intent, and player interactivity, allowing authors to maintain certain control while also leveraging the randomness introduced by LLMs.

Then, we propose a creative approach for authoring LLM-powered IDN. We name this approach *Orchid*, symbolizing the flowers' growth process. Orchid relates to *Storylet* and *Novella* introduced in subsection 2.2, serving as an abstraction of IDN authoring rather than a specific system. The innovation of the Orchid approach lies in the focus of both authorial agency and narrative contingency, allowing authors to maintain certain control while also leveraging the randomness introduced by LLMs. Orchid defines an IDN as a directed tree, where each node corresponds to a specific stage in the narrative.

Figure 1 demonstrates the details of *Orchid*, and how the ecosystem of author-player-LLM brought by it. In the **authoring phase** (the far left on Figure 1), the author constructs interactive narratives based on the five AEs (see subsubsection 3.3.1). These five AEs work together to specifically define the following three aspects

:

- The overall narrative structure, with organizing the relationship between different narrative stages and transition conditions that guide the narrative's development.
- The LLM-related story generation principles for each narrative stage, along with the basic settings that underpin the narrative generation within each stage.
- How player interactions within each stage influence the unfolding story.

Metaphorically, in this authoring phase, the author defines the stem of the flower, the the buds remains unopened. Each bud contains different information, leading to various possibilities for blooming.

In the **playing phase**(the middle on Figure 1), once the entire IDN program is defined, the player begins at the initial node and interacts according to the methods defined by the author. The LLM generates narratives in real-time based on the approach set by the author. The narrative content accumulates through the player's interactions in turns until the progression meets the transition conditions of a nearby node, allowing the player to enter. At this point, the LLM's settings are updated, and the player interacts in next node, with the narrative content constrained by the author's settings for that node. Metaphorically, the player, represented as a bee, serves as a catalyst. Although each bud has its original settings, the likelihood of blooming is influenced by the player's interactions, just as different catalysis methods can yield distinct flowers.

As it gradually comes together into a **narrative whole**, the player navigates through the directed story tree according to the narrative approach created by the author. Each player experiences a unique narrative shaped by their emergent behaviors and the LLM's creativity. Metaphorically, the bee continually interacts with different nodes and influences the blooming of the nodes, gradually resulting in a complete Orchid bouquet, thereby forming a comprehensive narrative experience.

Orchid emphasizes two key concepts: *authorial agency* and *contingency*. *Authorial Agency* in this approach highlights the author's control over the narrative structure and the rules governing player interactions. *Contingency* in *Orchid* arises from both player interactions and the contributions of the LLM. In traditional IDNs, emergent outcomes depend on player actions, but with LLMs, the machine element enhances unpredictability by introducing its own creativity, leading to significantly varied narratives generated in real-time. In *Orchid*, the authors need to consider both the emergent behaviors of players and the openness of LLM-generated text.

At the time of putting forward the conceptual approach, we are unsure how authors will navigate this openness. Thus, a critical inquiry emerges: What levels of *authorial agency* and *contingency* do authors desire, and what specific needs do they have to support both concepts? With this in mind, we developed a technology probe implementing the *Orchid* approach to further explore its creative method and uses, described in the next section.

4 Design and Development of the Probe exploring Orchid

Playful authoring. From the existing tools that applied generative mechanisms, and playful writing in an interesting aspect, the story creation process is changed into an easy, playful interaction [3, 83, 92, 107]

Given that LLMs are a relatively new technology in narrative creation, few authors have direct experience in applying LLMs in interactive narrative design. As such, it is difficult to contextualize LLM-driven narratives in the context of a design study. We thus implemented our contextual approach into a functional system that we use as a probe to further investigate the balance between authorial agency and contingency within *Orchid*, as well as to identify specific user needs. We use this probe in a study to explore the potential of *Orchid* collaboratively with authors. Additionally, it will serve as a foundation for our future system development, allowing us to validate the proposed conceptual approach.

4.1 Design Requirements

We developed the probe to demonstrate our conceptual approach effectively and to integrate the strengths of existing tools from **??** with the characteristics of LLMs prompt-engineering. We thus establish the following seven primary design requirements (DR), following the objectives of the conceptual approach developed in subsection 3.3.

Number: This is primarily characterized by parameterized story generation instead of player's active choices.

- **DR1:** Supporting story graph construction for easy story structure creation.
- **DR2:** Easy stage switching condition creation, allowing authors to establish simple logic for transitions without complex coding.
- **DR3:** Supporting authors in designing the mapping between player interaction inputs and the narrative progression.
- **DR4:** Providing a clear and intuitive graphical prompt engineering workflow that establishes a clear link between the elements on the interface and the operational logic of LLM.
- **DR5:** Enable live testing of the authoring outcomes, allowing users to flexibly adjust inputs that simulate player behavior and experience text generation from the player's perspective.
- **DR6:** Providing familiar narrative elements from the existing IDN authoring methods.
- **DR7:** Ensuring that users do not perceive this probe as designed for non-interactive automatic story creation.

While DR1 to DR5 arise directly from the conceptual approach, DR6 and DR7 are additional requirements to align with the author's previous narrative construction habits and to differentiate the probe from existing LLM-driven tools.

4.2 Interface Design and Narrative Construction Process

Based on the design requirements outlined above, we implement the probe as a web application split into three interconnected pages, each corresponding to a specific step in the narrative creation process. The first page focuses on narrative content authoring. Authors then arrange the content into a storygraph and define the transitions and triggers between the story stages on the second page. Finally, they can test the resulting generative narrative on the third page. Figure 4 illustrates the interaction between the front end and back end of the probe, as well as the flow of the system.

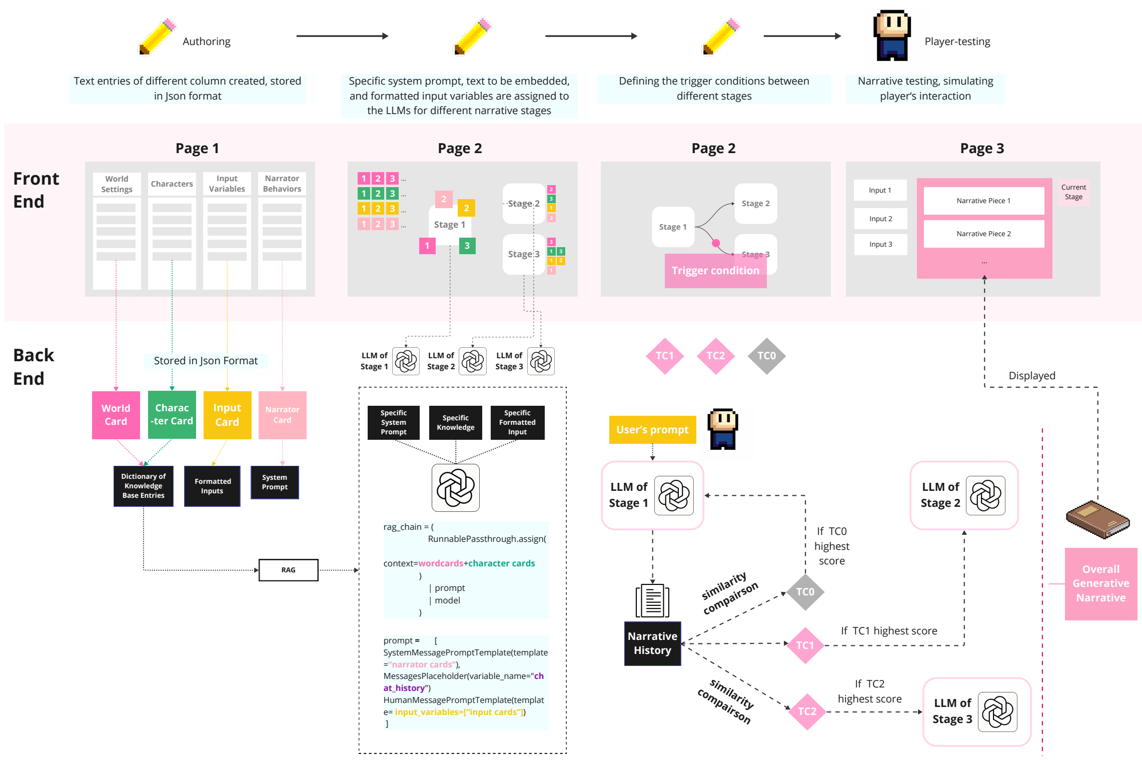

Fig. 4. Front end and Back end of the probe

4.2.1 *Page 1. Textual Narrative Content Authoring.*

In this page, authors can define the LLM-related story generation principles for each narrative stage as well as how player interactions within each stage influence the unfolding story, namely the first and third component of *Orchid*'s creative method defined in **??**.

The page is divided into four columns. The first and second columns contain text related to the *world settings* of the story and *character* definitions, which are essential to the IDN authoring process (addressing **DR6**). This information will be used as text embeddings to form a knowledge base for narrative generation [8], which the LLM will use during the narrative organization process. The third column focuses on *narrator behavior*, addressing **DR4**. Authors input

[8] https://platform.openai.com/docs/guides/embeddings

direct instructions detailing how the narrator should behave, similar to system prompts in LLMs. Authors can define the narrative perspective, style, and content requirements in this column, and these instructions will be integrated as the system prompt for the LLMs. The last column is dedicated to *input variables*, addressing **DR3**, allowing authors to specify how these inputs influence the narrative. For example, if an author wishes to introduce light intensity to affect the emotional tone of the narrative, they can define it as follows:

Variable: Light Intensity Variable Description: A numerical value ranging from 1 to 10; a higher value results in a more positive narrative tone, while a lower value leads to a more negative tone.

