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

Zhen Wu*
Division of Integrative System and Design
Hong Kong University of Science and Technology
HongKong, Hong Kong
zwuch@connect.ust.hk

Serkan Kumyol
Department of Computer Science and Engineering
Hong Kong University of Science and Technology
Hong Kong, Hong Kong
skumyol@connect.ust.hk

Shing Yin Wong
Department of Mathematics
Hong Kong University of Science and Technology
Hong Kong, Hong Kong
warrenwong641@gmail.com

Xiaozhu Hu
Division of Integrative Systems and Design
Hong Kong University of Science and Technology
Hong Kong, Hong Kong
xhubk@connect.ust.hk

Xin Tong
Thrust of Computational Media and Arts
Hong Kong University of Science and Technology (Guangzhou)
Guangzhou, China
xint@hkust-gz.edu.cn

Tristan Braud†
Division of Integrative Systems and Design
Hong Kong University of Science and Technology
Hong Kong, Hong Kong
braudt@ust.hk

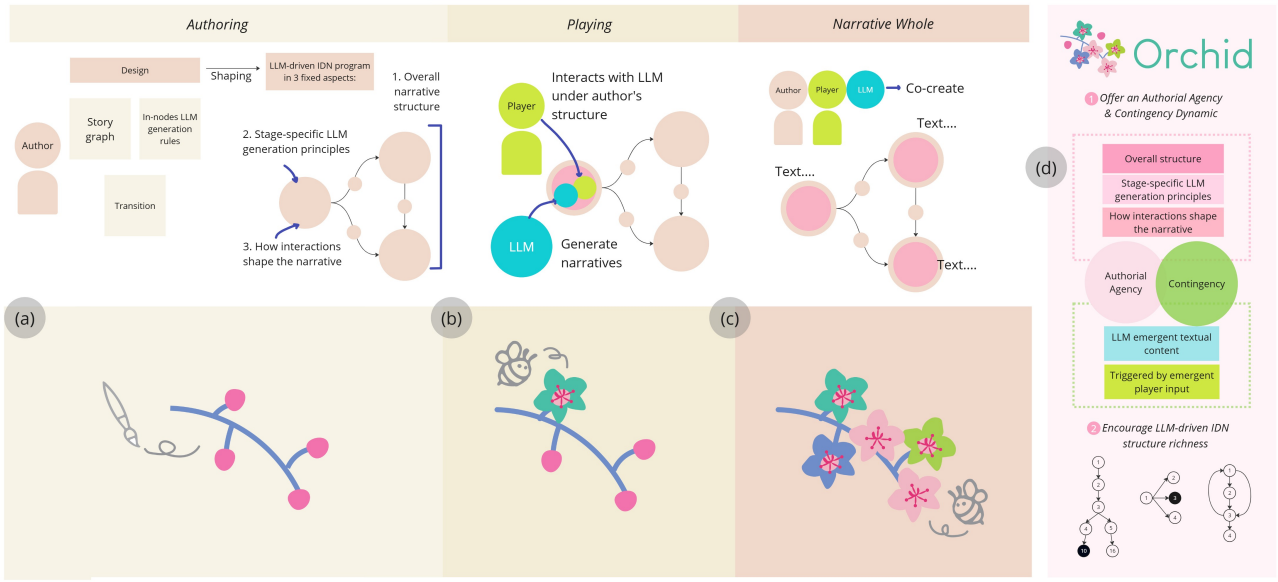

Figure 1: The creative approach, Orchid, introduces human-AI co-creation in three stages. LLMs are represented as different buds, with players' prompts serving as the catalyst that triggers narratives to bloom from within. (a) Authoring. The author creates the structure of the generative IDN, especially providing 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.

*Main contributor
†Corresponding author

C&C '25, Virtual, United Kingdom

ACM ISBN 979-8-4007-1289-0/25/06
https://doi.org/10.1145/3698061.3726906

Abstract

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

• **Computing methodologies** → **Natural language generation**; • **Human-centered computing** → **User studies**; **Interaction devices**; • **Applied computing** → **Computer games**.