We utilize LangChain's native input variable definition approach [9], which is part of its prompt templates. Once an input variable is defined, it will be required for each generation round within the specific narrative stages and influence the narrative based on its description. Users can create multiple entries within each column, and each entry will be instantiated as a card assigned to different stages on Page 2. An auto-content button is designed to provide examples of what can be filled in each column.

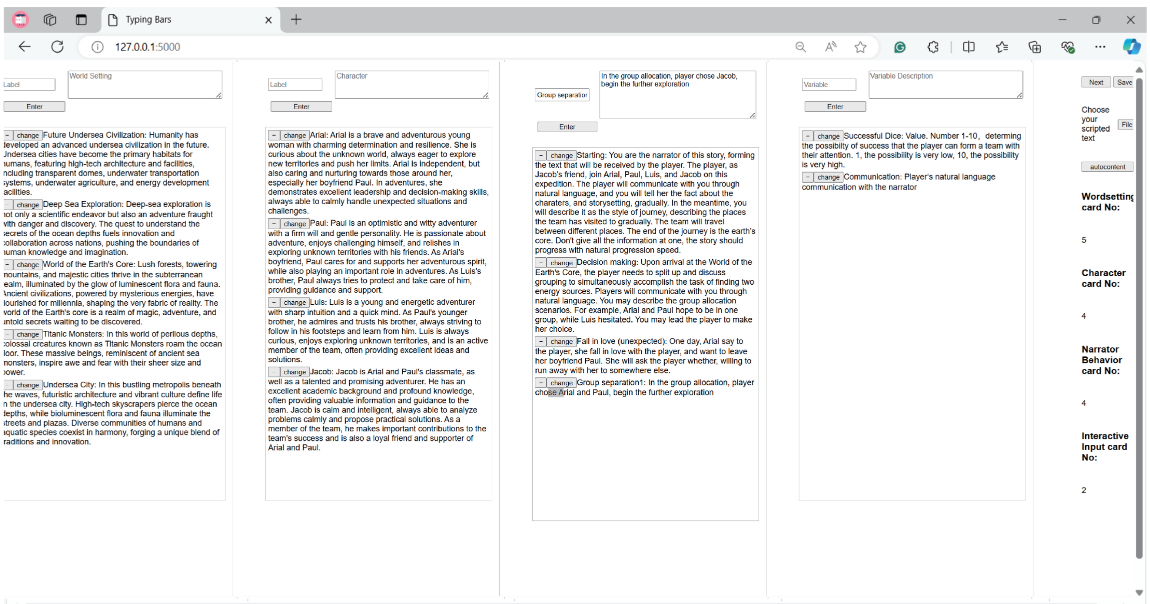

Fig. 5. Probe Page 1. Textual Narrative Content Authoring. There are 4 columns where the author can put world setting, character, narrator behavior, and player input separately. A screenshot from the actual system.

4.2.2 Page 2. Defining the Storygraph and the Trigger Conditions. This page relates to the second component the author will define under the creative method of *Orchid*, which involves creating the overall narrative structure. The design of this page addresses both **DR1** and **DR2**. When users click the "add stage" button, a block is created, allowing authors to the place cards created on Page 1. Each block represents a narrative stage, and users can assign different cards to each block. The same card can be placed in multiple narrative stages. Each narrative stage features an LLM, with the assigned cards providing essential information for that stage's LLM.

[9] https://api.python.langchain.com/en/latest/prompts/langchain_core.prompts.prompt.PromptTemplate.html

By double-clicking the "+" arrow between two stages, an input box will appear for users to define the transition conditions between the two stages in natural language. Users can type in natural language statements, such as “The sisterhood has taken control of the world,” or “when the keyword "love" is mentioned". The transition condition is based on a semantic similarity function that performs text-embedding similarity searches. Each branch’s conditions are compared against the existing generated narrative to select the highest-scoring option. A trigger condition, "stay in current" (TC0 in Figure 4), is included in each stage and compared alongside other stages. If none of the stages meet their conditions, the narrative may remain in the current stage. As the story transitions from one block to the next, the prompt engineering settings of the LLM are updated based on the assigned cards.

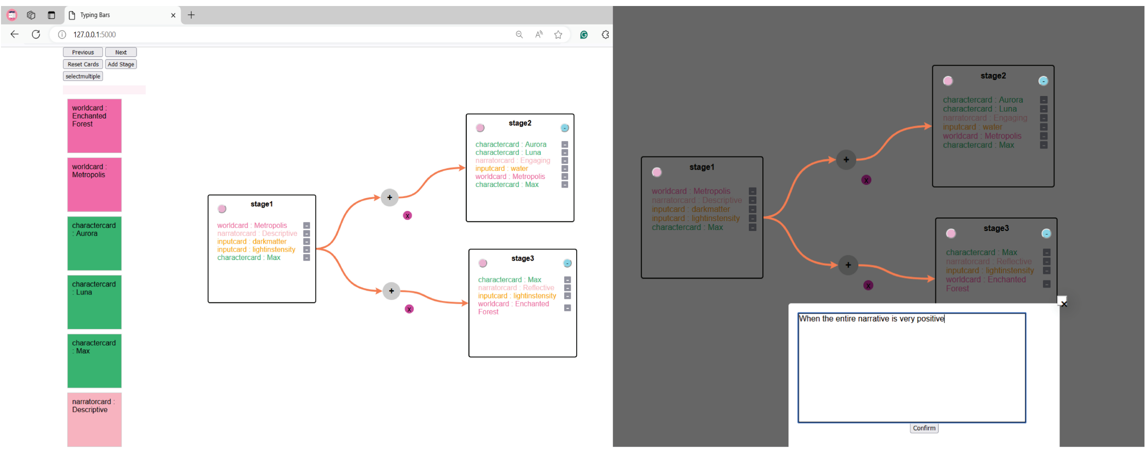

Fig. 6. Probe Page 2: Left: Defining the Story Graph. On this page, the author can create narrative stage blocks and allocate cards into the different blocks. They can then create connections between each block. Right: By clicking the buttons on the edges, participants can type the trigger conditions using natural language. Two screenshots from the actual system.

4.2.3 Page 3. Interactive narrative testing. The design of Page 3 primarily addresses **DR5**, allowing for testing of the authored outcome. In the previous pages, the author focuses on content definition through on cards and blueprints. AI activation occurs only during narrative execution on Page 3. This separation between authoring and testing also addresses **DR7**, where LLM interaction is intended for the player’s experience rather than assisting the authoring process. The narrative testing begins at the first stage of the story graph, with the input field located in the second column of the page. Here, the inputs assigned to this narrative stage by the author will be displayed, along with their descriptions. The author can manually input content to mimic player actions and press enter. After the author decides on the input contents, they can click "generate", and the resulting narrative appears on the right side. The current narrative stage is displayed in the upper right corner. After a piece of the narrative is generated, authors may experiment with different inputs and continue the narrative generation. When the narrative stage switches based on the trigger condition, the items in the input column will update according to the inputs assigned to the new stage. All the generated narratives and their corresponding inputs will be stored in a message list. Authors can scroll through the list to view the history. Each narrative piece has a small icon beside it; clicking the icon will show the input that leads to that narrative piece. Authors can also provide instructions in the left-side text box if the generated narrative does not match their intentions. These instructions can serve as a reference when the author decides to return to previous pages for authoring content modifications.

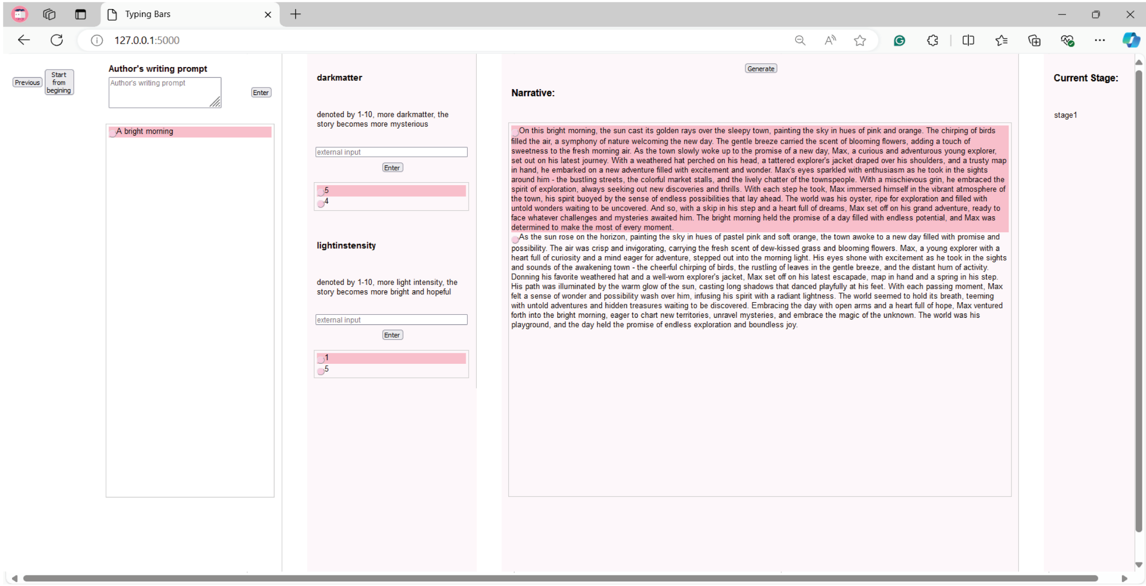

Fig. 7. Probe Page 3: Interactive Narrative Testing. On this page, the LLM will operate, and the author can assign values to the input variables they designed for the current stage. The generated narrative content will be displayed step by step. Selecting specific text will highlight the related inputs, and the current stage will be shown in the upper corner. A screenshot from the actual system.

4.3 Probe operation

Authors interact with the probe through the web interface. The data is then processed and sent to the server, which interfaces with the LLM to generate the narrative. Figure 4 demonstrates the communication between the user interface and the back end. We use GPT-3.5 for this probe as the LLMs for generating text in real-time and text-embedding-ada-002 for text embedding. The probe design is structured in a stage-by-stage manner, utilizing an algorithmic architecture tailored to LLM prompt engineering. This stage-by-stage approach allows for the dynamic allocation of system prompts and knowledge at specific narrative stages.

The entire backend program is coded in Python, leveraging Langchain [10] for text embedding and input variable definitions. The front end of the probe is built using the Flask APP approach [11] in Python with JavaScript and HTML.

On Page 1, the entries that the author types within each column are stored in JSON Format. On Page 2, when the author creates a new stage, a corresponding item is instantiated in the backend dictionary. As users drag cards into the stage, the values associated with that stage are formatted based on the information derived from the cards. World setting and character cards contribute to information retrieval, utilizing a RAG mechanism. The assigned narrator behavior is incorporated into the system prompt template for the LLM of the specific stage. Input cards are stored in a dictionary where the names of input variables serve as keys, while their descriptions become part of the prompt. Using Langchain's input variables definition approach, the LLM model guides how each input variable influences subsequent generations.

[10] https://www.langchain.com/
[11] https://flask.palletsprojects.com/en/3.0.x/

User-defined trigger conditions for all branches of a specific narrative stage are stored in the [TriggerConditions] list of that stage, including a TC0 condition that does not trigger any transition. When a piece of narrative has been generated, a function performs a similarity comparison, calculating the similarity score between the generated narrative and all the trigger conditions from the list and returning the name of the stage with the highest similarity score:

```
def stage_router(user_input):
            global stage
            embedding = embeddings.embed_query(narrative_history)
            sim_list = cosine_similarity([TriggerConditions], embeddings)
            best_similarity = sim_list[0]
            most_similar = trigger_conditions[best_similarity.argmax()]
            for template in current_trigger_list:
                if template[1] == most_similar:
                    label = template[0]
                    break
            stage = label
            return stage
```

On Page 3, the author simulates player behavior by inputting values for various input items. These values are formatted into user prompts and sent to the model for generation. Each narrative stage maintains its own history list. Whenever a piece of narrative is generated, a new entry is added to this history. When a trigger condition is activated and the narrative stage transitions, the history list is updated accordingly. This approach prevents inaccuracies in trigger conditions judgment. It ensures that the functionality of the trigger conditions operates as defined, based on the current narrative stage rather than the entirety of the previously generated story.

5 Probing Orchid through Design Activities

The probe we developed exemplifies the conceptual approach of *Orchid* in authoring LLMs-powered interactive narratives. Following this, we conducted a probe study in the form of IDN co-design workshops. This study aims to explore the creative experience of *Orchid* from the author's perspective and highlight gaps and future directions for this creative approach.

Specifically, this study focuses on the following research questions:

- **RQ1:** How do authors create interactive narratives under Orchid?
- **RQ2:** How do authors perceive *Orchid* in authoring interactive narratives, particularly regarding their perceptions of authorial agency and contingency?
- **RQ3:** What are the design implications that can support the creative method of *Orchid*?

This creative approach aims to encourage richer creativity for authoring LLM driven IDN, and raise the problem of balancing authorial agency and machine contingency while not aiming at solving it. To support this aim, we developed the creative approach into a technology probe [42], to investigate the critical inquiry from Orchid, specifically, we raised the following research questions:

- RQ1: What levels of authorial agency and contingency do authors desire when creating IDN using Orchid?
- RQ2: What type of narrative structure that author want to use Orchid to create, or does Orchid encourage varieties of narrative structure?

- RQ3: What are the specific needs for future systems supporting the creative approach of Orchid?

Through the study using technology probe, we observed how 12 IDN authors constructed new types of narratives using the Orchid approach, and summarized insights into future design implications emerged from the study.

5.1 Study setup and procedure

12 participants (8 females and 4 males) volunteered for our study. All participants were native speakers of the same language as the researcher. Their ages ranged from 20 to 28, with an average age of 23.25. Participants were recruited through social media and personal connections, and they were selected based on their experience in at least one of the following experiences: 1) Traditional writers interested in using text as their primary medium for producing written narratives. 2) Game designers or developers with experience in making IDN games. 3) Artists who have experienced in applying IDNs in their practice. Table 1 summarizes their demographic information. We also encourage participants not fitting this background but interested in the combinaion of LLM and IDNs. All participants have experience interacting with LLMs, and three of them participated in the initial exploration interviews. The participants' backgrounds span interactive storytelling, AI game research, narrative art, and various aspects of human-computer interaction.

Table 1. Demographic Background of Participants

Participant No	Gender	Career	Past IDN Experience
P1	Male	Designer	has applied IDNs in design.
P2	Female	Game Designer	has experience of LLM in IDNs.
P3	Female	Student in Chinese Literature	table top RPG writer.
P4	Female	Virtual Interactive Film Production.	applies to interactive film production.
P5	Male	Researcher in AI and Games	has designed several interactive narrative games.
P6	Female	Interaction Design student	experienced in hypertext games.
P7	Female	Curator	applies IDNs in curation practice
P8	Male	Social Design student	focuses on tangible interaction.
P9	Female	Narrative Art student	conducted a GenAI project in interactive narratives.
P10	Female	Creative Writer	player and writer of interactive narratives.
P11	Male	Web3 Designer	applied narrative design in projects.
P12	Female	Student in Chinese Literature	worked in a company for children's storytelling.

LLM, template. templatewordtemplateprobe, participantsparticipants

On the day of the user study, participants signed a consent form allowing for the collection of personal data and for interview recording, which would be included anonymously in future study presentations. LLM, LLM We created a demonstration video outlining the entire probe usage process. Participants were also given a writing template to prepare their text and story approach in advance, which they could use during the formal study. These materials were distributed to participants after they signed the consent form.

In the formal study, We established a remote network link allowing participants to access the probe online. Meanwhile, we collected the generated data on the local server. Before testing, we demonstrated the probe's usage process again to ensure participants' understanding. The entire process was conducted via Zoom and was video recorded. Once participants began using the probe to construct and test interactive narratives, we encouraged them to think aloud and express any thoughts or questions during the process. When participants felt their story was complete, we conducted a semi-structured interviews and asked them to fill a questionnaire.

Our semi-structured interviews focused on several topics:

1) *Description of the creative process*, addressing **RQ1**. This topic explores how *Orchid* expands what authors can achieve in interactive narrative (IDN) creation. Questions include: Describe your understanding of the system and introduce the IDN you developed. How does this creative method impact your narrative ideation? How do you perceive differences compared to your previous method of writing an interactive/static story?

2) *Role of AI's participation in the creative process*, which also addresses **RQ1**: Do you think AI is more of a creative asset or a writing assistant? How do you think of this shift?

3) *Authorial Agency and Contingency*, addressing **RQ2**. Before asking these questions, we provided participants with an explanation of the concepts of authorial agency and contingency as outlined in the conceptual approach. The questions under this topic include: Describe how your authorial agency is manifested when using this system to construct interactive narratives. In what ways do you see contingency reflected in story generation, and how do you understand this contingency? How do you perceive the balance between authorial agency and contingency when using this tool? How have your expectations of authorial agency changed compared with the existing authoring method?

4) *Perceived difficulties, advantages, and open suggestion*, addressing **RQ3**. This topic includes how the conceptual approach can be applied to participant's current workflows, anticipated features, and open feedback.

The questionnaire consists of six questions providing quantitative support for the study, inviting participants to rate the level of the following items on a scale of 1 to 5:

- The focus on the authorial agency when using the system.
- The focus on the contingency when using the system.
- Quality of the narrative outcome.
- Willingness to iterate.
- The extent to which they believe the system can support various narrative structures.
- The extent to which the system has changed participants' understanding of creating interactive narratives.

The entire study lasted 1 to 2 hours per participant, depending on the complexity of the IDNs that they constructed. Due to the specific local context in which the study was conducted (bilingual Chinese-English), participants were allowed to use both English and Chinese to construct the narrative.

5.2 Analysis

We used the automated transcription service provided by Zoom to convert the interview recordings to text, reviewed the results compared to the original audio, and corrected the transcription errors. Firstly, a qualitative analysis was conducted on the story that participants created during the study, with their prepared templates as support. Observations focused on how participants designed various inputs and how these inputs influenced the narrative and interactivity, as well as the overall narrative structure that the authors designed. This phase also analyzed the level of detail in the information provided by participants when constructing the narrative (the information they filled in on each card), assessing whether it was complex or relatively simple. Secondly, we conducted a qualitative analysis of both the recordings of participants' think-aloud processes and their responses to the interview questions. We documented participants' questions and thoughts as they engaged with the system, paying particular attention to their desired story structures and any challenges they faced. This analysis helped identify functionalities that could better support their narratives and highlighted design issues that might hinder user intentions. We used thematic coding in this part [14]. Two researchers participated in the coding. Each researcher independently coded the data before sharing results to identify common codes, which were then organized into axial codes for further analysis. This process included a second round of coding and inductive refinement of subcodes within each thematic area. Finally, selective coding was conducted to create a comprehensive codebook summarizing the key insights. The coding was conducted using Taguette [12]. A total of 5 themes and 105 codes were identified and agreed among the two researchers. Lastly, we calculated the value obtained through the questionnaire.

6 Results and Findings

6.1 Summary of Participant-Created Stories

We summarized the interactive narratives created by the participants, focusing on key aspects such as narrative structure, trigger conditions, player input, worldbuilding, and narrator behavior. Given the significance of narrative structure, we will illustrate this point with specific examples.