Keywords

LLMs, Interactive Digital Narrative, authorial agency and machine contingency, technology probe, emergent narrative

ACM Reference Format:

Zhen Wu, Serkan Kumyol, Shing Yin Wong, Xiaozhu Hu, Xin Tong, and Tristan Braud. 2025. Orchid: A Creative Approach for Authoring LLM-Driven Interactive Narratives. In *Creativity and Cognition (C&C '25), June 23–25, 2025, Virtual, United Kingdom.* ACM, New York, NY, USA, 18 pages. https://doi.org/10.1145/3698061.3726906

1 Introduction

Interactive digital narratives (IDNs) are stories where users influence outcomes through their choices and actions [59]. To keep the narrative engaging, authors define possible user choices and their pre-defined responses, balancing storytelling with interactivity [58]. 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 [79, 80]. 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 [50] 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 utilised. Works like *1001 Nights* [68] 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' homogenised lexical style, often resembling interactions with a corporate chatbot.

The uniqueness of LLMs for IDNs requires the development of new creative approaches, abstracting the authoring process rather than focusing on specific systems. Introducing a creative approach could encourage the emergence of more works. This would enhance our understanding of how to support those creations, as exemplified by works such as Storylet[43] and Novella[22]. Creative approaches are also embodied in authoring tools, such as the 'natural language programming' in Inform 7 [30] and the 'hypertext linking' in Twine [31]. Creative approach in these tools contributes 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 [61], limited research has explored this aspect in LLM-driven IDNs, with a major focus on game creation. The current application of LLMs in IDNs reveals a need for a new creative approach to introduce new human-AI co-creation dynamics and encourage richer IDN structures in this field.

In this paper, we introduce *Orchid*, a new creative approach for LLM-driven IDN authoring, emphasising 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* raises two primary hypotheses: (1) LLM-based IDNs created with Orchid facilitate an open-ended collaboration between authors, machine creativity, and player interactivity to generate emergent textual content; (2) Orchid supports the creation of richer narrative structures beyond traditional linear formats commonly found in many LLM-related experiences, such as in AI Dungeon.

We designed and developed a technology probe as a computational realisation of *Orchid* for investigating users' perceptions and uncovering their needs [35]. The probe embodies *Orchid*'s core concepts into simple functions, encouraging exploratory use rather than usability. Then, we conducted a study involving twelve IDN authors who interacted with the technology probe to create complete pieces of IDNs using *Orchid*. Through the design activities, participants created novel IDN structures, 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 summarised 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.

2 Background and Related Work

This work takes place at the intersection between LLMs, IDNs and authoring processes. After reviewing the primary works related to LLMs usage in narrative creation, we present the most prominent works on IDN authoring and the main creativity support methods and tools. We finally discuss how authorial agency and machine contingency have been considered in the literature.

2.1 LLM in Narrative Creation and IDNs

Large Language Models(LLMs) can help rapidly generate creative content in natural language [25], facilitating the transition from narrative ideas to output [5, 54]. As such, diverse creative support tools have been developed for LLM-assisted authoring [83]. Commercial applications such as Talefy [71], Dream Gen [20], and Storyscape [66], mainly focus on the complete narrative construction process, guiding users to create stories through defining different settings [48, 83]. Research work also presents novel approaches to facilitating creativity, such as using mind maps to guide Chat-GPT's generation [7, 57], employing metaphors such as sketching lines to automatically generate narrative arcs [16], and using magnets and dust to visually assist LLM-driven word-building [17].

IDNs combine storytelling with user interactivity through digital media, allowing users to shape the narrative they finally receive [28, 58, 59]. While traditional IDNs require authors to script all possible content and rules, which can be limiting, error-prone, and time-consuming, applying LLMs can simplify and enhance flexibility in the process [70], surpassing hard-coded limitations [49]. This is addressed in works including using LLMs in making plans for the event [55], guiding NPCs' dialogues without human intervention [4, 6, 13, 38, 46, 53, 75, 82], and generating game levels governing complex functional constraints and spatial relationships [8, 34, 45, 74]. Researchers have also explored interactive possibilities in LLM-based IDN applications. For instance, Sun et al. used LLM-driven chatbots as fictional game characters to interact with real people on social media [69]. Games like *1001 Nights* encourage players to input their own stories and generate weapons for battles, and *I Light U Up* uses light to communicate with virtual beings within a narrative [80]. In *Turlock Holmes*, AI plays the role of a host, allowing players to inquire about clues based on a given context [73]. Although these LLM implementations are well-suited to their specific contexts, the relationship between player interactions and the narrative is loose, resembling a separation of player interactions from narrative generation. For instance, in 1001 Nights, the LLM is only used to generate items rather than directly influencing the progression of the narrative. IDNs in the mainstream market, such as *AI Dungeon*, which uses LLMs to generate narratives based on player input, and AI Roguelite [50], which creates dynamic quests and dialogues, provide full narrative generation. However, the outcome they generate often faces disruptions and dispatching from the initial settings, as seen in Steam player comments: "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 summary, while LLMs hold significant potential for facilitating IDN creation, there remains a need for a more cohesive and comprehensive IDN creation approach that fully leverages their capabilities in combining interactivity with narrative.

2.2 IDN and authoring

Creative support and authoring methods are important topics for IDN researchers [24]. These are reflected in two aspects: developing authoring tools that provide a complete workspace for creating IDNs [61] and proposing IDN creative approaches that are open to any computational method to realise it in IDN. In this paper, we refer to creative approaches as the description of authoring methods that help authors analyse and understand how to create a specific type of IDN [1, 10]. For example, Novella [22, 23] offers a game-centric IDN creation model focused on story objects and narrative states tracked through variables and flow graphs. Storylet [43] employs modular chunks to define scenes, allowing authors to set transition conditions and variables, while players choose their paths. Authoring tools and approaches for IDNs are often genre-specific. For example, script-based branching tools like Inklewriter [36], Articy [32], and Storyspace [67] create narrative branches, while procedural narrative generation involves dynamic interactions and real-time user responses, granting players significant decision-making power in games like Dwarf Fortress [2], No Man's Sky [29], Road 96 [19], and Elite Dangerous [21]. In contrast, emergent narrative genres produce outputs such as game logs and NPC behaviours, relying on complex definitions emphasising player participation, and high-level plot development, exemplified by FEARNOT! [44], which defines goals, conditions, and emotional reactions.

While recent attention to LLMs in IDNs has primarily focused on content creation, there is a lack of research into authoring approaches supporting new types of creativity in LLM-driven IDNs. This gap has inspired us to explore the authoring aspect and propose a creative approach for the integration of LLMs and IDNs.

2.3 Nodes and Graphs

Nodes and graphs are popular for defining stories, as seen in various IDN creation methods. Storygraphs are frequently used in script-based branching, allowing authors to create blueprints and write detailed content within nodes. Tools like Twine [31] and ChoiceScript [33] visualise relationships between different nodes through built-in syntax, while Undum/Raconteur [77] and StoryNexus [65] enable branching based on game state variables. In LLM-assisted narrative generation, works like VISAR [84] and XCreation [81] leverage nodes and graphs to represent story elements and their interconnections, enabling an intuitive understanding of narrative structure and guidance for LLM generation. Graphs and nodes are also efficient in a broader generative context. For instance, ComfyUI provides pipelines for open-weight image generation models using a graph/nodes/flowchart-based interface [2]. In the LLM-IDN

[1] https://store.steampowered.com/agecheck/app/1519310/
[2] https://github.com/comfyanonymous/ComfyUI

context, nodes and graphs present two key benefits. Firstly, nodes and graphs suit the creation of diverse IDN forms, allowing authors to generate different narrative structures by focusing on narrative events and their transitions. On the other hand, nodes and graphs create a visual structure that allows users to intuitively understand the story's logic, enhancing comprehension of the generation process and providing better control and transparency. This simplifies the understanding of text generation, boosting creative efficiency and reducing the user's cognitive load.