6.1.1 Narrative structure. Three types of narrative structures arose from the study.

Linear progressive. P5, P7, P8, P9, P10 designed this type of narrative structure. This narrative structure features progressions from one stage to the next, and aligns with traditional plot development and linear storytelling from film and drama, while allowing for precise scene segmentation. P7's interactive narrative is an example of this narrative structure, adapted from the romantic film *Before Sunrise*, simulating a cinematic scene with segmented shots and interactive plot development. Players take on the role of the female protagonist, experiencing events that unfold on a train. The narrative is divided into three stages: Stage 1: In this initial stage, the player witnesses a couple arguing on the train. The player can choose to intervene in the argument, and their actions can lead to the couple leaving the train compartment, transitioning them into Stage 2. Stage 2: This stage expands the background story, revealing the protagonist's identity as Selina, the female protagonist. Players gain insights into Selina's backstory. Stage 3: When the train comes to a stop, players enter the final stage, where they encounter Jessie, the male protagonist, and engage in conversation.

Branching. Under this structure, the narrative begins with a clear starting, which then branches out into multiple paths, effectively illustrating a multipath structure. Distinct branches often lead to significant shifts in the background settings, while the characters, especially the protagonist, may remain unchanged. For example, in P2's design, the story revolves around Neo, a virtual character living on YouTube. Players take on the role of commenters in a YouTube discussion, providing feedback on Neo's actions. The narrative features two branches: if players leave very negative comments, Neo will head towards a desperate desert; conversely, if the comments are positive, Neo will reach a new space dimension called Vtube. This structure was applied by P1, P2, P5, P11, and P12.

Looping. This narrative structure features shifting perspectives and the characteristic of going back and forth, allowing for a dynamic exploration of the story's development. Each narrative stage features distinct narrator behaviors. For example, in P4's story, which follows a protagonist switching identities across different times and spaces, three narrative stages are designed, each with unique narrator behavior. Stage 1, Omniscient Mode: In this mode, the story reveals numerous details to the player, providing a comprehensive understanding of the narrative context. Stage 2,

[12]https://app.taguette.org/project/69964/highlights/

Multiple Perspectives Mode: Here, the story presents information from various angles, creating a dialectical experience that includes both truths and falsehoods. Stage 3, Negative Mode: Inspired by Black Mirror, this mode contrasts player actions; when the player expresses positivity, the narrative turns negative, creating an ironic effect. Players will navigate through different narrative stages, cycling to discover their true identity.

ACM Conference Proceedings Primary Article Template/Figure/Examples.jpg

Fig. 8. The three types of narrative structures created by participants using the probe, along with screenshots of the example story graphs. The linear progressive structure is from P7, the branching structure is from P2, and the looping structure is from P4

6.1.2 Types of trigger condition. Most trigger conditions are based on the semantic analysis of the generated narrative and can be categorized into: the occurrence of specific events or plots, numerical values (such as elapsed time or certain overall metrics), character developments, semantic similarities (such as whether the narrative progresses to positive outcomes).

6.1.3 Types of input. We found that user-designed inputs can be categorized into three types: Natural Language: This invites players to interact with the narrative through dialogue or comments. For example, "comments" in P2's narrative. Number: This is primarily characterized by parameterized story generation instead of player's active choices. Specific aspects of the story, such as relationship dynamics, choices, and tension levels, are parameterized. This influences particular events and character actions within the narrative. For example, P11 designed a fighting story, introducing the value of explosions and sparks, the higher the value, the more wars. Bool: Similar to value inputs, this consists of simple yes/no choices. For instance, in a story designed by P1 about a judge's rebellion, the protagonist's choice to resist or not is determined by a yes or no input, affecting the plot's direction. Overall, participants design inputs that impact the plot's direction and the content of various narrative stages. However, some participants mistakenly linked input values with stage trigger conditions, believing these values were derived from the generated narrative. For example, in P7's narrative design, if the argument rounds exceed three, the couple leaves the train car. P7 initially allowed this number to be input by the player, but it should be derived from the narrative rather than entered by the player.

6.1.4 Narrative worldbuilding. . Depending on the author's intent and past writing experience, the information they provide can be long or short. Some authors focus more on the relationships between the characters, with half of them making these relationships a central focus of the information building on Page 1. Authors with strong writing backgrounds, such as P10, adapted past writing material and provided extensive text in the worldbuilding.

6.1.5 Narrator behavior defining. . In defining the narrator behavior, seven participants mentioned specific styles, genres, and desired narrative perspectives. Some participants also defined particular events they wanted to see in the story and the perspective of the narrative. Participants tended to input their preferred perspectives and writing styles directly, rather than in the typical wording when defining system prompts of ChatGPT. These insights from the story analysis inform the current limitations of the probe and indicate what functionalities future systems should provide.

6.2 Findings from coding

During the coding, five themes were identified. Table 2 summarizes the themes and subthemes from the result of coding.

Table 2. Themes and Subthemes

Theme	Subthemes	Code
Authorial Agency and Contingency	- Perceptions of contingency - When/where authorial agency is needed - Highlighting authorial agency of the system - Insufficient authorial agency of the system	22
How authors perceive the creative process demonstrated by the probe	- User perceptions of AI involvement - Concerns about this creative method - Perceived differences with other methods - Initial perceptions of this method	28
Narrative ideation	-Based on existing literature -AI-ization of previous game ideas -Combine AI with narrative	7
Description of the ideal narrative	-Coherence -Accuracy	8
Suggestions for the tool/Specific Needs	-Ideal functions -Where the current probe design confuses users -Advantages of the current design of probes	40

6.2.1 Perception of Authorial Agency and Contingency. Seven participants appreciated the contingency introduced by the probe, expressing a desire for the system to generate unexpected narrative outcome. They believed that the appeal of this creative method lies in the contingency provided by the LLM. While considering the contingency of LLM, participants also link it with the player authoring agency. P9 remarked that this contingency can "encourage authors to consider the emergent behaviors of users and transfer authoring agency to them."

However, observed from the interview result, we found that contingency in *Orchid* is conditional. The first condition is the author's design intent, which depends on whether the author focuses on playful mechanics or a specific theme they wish to express. This aspect was mentioned by participants P1, P3, and P7. Depending on different design intents, expectations regarding contingency may vary. Another condition is the premise of accurate story structure delivering.

If an author finds the generated content deviating from their established approach, their openness to contingency may significantly diminish. For example, P1 stated, "If I have a clear narrative structure, I don't want unexpected elements to disrupt it; otherwise, I might view the LLM as quite stupid."

Regarding when and where agency is needed, ten participants emphasized the importance of maintaining control over the main storyline rather than focusing on the details. P10 remarked, "I wish I could control the main storyline while leaving the descriptive text to the GPT." However, participants also expressed concerns about insufficient authorial agency presented in the probe. They were particularly attentive to whether the established plot and information were accurately conveyed by the GPT, rather than fabricated. This highlights a focus on ensuring that the GPT understands the world-building information defined by the author. This topic also reveals deficiencies in the current probe, as some information can be confusing and details often lack specificity. For instance, five participants experienced uncertainty regarding character relationships. P10 designed a family story featuring a character named Lin, but during the story generation, the LLM replaced Lin's name with "K" and altered her role within the family. This was unacceptable to the author. Additionally, in terms of story progression, users expressed a desire for agency over the pacing of the narrative. Eight participants reported a loss of authorial agency when trigger conditions deviated from their original definitions. P5 stated, "The transitions in the story are the aspects where we, as authors, need to retain our own thinking. This is something I believe should not be left to AI."

Meanwhile, participants emphasized the story graph design of the probe as a means to foster a strong sense of authorial agency. The card sorting feature enhances the logical flow of the process, clarifying the author's creative guidelines. P1 described the relationship between authorial agency and contingency in *Orchid* as "unexpected yet reasonable."

6.2.2 Perception of Creative Process of Orchid. This theme reveals author's thinking of *Orchid*, what difference in IDN authoring experience this conceptual approach brings. The first aspect is User perceptions of AI involvement in the creative process. Five participants viewed AI as an asset instead of an assistant. For instance, P9 explained that AI should not merely be seen as a tool for extending writing, but rather as an integral part of the overall creative experience. However, some participants expressed concerns about the creative process, also linked to the participation of AI. 3 participants worried that incorporating AI mechanisms might sacrifice original ideas, while others were concerned about the easily made logical errors, emphasizing the necessity of a thorough understanding of the system for achieving comfort and confidence in its use.

The second aspect is the comparison with existing authoring experience, participants emphasized the need to consider contingency, and the distinction between writing a complete piece versus creating an outline. For example, P8 noted that the system promotes a more open-ended creative model, encouraging user participation and co-creation. P1 remarked, "The entire story feels collaborative, striking a balance between the unexpected and the logical."

Initial perceptions of this creative process included its flexible yet structured nature (noted by 3 participants), the ability to "simplify complexity," and the potential to lower barriers to creation. P5 noted that "Using the probe to create, I have a clear sense that I am creating a approach." 11 participants mentioned the ease of adapting to their current workflows and the possibilities for experimental writing education and tangible interaction.

6.2.3 Narrative ideation. Participants conceptualized the probe's use in various ways, primarily aligned with AI-generated narrative elements (7 participants). They kept the "system prompt" in mind while creating the basic narrative settings, often drawing on past experiences of prompt-engineering LLMs. Three participants emphasized trigger logic in their narrative structuring, utilizing the probe to design narrative approaches, such as P4's looping structure. Some participants referenced existing literature and adapted their previous game ideas to fit the "AI" context. For example, P2 redesigned the Neo adventure, which had previously been a non-generative, pre-scripted narrative. This time, the *Orchid* approach encouraged P2 to present an open-ended version of her pre-scripted story.

6.2.4 Description of the ideal outcome. Based on the participants' feedback, they have the following expectations for the final narrative generation results:

Coherence: Four participants emphasized the importance of coherence in the narrative, highlighting the need for contextual connections of narrative between different stages. They pointed out that sometimes the plot does not progress or fails to connect with previous narrative pieces, leading to a disjointed overall experience. Additionally, due to the current design of the probes, which update the history at different stages, the coherence of the narrative can be further impacted.