2.4 Authorial Agency and Machine Contingency

Agency refers to exercising or manifesting the capacity to act [60]. In IDNs, the authors' agency emphasises their ability to shape the storyline [26]. In conventional IDNs, balancing game interactivity with maintaining narrative consistency and author agency is considered challenging, especially when considering complex story structures [26, 27, 42, 52]. AI has an open-ended nature that shares the creative agency in humans' authoring experience, and the outcome is a co-creation between the user's inputs and AI's creativity [56]. The attention to agency in the author-LLMs relationship can be associated with a broader debate in Human-AI Co-creation. For instance, in the work introducing AI to facilitate writing activities, particularly in machine-in-the-loop creative writing, it helps to spark ideas, as mentioned by [18, 64, 78]. This was also addressed in drawing [47] and language art [39].

While LLMs have proven popular in narrative generation, it is crucial to reflect on the authority humans maintain in the context of "accelerationism" [11] 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 [37, 40]. *College Ruled* employs causality weighting to predict and select plot fragments generated by LLMs, allowing for finer authorial control throughout the storytelling process [76]. Wang et al. proposed a creative workflow for IDN that enables writers to define high-level plot outlines that are later transformed into concrete character action sequences via an LLM-based narrative planning process [79]. However, gaps remain in the singular interactivity that those methods can offer, questioning the applicability to complex player interactions.

Contingency in this work refers to the randomness and unpredictability in narrative outcomes. Contingency plays a crucial role in emergent narrative [14, 44]. Unpredictability in emergent narratives creates a fragile yet engaging experience [9], and combines pre-authored structures with user freedom, allowing meaningful player interaction within a story [51, 62, 63]. Introducing LLMs adds a machine element to 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 results in emergent outputs that are out of the author's expectation, *i.e. agents killing each other for resources* [15]. 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 [41, 72].

The relationship between authorial agency and contingency is an important topic, both in IDN authoring and mixed-initiative machine creativity in interactive systems, which our proposed creative approach will focus on.

2.5 Gaps and Opportunities

Through the above related work, we identify three opportunities in the creation of LLM-driven IDNs: **1) The current authoring approach for LLM-powered narratives often restricts the author's role to the initial setup.** For instance, in *AI Dungeon*, the author provides the setting, while the narrative primarily evolves through player prompts, diminishing authorial presence in extended interactions (see Figure 2, lower right). **2) The underutilization of LLMs' real-time narrative creativity**. Looking at subsection 2.1, the current focus of LLM generation is on making plans and guiding NPC dialogues rather than narrating a complete and coherent story. There is also limited creative support for achieving this. **3)The interaction between players and LLMs is loose or poorly designed.** subsection 2.1 points out that some player interactivity does not drive LLM generation but instead governs game mechanics. In cases where player interactivity does drive generation, the input is limited to natural language, which is quite narrow, reducing the potential for other meaningful engagements.

3 The Orchid Creative Approach

This section details the *Orchid* creative approach. After defining the core concepts of authorial agency and machine contingency, we describe the main aspects of the approach and how they address the identified opportunities(subsection 2.5) before explaining their relation to the orchid metaphor.

3.1 Authorial Agency and Machine Contingency in Orchid

Orchid provides a new approach for integrating LLM's narrative power, emphasising two key concepts: *authorial agency* and *machine contingency*. *Authorial Agency* highlights the author's control over the narrative structure and the rules governing player interactions. Existing LLM-authored approaches focus on the initial setting, diminishing author agency as LLM generation and player inputs take over the narration. *Contingency* in *Orchid* arises from the contributions of the LLM, triggered by player interactions. In traditional IDNs, emergent outcomes are primarily driven by player actions or designed around those actions. LLMs add a machine element to this contingency, further amplifying unexpectedness and resulting in narratives that can differ significantly with each generation. For the creation under the *Orchid* framework, authors consider both the emergent behaviours of players and the openness of LLM-generated text.

3.2 Design Rationale

We designed *Orchid* as a new creative approach to author LLM-driven IDNs as shown in Figure 2. *Orchid* defines an IDN as a directed tree, where each node corresponds to a specific stage in the narrative. The LLM combines the content of each node, defined by the author, with the player's input to create the narrative. Similar to *Storylet* and *Novella*, *Orchid* serves as an abstraction for the IDN

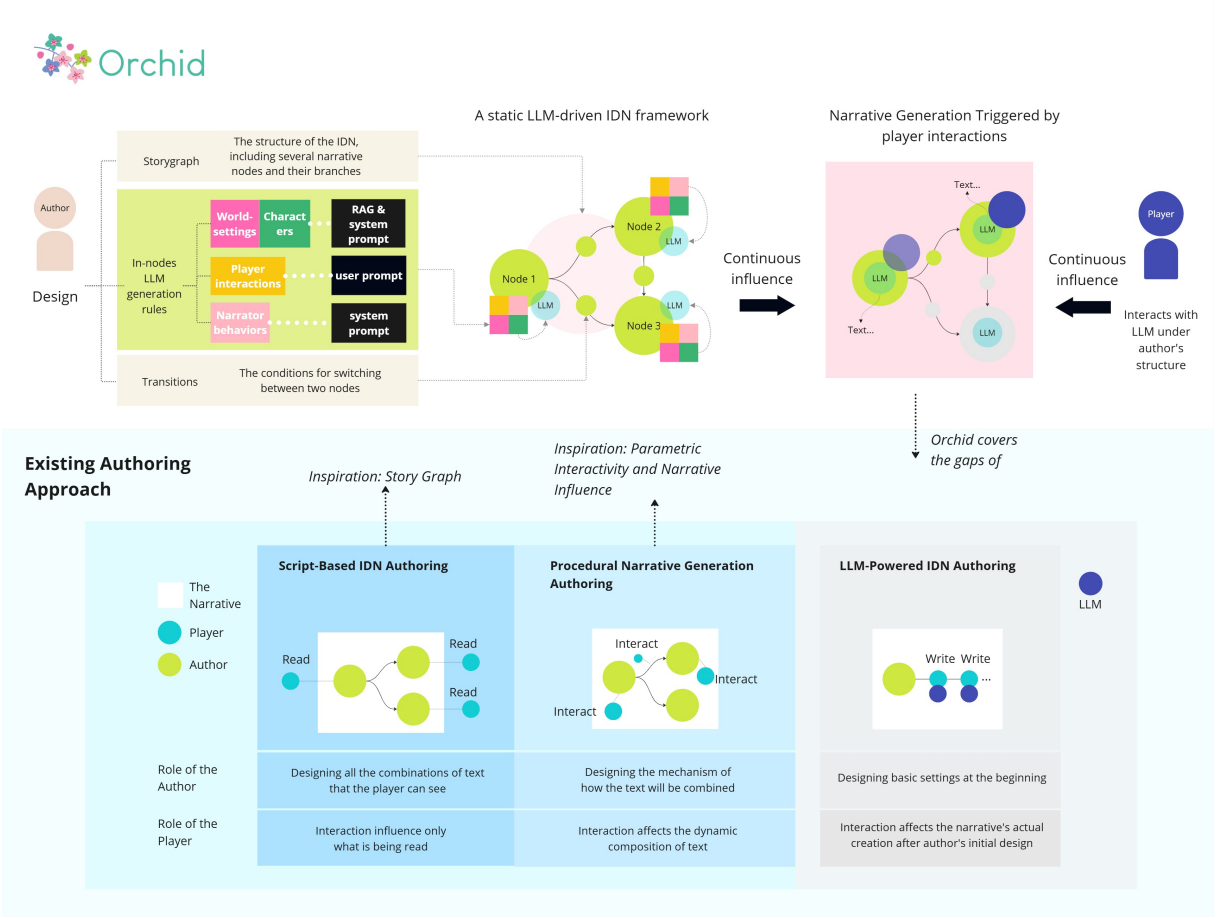

Figure 2: Upper Part: Orchid is a creative approach for authoring LLM-powered IDNs. By creating a story graph, defining transitions, and establishing in-node LLM generation rules, authors develop a static framework. When delivered to the player, their input triggers LLM text generation within each node. Lower Part: Orchid is inspired by script-based IDN authoring and procedural narrative generation, addressing gaps observed in existing LLM-powered IDN authoring.

authoring process, rather than targeting a specific system. The creative approach revolves around the following three concepts:

- Story graph: The overall narrative structure, organizing the relationship between different narrative stages as nodes and their sequences that guide the narrative's development.
- Transitions: The conditions for the narrative to switch from the current node to another adjacent node.
- In-nodes LLM generation rules: The LLM-related story generation principles for each narrative stage (node), including 1) world settings, characters, 2) narrator behaviours, and 3) player interactions, that underpin the narrative generation within each stage.