Plot Progression: Similarly, four participants expressed concern about whether the plot was advancing and identified a lack of progression as a significant issue. They want the narrative to stay aligned with their original intent, ensuring that the overall development feels coherent. For example, in P10's actual story, the system revealed all the information from the first narrative stage, which disrupted the overall pacing of the narrative. This imbalance can create a sense of the story being "out of control."

Perspective Differences: Participants also mentioned the differing perspectives of themselves as authors and potential players. Two participants noted that players may not pay much attention to the overall coherence of the narrative while playing. In contrast, as authors, they place great importance on the accuracy of key plot points, insisting that these must align with their basic setting.

6.2.5 Design Implications. Several user needs were identified, from the interview record linking to the ideal functions mentioned by the participants, the report of where the current probe design confuses users and the lighted advantages of the current design of probes. These insights serves as design implications for future work.

- **Stricter Trigger Judgment.** Five participants suggested improvements to the trigger mechanism and clearer control over triggers. Three participants noted that the narrative pace is often too fast due to inaccuracies in trigger judgment. P5 indicated that trigger judgments should not be left to similarity comparisons, believing trigger conditions are a core aspect of authorial control over the narrative. Relying on AI for this process poses risks, as it can lead to a loss of responsibility and potential plot instability. Most participants stressed that trigger judgments should be strict, rather than depending on a "black box" mechanism depending on similarity comparison.
- **Improving narrative coherence.** Memory emerged as a major concern, with five participants mentioning the importance of contextual coherence. They pointed out that during dynamic generation and stage transitions, if the entire memory is rewritten, it can lead to problems. Since the current probe lacks dynamic memory management and updates information at each stage, this results in issues with overall narrative coherence.
- **Providing tutorials and narrative templates.** Six participants expressed the need for guidance on defining variables and input parameters. They noted that the wide variety of design options makes it challenging to clearly define inputs, highlighting the differences between this process and their experiences with other interactive narrative projects. They suggested the implementation of predefined templates to support narrative structuring, allowing users to easily load these templates to establish different narrative approaches.

- **Dynamic Text Modification and Retention.** Participants reported the intention of the ability to directly modify the generated text in a way that influences subsequent iterations of the output. Instead of requiring authors to manually remember changes and re-enter them later, the model should seamlessly incorporate these adjustments into its understanding. As suggested by P1 and P5, when authors are dissatisfied with a specific segment, they should be able to edit that text directly. This modification should be retained by the model, allowing the narrative to evolve in alignment with the author's intent as the story progresses. Such an approach can enable a more intuitive and responsive interaction between the author and the mode.
- **Integration with Interactive Input/Output.** Participants expressed a strong desire for the tool to support connections to real interactive input and output systems. From the perspective of interactive input, participants envision direct data integration with commonly used game engines, such as Unity. This would allow for a seamless flow of information and enhance the tool's utility in dynamic environments(P2, P3, P5). On the output side, there is a call for integration with visual development tools like TouchDesigner. Additionally, participants, including P9, expressed interest in connecting with other generative AI platforms focused on visuals, such as Stable Diffusion and Midjourney. This integration would facilitate a more cohesive workflow, allowing the creative method of *Orchid* to seamlessly blend into existing processes and significantly enhance its practical applications.
- **Enhancing explainability** Participants expressed a need for improved explainability in the content creation process. They would like examples that illustrate the appropriate level of detail when defining a card, as it is currently unclear how much content should be included to ensure high-quality output in later stages. P11 and P12 both suggested that to support users unfamiliar with prompt engineering, another AI assistant could be implemented to clarify the relationship between the initial content defined in the card and the stories generated thereafter. This assistant would provide guidance on optimizing card definitions, making the process more user-friendly and enhancing overall content quality.

7 Discussion

7.1 Summary of Insights

The exploration into authorial agency and contingency within the *Orchid* approach reveals several key insights. Authors express expectations for authorial control primarily concerning the overall structure and the criteria for narrative transitions, as well as the precision in predefined details. Their anticipation regarding contingency focuses on the system's ability to generate descriptive details and account for players' unpredictable behaviors. Regarding the perception of the creative process, participants identified that the story graph planning of narrative branches, enables authors to strategize AI behavior at each stage. Additionally, participants highlighted the ability to define player interactions in various forms, mapping them as input variables that influence the narrative, rather than limiting the experience to simple dialogue, as seen in traditional AI dungeon games. Pariticipants also held This approach holds potential for transmedia applications, suggesting that games and interactive installations could be the next steps in utilizing text generation.

7.2 Future Directions

Building on the design implications, it is essential to outline specific future directions of *Orchid*. Firstly, to enhance narrative quality, more rigorous research is needed, focusing on various prompt engineering strategies and iterative evaluations of authors' subjective assessments. Our current paper serves as a proof of concept, highlighting significant room for improvement, especially in enabling GPT to produce cohesive narratives aligned with the author's intentions. Additionally, implementing a dynamic memory management database is crucial for improving narrative coherence. This memory should represent narrative progression and utilize effective approaches like knowledge graphs. By helping the model understand the author-defined structure, we can enhance coherence through a hierarchical information representation. Knowledge graphs can guide the generation process by storing information and transforming it into comprehensive prompts, thus improving overall coherence. Secondly, future development should focus on designing trigger conditions. The goal is to create a user-friendly strategy, reinforce authorial agency, and ensure that trigger definitions closely align with narrative progression. Lastly, there is some confusion regarding how this system will function within an actual interactive context. This highlights the necessity of developing an API to demonstrate its practical application in real gaming scenarios.

7.3 Role of the player

Orchid has sparked important discussions about the role of the player. The relationship between the reader and the author can often be unstable and dualistic. During the study, some authors reported confusion about their roles, oscillating between being the creator and the player. This dynamic aligns with the concept of participatory reading. In terms of player experience, it is crucial to explore how players perceive and engage with the interactive narratives produced by *Orchid*. This calls for further investigation into the ecological role of players and their interactions with narrative structures, highlighting the need to understand their engagement beyond conventional gaming paradigms.

7.4 Comparing Orchid with Existing IDNs creative approach

Orchid situates different from exsiting enhances authorial agency and contingency in several ways. In script-based narratives, player interactions do not contribute to the overall narrative creation. Instead, players navigate through a static database of stages without altering the plot [81]. In this context, authors maintain full agency over the narrative. For example, in the interactive fiction, "Afternoon, a story"[13], players explore different plot points by clicking on hyperlinks, engaging with the narrative without affecting its fundamental structure. The *Orchid* approach incorporates the concept of story graphs and branching paths for enforcing authorial agency while allowing the players to directly contribute to the narrative creation.

In contrast to script-based narratives, procedural narrative generation involves the author establishing a approach while player interactions influence the dynamic composition of content. The narrative remains modular and parametric but is ultimately authored. For example, in Roguelike games like *Faster Than Light*[14], random events introduce variability. *Orchid* draws inspiration from this parametric interaction and narrative influence relationship to provide both authorial agency and contingency. This approach is reflected in how authors define the impact of player interactions on the narrative in *Orchid*.

LLM-powered narratives grant players considerable freedom to influence the story by prompt, often reducing the author's role in story generation past the initial setting. For example, in *AI Dungeon*, the author sets the exposition stage and provides foundational details. The rest of the narrative revolves primarily through player prompting with the LLM, leading to a diminished authorial presence along the narrative generation. Players can easily stray away from the initial setting through their actions. For instance, a fantasy setting can turn into a completely different experience if the

[13] https://ifdb.org/viewgame?id=oez7iutmip7gj79f
[14] https://subsetgames.com/ftl.html

player inputs "I take a spaceship to Mars", with no mechanisms to bring the narrative back to the setting expected by the author. However, the opportunities that LLMs offer for open-ended narrative experiences and their prompt-based interaction nature create substantial possibilities for contingent storytelling. *Orchid* inherits this approach, leveraging LLMs to enforce the contingency in narrative experience.

Orchid offers a new form of authorial agency-contingency for LLM-driven interactive narratives, In our conceptual approach, contingency arises from both player interactions and the LLM's generative nature. Authorial agency and contingency exist in a dynamic relationship shaped by how the author defines the rules for player and LLM engagement. Rather than being in simple opposition, authorial agency and contingency demonstrate an instability that depends on the author's design choices.

8 Conclusion

This study introduces *Orchid*, a new approach to facilitate collaborative narrative creation between human authors and LLMs. Orchid emphasizes a balance between author agency and contingency, enabling the user to actively participate in the narrative creation while enforcing the scope of the narrative defined by the author. *Orchid* was designed through a co-creation approach. We first interviewed authors of IDN narratives to understand their usage of existing tools and their needs for LLM-driven tools. Following these interviews, we designed our approach, which we implemented into a functional probe for authors to design their own narratives. We finally presented the probe to 12 participants, asking them to design an interactive narrative. Authors using *Orchid* valued the control over narrative structure and details, provided as natural language instructions. They also appreciated the system's ability to handle unexpected player actions. This study lays the groundwork for future LLM authoring tools that foster co-creation between player and author, allowing players to contribute to the narrative through their actions while preserving the author's intended vision.

References

[1] Espen Aarseth. 2012. A narrative theory of games. In *Proceedings of the International Conference on the Foundations of Digital Games* (Raleigh, North Carolina) *(FDG '12)*. Association for Computing Machinery, New York, NY, USA, 129–133. https://doi.org/10.1145/2282338.2282365

[2] Tarn Adams and Zach Adams. 2006. Dwarf Fortress. https://www.bay12games.com/dwarves/ Accessed: 2024-12-30.

[3] AIDungeon 2023. AI Dungeon. https://aidungeon.com/.