3.2.1 Story graph and transitions. The **story graph** and **transition** address the first gaps. They provide authors with greater control over the narrative progression. Contrary to tools like *AI Dungeon* that focus on the initial setting, authors define a blueprint of the entire narrative structure, linking characters, places, and events. The players engage with the narrative without altering its fundamental structure. (see Figure 2, lower left). As such, *Orchid* maintains significant author agency across the entire experience.

The **story graph** structure is commonly used in traditional IDNs. Authors define the main narrative stages as nodes and link them with transitions. This approach allows the development of multiple branching narratives, driven by the player's choices. *Orchid* adapts this approach to LLM-driven narrative for authors to maintain more precise control over the AI generation process. Authors define nodes and transitions in written text. The LLM then generates the details in each step, dynamically adapting to the players' input while enforcing the narrative structure.

In traditional IDNs, **transitions** depend on previous developments in the narrative. *Orchid* adopts this approach for LLM-generated narratives. The "agent memory" of the LLM is evaluated at each turn of player input. Transitions are triggered when the cumulative narrative content in the current node meets the transition conditions of a nearby node. Upon transition, LLM's settings are updated, and the player interacts in the next node, with the

narrative content constrained by the author's settings defined in that specific node.

3.2.2 In-nodes LLM generation rules. **In-nodes LLM generation rules** aim to address the second gap. We aim to lower the barrier and scaffold the process for defining LLM generation rules that are tailored to IDN design. Authors design the specific generation rules within each node, in three categories, drawing inspiration from the concept of story breaking [57]: **world-settings and characters**, **narrator behaviour**, and **player interactions**.

World-settings and characters encourages authors to create detailed descriptions of the IDN's background information, forming the narrative's foundation that will serve as the knowledge base for the LLM. LLMs often present a limited context size. World-setting and characters are thus relying on Retrieval Augmented Generation (RAG) [3], which retrieves relevant information from external sources and integrates it with user prompts through text embeddings and contextualization.

Narrator behaviour encompasses specific instructions for the LLM narration. It includes content such as narrative perspective and style, and specific details that the LLM needs to mention within the node. In LLMs, this relates to the "system prompt", which defines how the AI agent should behave. In the context of IDN, the system prompts guide the AI narrator, instructing the LLM on what content to present and how to present it.

Finally, **player interactions** address the lack of structured designs for how player interactivity influences LLM generations(the third gap). Authors describe the types of inputs players can provide and how they affect narrative progression, mapping player actions to narrative parameters. In LLMs, this refers to the "user prompt." *Orchid* suggests a parametric authoring method that first describes the interactive input in terms of its typology and range, then explains how this input influences the narrative. For example, players may choose a number from 1 to 10, with higher numbers leading to a more positive narrative outcome generated by the LLM. This method is inspired by procedural narrative generation mentioned in subsection 2.2, which typically involves a parametric generation framework (see Figure 2, lower middle).

3.3 Orchid Metaphor in the author-AI-player Co-Creation Dynamic

The proposed creative approach uses the metaphor of the orchid flower's growth and blooming process to scaffold the co-creation dynamic between authors, players, and LLMs. Orchids branch from the stem, with each branch featuring several buds that bloom into flowers. Similarly, *Orchid* defines the narrative using "branches", "buds", and "flowers". A bud embodies uncertainty while blooming signifies a definite outcome. The author sets the "stem" (narrative branches) and "buds" (LLM-powered narrative nodes) in an unbloomed state, controlling the fundamental settings and instructions (Figure 1.(a)). However, the final narrative (bloomed flower) depends on the LLM and the player's input. The player, represented as a "bee" in Figure 1. (b), influences the flowering process, affecting the likelihood and appearance of blooming. We can use this to imagine that the player's actions drive the narratives from the node through interactive inputs, independent of the author's or AI's creativity. These interactions influence the "blooming of the nodes," resulting in a unique narrative (Figure 1.(c)). We can imagine this as, in the branch of an orchid, some buds bloom, while some buds do not, and the final flower shape comes out from the blooming buds.