[4] Nader Akoury, Ronan Salz, and Mohit Iyyer. 2023. Towards Grounded Dialogue Generation in Video Game Environments. In *Creative AI Across Modalities Workshop, AAAI*. AAAI.

[5] Arwa I. Alhussain and Aqil M. Azmi. 2021. Automatic Story Generation: A Survey of Approaches. *ACM Comput. Surv.* 54, 5, Article 103 (may 2021), 38 pages. https://doi.org/10.1145/3453156

[6] Alberto Alvarez. 2023. ChatGPT as a Narrative Structure Interpreter. In *Interactive Storytelling: 16th International Conference on Interactive Digital Storytelling, ICIDS 2023, Kobe, Japan, November 11–15, 2023, Proceedings, Part II* (Kobe, Japan). Springer-Verlag, Berlin, Heidelberg, 113–121. https://doi.org/10.1007/978-3-031-47658-7_9

[7] Trevor Ashby, Braden K Webb, Gregory Knapp, Jackson Searle, and Nancy Fulda. 2023. Personalized Quest and Dialogue Generation in Role-Playing Games: A Knowledge Graph- and Language Model-based Approach. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 290, 20 pages. https://doi.org/10.1145/3544548.3581441

[8] Claus Atzenbeck, Sam Brooker, and Daniel Roßner. 2023. Storytelling Machines. In *Proceedings of the 6th Workshop on Human Factors in Hypertext* (Rome, Italy) *(HUMAN '23)*. Association for Computing Machinery, New York, NY, USA, Article 4, 9 pages. https://doi.org/10.1145/3603607.3613481

[9] Maren Awiszus, Frederik Schubert, and Bodo Rosenhahn. 2022. Wor (l) d-GAN: Toward Natural-Language-Based PCG in Minecraft. *IEEE Transactions on Games* 15, 2 (2022), 182–192.

[10] Ruth Aylett. 1999. Narrative in virtual environments-towards emergent narrative. In *Proceedings of the AAAI fall symposium on narrative intelligence*. USA, 83–86.

[11] Mark Bernstein. 1998. Patterns of hypertext. In *Proceedings of the ninth ACM conference on Hypertext and hypermedia: links, objects, time and space—structure in hypermedia systems: links, objects, time and space—structure in hypermedia systems*. 21–29.

[12] Jay David Bolter. 2001. *Writing Space: Computers, HyperText, and the Remediation of Print* (2nd ed.). L. Erlbaum Associates Inc., USA.

[13] Benjamin H. Bratton. 2016. *The Stack: On Software and Sovereignty*. The MIT Press. https://doi.org/10.7551/mitpress/9780262029575.001.0001

[14] Virginia Braun and Victoria Clarke. 2019. Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health* 11, 4 (2019), 589–597. https://doi.org/10.1080/2159676X.2019.1628806

[15] Sarah Anne Brown, Sharon Lynn Chu, Francis Quek, Pomaikai Canaday, Qing Li, Trystan Loustau, Sindy Wu, and Lina Zhang. 2019. Towards a Gesture-Based Story Authoring System: Design Implications from Feature Analysis of Iconic Gestures During Storytelling. In *Interactive Storytelling: 12th International Conference on Interactive Digital Storytelling, ICIDS 2019, Little Cottonwood Canyon, UT, USA, November 19–22, 2019, Proceedings* (Little Cottonwood Canyon, UT, USA). Springer-Verlag, Berlin, Heidelberg, 364–373. https://doi.org/10.1007/978-3-030-33894-7_38

[16] Chris Callison-Burch, Gaurav Singh Tomar, Lara J. Martin, Daphne Ippolito, Suma Bailis, and David Reitter. 2022. Dungeons and Dragons as a Dialog Challenge for Artificial Intelligence. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, Yoav Goldberg, Zornitsa Kozareva, and Yue Zhang (Eds.). Association for Computational Linguistics, Abu Dhabi, United Arab Emirates, 9379–9393. https://doi.org/10.18653/v1/2022.emnlp-main.637

[17] Rogelio Cardona-Rivera and Chris Martens. 2017. Procedural Narrative Generation. In *Proceedings of the Game Developers Conference*. North Carolina State University. https://www.gdcvault.com/play/1024143/Procedural-Narrative Accessed: 2024-12-30.

[18] Simon Chauvin, Guillaume Levieux, Jean-Yves Donnart, and Stéphane Natkin. 2015. Making sense of emergent narratives: An architecture supporting player-triggered narrative processes. *2015 IEEE Conference on Computational Intelligence and Games (CIG)* (2015), 91–98. https://api.semanticscholar.org/CorpusID:2127909

[19] Weize Chen, Yusheng Su, Jingwei Zuo, Cheng Yang, Chenfei Yuan, Chen Qian, Chi-Min Chan, Yujia Qin, Yaxi Lu, Ruobing Xie, et al. 2023. Agentverse: Facilitating multi-agent collaboration and exploring emergent behaviors in agents. *arXiv preprint arXiv:2308.10848* 2, 4 (2023), 6.

[20] John Joon Young Chung, Wooseok Kim, Kang Min Yoo, Hwaran Lee, Eytan Adar, and Minsuk Chang. 2022. TaleBrush: Sketching Stories with Generative Pretrained Language Models. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (<conf-loc>, <city>New Orleans</city>, <state>LA</state>, <country>USA</country>, </conf-loc>) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 209, 19 pages. https://doi.org/10.1145/3491102.3501819

[21] John Joon Young Chung and Max Kreminski. 2024. Patchview: LLM-powered Worldbuilding with Generative Dust and Magnet Visualization. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology* (Pittsburgh, PA, USA) *(UIST '24)*. Association for Computing Machinery, New York, NY, USA, Article 77, 19 pages. https://doi.org/10.1145/3654777.3676352

[22] DigixArt. 2021. Road 96. https://road96.com/

[23] DreamGen. 2025. AI Role-Play and Story Generator | DreamGen. https://dreamgen.com/ Accessed: 2025-01-02.

[24] Failbetter Games. 2009. Fallen London. https://www.fallenlondon.com/login Accessed: 2025-01-05.

[25] Jérémy Frey, Gilad Ostrin, May Grabli, and Jessica R. Cauchard. 2020. Physiologically Driven Storytelling: Concept and Software Tool. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376643

[26] Frontier Developments. 2021. Elite Dangerous. https://www.elitedangerous.com/

[27] Galactic Cafe. 2013. The Stanley Parable. https://www.stanleyparable.com/ Accessed: 2025-01-05.

[28] Daniel Green. 2018. Novella: An authoring tool for interactive storytelling in games. In *Interactive Storytelling: 11th International Conference on Interactive Digital Storytelling, ICIDS 2018, Dublin, Ireland, December 5–8, 2018, Proceedings 11*. Springer, 556–559.

[29] Daniel Green, Charlie Hargood, and Fred Charles. 2019. Novella 2.0: A Hypertextual Architecture for Interactive Narrative in Games. In *Proceedings of the 30th ACM Conference on Hypertext and Social Media* (Hof, Germany) *(HT '19)*. Association for Computing Machinery, New York, NY, USA, 77–86. https://doi.org/10.1145/3342220.3343655

[30] Daniel Green, Charlie Hargood, and Fred Charles. 2021. Use of Tools: UX Principles for Interactive Narrative Authoring Tools. *J. Comput. Cult. Herit.* 14, 3, Article 41 (jul 2021), 25 pages. https://doi.org/10.1145/3458769

[31] D. Green, Charlie Hargood, Fred Charles, and A. Jones. 2018. Novella: A Proposition for Game-Based Storytelling. In *Narrative and Hypertext 2018*. ACM. http://eprints.bournemouth.ac.uk/30690/ This conference paper is embargoed as it has not yet been presented at the 9 July 2018 Workshop on Narrative Systems to be held at Narrative and Hypertext 2018, Baltimore..

[32] Muhammad Usman Hadi, Qasem Al Tashi, Abbas Shah, Rizwan Qureshi, Amgad Muneer, Muhammad Irfan, Anas Zafar, Muhammad Bilal Shaikh, Naveed Akhtar, Jia Wu, et al. 2024. Large language models: a comprehensive survey of its applications, challenges, limitations, and future prospects. *Authorea Preprints* (2024).

[33] Jessica Hammer. 2007. Agency and authority in role-playing "texts". *A new literacies sampler* 29 (2007), 67–94.

[34] D Fox Harrell and Jichen Zhu. 2009. Agency Play: Dimensions of Agency for Interactive Narrative Design.. In *AAAI spring symposium: Intelligent narrative technologies II*. 44–52.

[35] N Katherine Hayles. 2001. The transformation of narrative and the materiality of hypertext. *Narrative* 9, 1 (2001), 21–39.

[36] Hello Games. 2016. No Man's Sky. https://www.nomanssky.com/

[37] https://ganelson.github.io/. 2024. Inform 7 | Inform is a natural-language-based programming language for writers of interactive fiction. https://ganelson.github.io/inform-website/.

[38] https://twinery.org/. 2024. Twine: An Open-Source Tool for Telling Interactive, Nonlinear Stories.

[39] https://www.articy.com/. 2024. Articy Draft: Interactive Storytelling and Game Design Tool. https://www.articy.com/en/.

[40] https://www.choiceofgames.com/. 2024. ChoiceScript: A Simple Language for Writing Multiple-Choice Interactive Novels. https://www.choiceofgames.com/make-your-own-games/choicescript-intro/.