[3]https://python.langchain.com/docs/concepts/rag/

4 Probe Design and Development

We aim to study how authors navigate the new co-creation dynamic offered by *Orchid*, specifically how they create IDNs using *Orchid* and their needs related to *authorial agency* and *contingency*. Exploring these concepts is challenging without a system incorporating LLM generation capabilities. We thus implement Orchid into a functional technology probe [35] focused on the key authoring aspects described in subsection 3.2.

4.1 Front-end Interface Design and Features

We designed the probe as a web application split into three interconnected pages. Each page corresponds to a specific step in the narrative creation process. The first page focuses on **in-nodes LLM generation rules**. Authors then arrange the content into a **story graph** and define the **transition conditions** between the nodes on the second page. Finally, they can test the resulting IDN on the third page. The upper part of Figure 7 demonstrates the entire user flow of the probe.

4.1.1 Page 1. Defining In-nodes LLM generation rules. This page allows authors to input the main atomic IDN entries, divided into four categories: world setting, character, narrator behaviour, and input variables (sequence from left to right in Figure 3. The page offers the following features:

- Entry Creation: Authors create new entries at the top of each column, (Figure 3-A). Each entry is characterised by a short name for easier identification, and the description influences LLM generation.
- Adding and Removing Entries(Figure 3-B): Authors can create multiple entries within each container. The "-" button allows authors to delete an entry, while the "Change" button enables editing an existing entry.
- Card Instantiation and Card Count (Figure 3-C): Upon creation, each entry is instantiated as a card for future steps.
- Save and Auto-Content(Figure 3-D): Users can save and restore their previous entries as JSON files. An 'Auto-Content' button automatically populates the text boxes with example entries for each column.

4.1.2 Page 2. Defining the Story graph and the Trigger Conditions. This page supports the design of the **Storygraph** and **Transitions** as follows:

- Node Creation: Users can create new nodes by clicking the "Add Stage" button(Figure 4-A). A new block representing the narrative node appears on the right side canvas (Figure 5, A to B-1).
- Assigning Cards: Each entry defined in Page 1 is instantiated as a card, displayed on the left side of this page (Figure 4-B). Users can assign the cards by drag-and-dropping them onto nodes (Figure 5-B-2). Each card can be assigned to several