[41] Chengpeng Hu, Yunlong Zhao, and Jialin Liu. 2024. Game Generation via Large Language Models. In *2024 IEEE Conference on Games (CoG)*. 1–4. https://doi.org/10.1109/CoG60054.2024.10645597

[42] Hilary Hutchinson, Wendy Mackay, Bo Westerlund, Benjamin B. Bederson, Allison Druin, Catherine Plaisant, Michel Beaudouin-Lafon, Stéphane Conversy, Helen Evans, Heiko Hansen, Nicolas Roussel, and Björn Eiderbäck. 2003. Technology probes: inspiring design for and with families. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Ft. Lauderdale, Florida, USA) *(CHI '03)*. Association for Computing Machinery, New York, NY, USA, 17–24. https://doi.org/10.1145/642611.642616

[43] Inklewriter Tool 2024. Inklewriter: Interactive Story Writing Tool. https://www.inklestudios.com/inklewriter/.

[44] Michael Joyce. 1987. Afternoon, a Story. Interactive Hypertext Fiction. https://directory.eliterature.org/individual-work/309

[45] Mubbasir Kapadia, Jessica Falk, Fabio Zünd, Marcel Marti, Robert W. Sumner, and Markus Gross. 2015. Computer-assisted authoring of interactive narratives. In *Proceedings of the 19th Symposium on Interactive 3D Graphics and Games* (San Francisco, California) *(i3D '15)*. Association for Computing Machinery, New York, NY, USA, 85–92. https://doi.org/10.1145/2699276.2699279

[46] Jack Kelly, Alex Calderwood, Noah Wardrip-Fruin, and Michael Mateas. 2023. There and back again: extracting formal domains for controllable neurosymbolic story authoring. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, Vol. 19. 64–74.

[47] Jack Kelly, Michael Mateas, and Noah Wardrip-Fruin. 2024. Paradise: An Experiment Extending the Ensemble Social Physics Engine with Language Models. In *Proceedings of the 19th International Conference on the Foundations of Digital Games* (Worcester, MA, USA) *(FDG '24)*. Association for Computing Machinery, New York, NY, USA, Article 64, 4 pages. https://doi.org/10.1145/3649921.3659841

[48] Taewook Kim, Hyomin Han, Eytan Adar, Matthew Kay, and John Joon Young Chung. 2024. Authors' Values and Attitudes Towards AI-bridged Scalable Personalization of Creative Language Arts. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 31, 16 pages. https://doi.org/10.1145/3613904.3642529

[49] 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

[50] Max Kreminski and Chris Martens. 2022. Unmet Creativity Support Needs in Computationally Supported Creative Writing. In *Proceedings of the First Workshop on Intelligent and Interactive Writing Assistants (In2Writing 2022)*, Ting-Hao 'Kenneth' Huang, Vipul Raheja, Dongyeop Kang, John Joon Young Chung, Daniel Gissin, Mina Lee, and Katy Ilonka Gero (Eds.). Association for Computational Linguistics, Dublin, Ireland, 74–82. https://doi.org/10.18653/v1/2022.in2writing-1.11

[51] Max Kreminski and Michael Mateas. 2021. A coauthorship-centric history of interactive emergent narrative. In *Interactive Storytelling: 14th International Conference on Interactive Digital Storytelling, ICIDS 2021, Tallinn, Estonia, December 7–10, 2021, Proceedings 14*. Springer, 222–235.

[52] Max Kreminski and Noah Wardrip-Fruin. 2018. Sketching a map of the storylets design space. In *Interactive Storytelling: 11th International Conference on Interactive Digital Storytelling, ICIDS 2018, Dublin, Ireland, December 5–8, 2018, Proceedings 11*. Springer, 160–164.

[53] Michael Kriegel, Ruth Aylett, Joao Dias, and Ana Paiva. 2007. An Authoring Tool for an Emergent Narrative Storytelling System.. In *AAAI fall symposium: intelligent narrative technologies*. 55–62.

[54] Vikram Kumaran, Dan Carpenter, Jonathan Rowe, Bradford Mott, and James Lester. 2023. End-to-end procedural level generation in educational games with natural language instruction. In *2023 IEEE Conference on Games (CoG)*. IEEE, 1–8.

[55] Vikram Kumaran, Jonathan Rowe, Bradford Mott, and James Lester. 2023. Scenecraft: Automating interactive narrative scene generation in digital games with large language models. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, Vol. 19. 86–96.

[56] Ben Kybartas and Rafael Bidarra. 2017. A Survey on Story Generation Techniques for Authoring Computational Narratives. *IEEE Transactions on Computational Intelligence and AI in Games* 9, 3 (2017), 239–253. https://doi.org/10.1109/TCIAIG.2016.2546063

[57] George P Landow. 1997. *Hypertext 2.0 (rev. ed.) the convergence of contemporary critical theory and technology.* Johns Hopkins University Press.

[58] Larian Studios. 2025. Baldur's Gate3. https://baldursgate3.game/ Accessed: 2025-01-05.

[59] Gaetan Lopez Latouche, Laurence Marcotte, and Ben Swanson. 2023. Generating video game scripts with style. In *Proceedings of the 5th Workshop on NLP for Conversational AI (NLP4ConvAI 2023)*. 129–139.

[60] Tomas Lawton, Kazjon Grace, and Francisco J Ibarrola. 2023. When is a Tool a Tool? User Perceptions of System Agency in Human–AI Co-Creative Drawing. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference* (Pittsburgh, PA, USA) *(DIS '23)*. Association for Computing Machinery, New York, NY, USA, 1978–1996. https://doi.org/10.1145/3563657.3595977

[61] Mina Lee, Katy Ilonka Gero, John Joon Young Chung, Simon Buckingham Shum, Vipul Raheja, Hua Shen, Subhashini Venugopalan, Thiemo Wambsganss, David Zhou, Emad A Alghamdi, et al. 2024. A Design Space for Intelligent and Interactive Writing Assistants. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–35.

[62] Carlos León and Pablo Gervás. 2014. Creativity in Story Generation From the Ground Up: Non-deterministic Simulation driven by Narrative. In *International Conference on Innovative Computing and Cloud Computing*. https://api.semanticscholar.org/CorpusID:12272687

[63] Boyang Li and Mark Riedl. 2015. Scheherazade: Crowd-powered interactive narrative generation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 29.

[64] Jialu Li, Yuanzhen Li, Neal Wadhwa, Yael Pritch, David E Jacobs, Michael Rubinstein, Mohit Bansal, and Nataniel Ruiz. 2024. Unbounded: A Generative Infinite Game of Character Life Simulation. *arXiv preprint arXiv:2410.18975* (2024).

[65] Hui Liang, Jian Chang, Ismail K Kazmi, Jian J Zhang, and Peifeng Jiao. 2017. Hand gesture-based interactive puppetry system to assist storytelling for children. *The Visual Computer* 33 (2017), 517–531.

[66] Max Loh. 2023. AI Roguelite. https://store.steampowered.com/app/1889620/AI_Roguelite/?rdt_cid=4771370055936078709 Accessed: 2024-12-22.

[67] Sandy Louchart. 2007. *Emergent Narrative-towards a narrative theory of Virtual Reality.* University of Salford (United Kingdom).

[68] Steven Lumpkin. 2019. Procedural Storytelling in Dungeons & Dragons. In *Procedural Storytelling in Game Design*. AK Peters/CRC Press, 257–268.

[69] Cat Manning. 2019. Tarot as Procedural Storytelling. In *Procedural Storytelling in Game Design*. AK Peters/CRC Press, 339–350.

[70] Lev Manovich. 2002. *The language of new media.* MIT press.

[71] Stacey Mason, Ceri Stagg, and Noah Wardrip-Fruin. 2019. Lume: a system for procedural story generation. In *Proceedings of the 14th International Conference on the Foundations of Digital Games*. 1–9.

[72] Eric Murnane. 2018. *Emergent Narrative: Stories of Play, Playing with Stories.* Doctor of Philosophy (Ph.D.). University of Central Florida, College of Arts and Humanities. http://purl.fcla.edu/fcla/etd/CFE0007061 If this is your thesis or dissertation, and want to learn how to access it or for more information about readership statistics, contact us at STARS@ucf.edu.

[73] Mark Nelson and Michael Mateas. 2021. Search-Based Drama Management in the Interactive Fiction Anchorhead. *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment* 1, 1 (Sep. 2021), 99–104. https://doi.org/10.1609/aiide.v1i1.18723

[74] Britney Ngaw, Grishma Jena, João Sedoc, and Aline Normoyle. 2024. Towards Authoring Open-Ended Behaviors for Narrative Puzzle Games with Large Language Model Support. In *Proceedings of the 19th International Conference on the Foundations of Digital Games* (Worcester, MA, USA) *(FDG '24)*. Association for Computing Machinery, New York, NY, USA, Article 43, 4 pages. https://doi.org/10.1145/3649921.3656975

[75] Jeongyoon Park, Jumin Shin, Gayeon Kim, and Byung-Chull Bae. 2023. Designing a Language Model-Based Authoring Tool Prototype for Interactive Storytelling. In *Interactive Storytelling*, Lissa Holloway-Attaway and John T. Murray (Eds.). Springer Nature Switzerland, Cham, 239–245.

[76] Joon Sung Park, Joseph O'Brien, Carrie Jun Cai, Meredith Ringel Morris, Percy Liang, and Michael S. Bernstein. 2023. Generative Agents: Interactive Simulacra of Human Behavior. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology* (San Francisco, CA, USA) *(UIST '23)*. Association for Computing Machinery, New York, NY, USA, Article 2, 22 pages. https://doi.org/10.1145/3586183.3606763

[77] Alejandro Ramirez and Vadim Bulitko. 2014. Automated planning and player modeling for interactive storytelling. *IEEE Transactions on Computational Intelligence and AI in Games* 7, 4 (2014), 375–386.

[78] Mark Riedl. 2021. An Introduction to AI Story Generation. *Medium* (January 2021). https://mark-riedl.medium.com/an-introduction-to-ai-story-generation-7f99a450f615 Accessed: 2024-12-30.

[79] Mark O Riedl and R Michael Young. 2006. Story planning as exploratory creativity: Techniques for expanding the narrative search space. *New Generation Computing* 24, 3 (2006), 303–323. https://doi.org/10.1007/BF03037337

[80] Daniel Roßner, Claus Atzenbeck, and Sam Brooker. 2023. SPORE: A Storybreaking Machine. In *Proceedings of the 34th ACM Conference on Hypertext and Social Media*. 1–6.

[81] Marie-Laure Ryan. 2004. Will New Media Produce New Narratives? Marie-Laure Ryan. *Narrative across media: The languages of storytelling* (2004), 337.

[82] Marie-Laure Ryan, James Ruppert, and John W Bernet. 2004. *Narrative across media: The languages of storytelling*. U of Nebraska Press.

[83] sagarpg.io 2023. Saga: AI Text RPG. https://sagarpg.io/.

[84] Markus Schlosser. 2019. Agency. In *The Stanford Encyclopedia of Philosophy* (Winter 2019 ed.), Edward N. Zalta (Ed.). Metaphysics Research Lab, Stanford University.

[85] Yotam Shibolet, Noam Knoller, and Hartmut Koenitz. 2018. A Framework for Classifying and Describing Authoring Tools for Interactive Digital Narrative. In *Interactive Storytelling*, Rebecca Rouse, Hartmut Koenitz, and Mads Haahr (Eds.). Springer International Publishing, Cham, 523–533.

[86] Emily Short. 2019. Creating Tools for Procedural Storytelling. In *Procedural Storytelling in Game Design*, Tanya Kremers (Ed.). A K Peters/CRC Press, 15.

[87] Tanya X Short and Tarn Adams. 2019. *Procedural storytelling in game design.* Crc Press.

[88] Joan Soler-Adillon. 2019. The open, the closed and the emergent: Theorizing emergence for videogame studies. *Game Studies* 19, 2 (2019).

[89] Callum Spawforth and David Millard. 2017. A framework for multi-participant narratives based on multiplayer game interactions. In *10th International Conference on Interactive Digital Storytelling (14/11/17 - 17/11/17)*. 150–162. https://eprints.soton.ac.uk/414445/

[90] StoryNexus 2024. StoryNexus: A Platform for Creating Interactive Worlds and Stories. https://www.ifwiki.org/StoryNexus.

[91] StoryScape. 2023. StoryScape AI. https://storyscapeai.app/ Accessed: 2025-01-02.

[92] StoryScapeAI 2023. StoryScape AI. https://storyscapeai.app/.

[93] Storyspace 2024. Storyspace: Hypertext Writing Environment. https://www.eastgate.com/storyspace/.

[94] Bethesda Game Studios. 2011. The Elder Scrolls V: Skyrim. https://bethesda.net/en/search

[95] Anne Sullivan, Mirjam Palosaari Eladhari, and Michael Cook. 2018. Tarot-based narrative generation. In *Proceedings of the 13th International Conference on the Foundations of Digital Games*. 1–7.

[96] Yuqian Sun, Zhouyi Li, Ke Fang, Chang Hee Lee, and Ali Asadipour. 2023. Language as reality: a co-creative storytelling game experience in 1001 nights using generative AI. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, Vol. 19. 425–434.

[97] Yuqian Sun, Hanyi Wang, Pok Man Chan, Morteza Tabibi, Yan Zhang, Huan Lu, Yuheng Chen, Chang Hee Lee, and Ali Asadipour. 2023. FictionalWorlds, Real Connections: Developing Community Storytelling Social Chatbots through LLMs. In *International Conference on Interactive Media, Smart Systems and Emerging Technologies (IMET)*, Nuria Pelechano, Fotis Liarokapis, Damien Rohmer, and Ali Asadipour (Eds.). The Eurographics Association. https://doi.org/10.2312/imet.20231261

[98] Nicolas Szilas, Ruud de Jong, and Mariët Theune. 2020. ExpReal: a Writing Language and System for Authoring Texts in Interactive Narrative. In *Proceedings of the 15th International Conference on the Foundations of Digital Games* (Bugibba, Malta) *(FDG '20)*. Association for Computing Machinery, New York, NY, USA, Article 12, 4 pages. https://doi.org/10.1145/3402942.3402949

[99] Talefy. [n. d.]. Random Story Mode | Talefy - AI Story Generator. https://talefy.ai/random-story Accessed: 2025-01-02.

[100] Pittawat Taveekitworachai, Febri Abdullah, Mustafa Can Gursesli, Mury F Dewantoro, Siyuan Chen, Antonio Lanata, Andrea Guazzini, and Ruck Thawonmas. 2023. What is waiting for us at the end? inherent biases of game story endings in large language models. In *International Conference on Interactive Digital Storytelling*. Springer, 274–284.

[101] Graham Todd, Sam Earle, Muhammad Umair Nasir, Michael Cerny Green, and Julian Togelius. 2023. Level Generation Through Large Language Models. In *Proceedings of the 18th International Conference on the Foundations of Digital Games* (Lisbon, Portugal) *(FDG '23)*. Association for Computing Machinery, New York, NY, USA, Article 70, 8 pages. https://doi.org/10.1145/3582437.3587211

[102] Mike Treanor, Ben Samuel, and Mark J. Nelson. 2024. Prototyping Slice of Life: Social Physics with Symbolically Grounded LLM-based Generative Dialogue. In *Proceedings of the 19th International Conference on the Foundations of Digital Games* (Worcester, MA, USA) *(FDG '24)*. Association for Computing Machinery, New York, NY, USA, Article 56, 4 pages. https://doi.org/10.1145/3649921.3656988

[103] Nicholas Sloss Treynor and Joshua McCoy. 2024. College Ruled: A Pathfinding Approach to Generative Storytelling. In *Proceedings of the 19th International Conference on the Foundations of Digital Games* (Worcester, MA, USA) *(FDG '24)*. Association for Computing Machinery, New York, NY, USA, Article 15, 10 pages. https://doi.org/10.1145/3649921.3649994

[104] Undum 2024. Undum: A Framework for Interactive Fiction and Narrative Games. https://idmillington.github.io/undum/.

[105] Ralph C.S. Walker. 1998. Contingency. (1998). https://doi.org/10.4324/9780415249126-N008-1

[106] Guanzhi Wang, Yuqi Xie, Yunfan Jiang, Ajay Mandlekar, Chaowei Xiao, Yuke Zhu, Linxi Fan, and Anima Anandkumar. [n. d.]. Voyager: An Open-Ended Embodied Agent with Large Language Models. 2023. *Comment: Project website and open-source codebase: https://voyager. minedojo. org/Cited on* ([n. d.]), 33.

[107] Timothy S Wang and Andrew S Gordon. 2023. Playing story creation games with large language models: Experiments with GPT-3.5. In *International Conference on Interactive Digital Storytelling*. Springer, 297–305.

[108] Yi Wang, Qian Zhou, and David Ledo. 2024. StoryVerse: Towards Co-authoring Dynamic Plot with LLM-based Character Simulation via Narrative Planning. In *Proceedings of the 19th International Conference on the Foundations of Digital Games* (Worcester, MA, USA) *(FDG '24)*. Association for Computing Machinery, New York, NY, USA, Article 55, 4 pages. https://doi.org/10.1145/3649921.3656987

[109] Stephen Ware, R. Young, Christian Stith, and Phillip Wright. 2015. Interactive Narrative Planning in The Best Laid Plans. *Proceedings of the AAAI Conference on Artificial Intelligence* 29, 1 (Mar. 2015). https://doi.org/10.1609/aaai.v29i1.9780

[110] Stephen G. Ware and R. Michael Young. 2014. Glaive: a state-space narrative planner supporting intentionality and conflict. In *Proceedings of the Tenth AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment* (Raleigh, NC, USA) *(AIIDE'14)*. AAAI Press, 80–86.

[111] Wikipedia contributors. 2023. Zork. https://en.wikipedia.org/wiki/Zork Accessed: 2025-01-05.

[112] Zhen Wu, Ruoyu Wen, Marcel Zaes Sagesser, Sum Yi Ma, Jiashuo Xian, Wenda Wu, and Tristan Braud. 2024. I Light U Up: Exploring a New Emergent Narrative Paradigm through Physical Data Participation in AI Generative Experiences. In *SIGGRAPH Asia 2024 Art Papers (SA '24)*. Association for Computing Machinery, New York, NY, USA, Article 22, 9 pages. https://doi.org/10.1145/3680530.3695446

[113] Qing Ru Yong and Alex Mitchell. 2023. From Playing the Story to Gaming the System: Repeat Experiences of a Large Language Model-Based Interactive Story. In *Interactive Storytelling*, Lissa Holloway-Attaway and John T. Murray (Eds.). Springer Nature Switzerland, Cham, 395–409.

[114] Xiao You, Pittawat Taveekitworachai, Siyuan Chen, Mustafa Can Gursesli, Xiaoxu Li, Yi Xia, and Ruck Thawonmas. 2024. Dungeons, Dragons, and Emotions: A Preliminary Study of Player Sentiment in LLM-driven TTRPGs. In *Proceedings of the 19th International Conference on the Foundations of Digital Games* (Worcester, MA, USA) *(FDG '24)*. Association for Computing Machinery, New York, NY, USA, Article 59, 4 pages. https://doi.org/10.1145/3649921.3656991

[115] Ann Yuan, Andy Coenen, Emily Reif, and Daphne Ippolito. 2022. Wordcraft: Story Writing With Large Language Models. In *Proceedings of the 27th International Conference on Intelligent User Interfaces* (Helsinki, Finland) *(IUI '22)*. Association for Computing Machinery, New York, NY, USA, 841–852. https://doi.org/10.1145/3490099.3511105

[116] ZA/UM. 2019. Disco Elysium. https://discoelysium.com/zh Accessed: 2025-01-05.