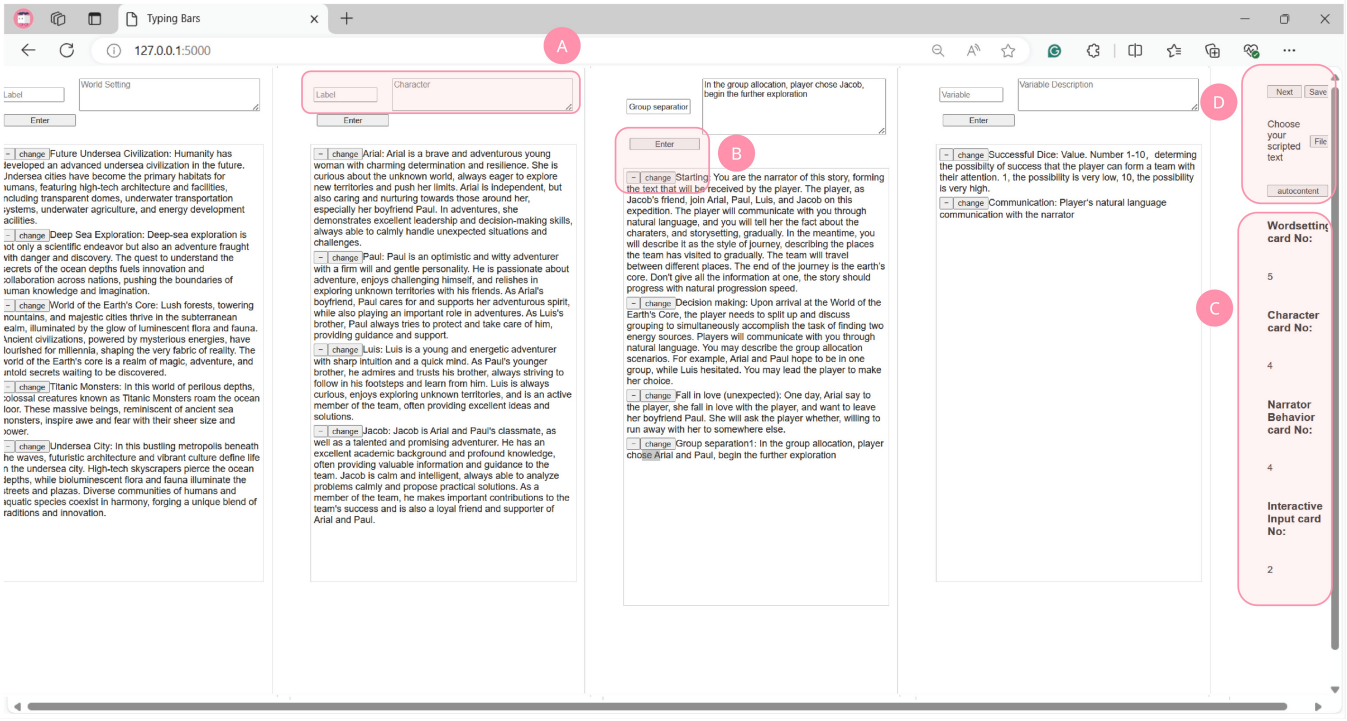

Figure 3: Probe Page 1. Textual Narrative Content Authoring. There are 4 columns where the author can put world setting, character, narrator behaviour, and player input separately. The background image is a screenshot from the actual system.

nodes, and each node can receive several cards. Users can also remove a card from a node by clicking the "-" button.

- Transition Creation: Each node features a pair of pink circles, representing entry and exit. Users can create transitions by linking these circles (Figure 5-C), creating a directed arrow between the nodes (Figure 5-D). The transition condition can be defined in natural language by clicking the "+" symbol on the arrow. This transition will occur when the generative narrative closely matches this description. Users can also delete the transition by clicking the "x" button(Figure 4-C).

4.1.3 Page 3. IDN testing. This page is designed for authors to test the IDN they defined in real-time. It includes the following components:

- Player input simulation: The author can type in the text box to simulate player actions. Pressing "enter" adds the text to the input container. Narrative testing starts at the first stage of the narrative, thus the first node of the story graph. When transitioning to a new node, the inputs update accordingly.
- After selecting input content, the author can click "Generate" at Figure 6-B. The LLM-generated narrative will appear as text entries at Figure 6-C. Each narrative turn is displayed as a separate entry, and all generated narratives are saved in a message list. To clear the narrative and restart, authors can click the button at Figure 6-G.
- Current stage: The current narrative stage is displayed at Figure 6-C and updates based on the transition conditions.
- Narrative history list checking: Authors can scroll through the list at Figure 6-C to view the narrative history. Each narrative piece has an icon next to it; clicking the icon highlights the input that generated that piece (Figure 6-E), helping authors to reference and reflect on it.
- Additional author instructions: If the generated narrative doesn't meet the authors' intentions, they can type instructions in the textbox at Figure 6-F. These instructions are saved in the entry history list, providing a reference when revisiting the previous pages for content changes.

4.2 Backend and Technical Implementation

Figure 7 illustrates the interaction between the front end and back end of the technology probe. The entire backend program is coded in Python, leveraging Langchain [4] for text embedding, RAG implementation, and input variable definitions. The front end of the probe is built using the Flask APP approach [5] in Python with JavaScript and HTML. We use GPT-3.5 for this probe as the LLM for generating text in real-time and text-embedding-ada-002 for text embedding and RAG.

The entries typed on Page 1 are sent to the backend in JSON format. *World setting* and *character* entries are processed into a dictionary for later text embedding (knowledge base, RAG). *Input variables* are processed by Langchain's definitions and added to the system prompt of the assigned node. The LLM will guide how

[4] https://www.langchain.com/
[5] https://flask.palletsprojects.com/en/3.0.x/

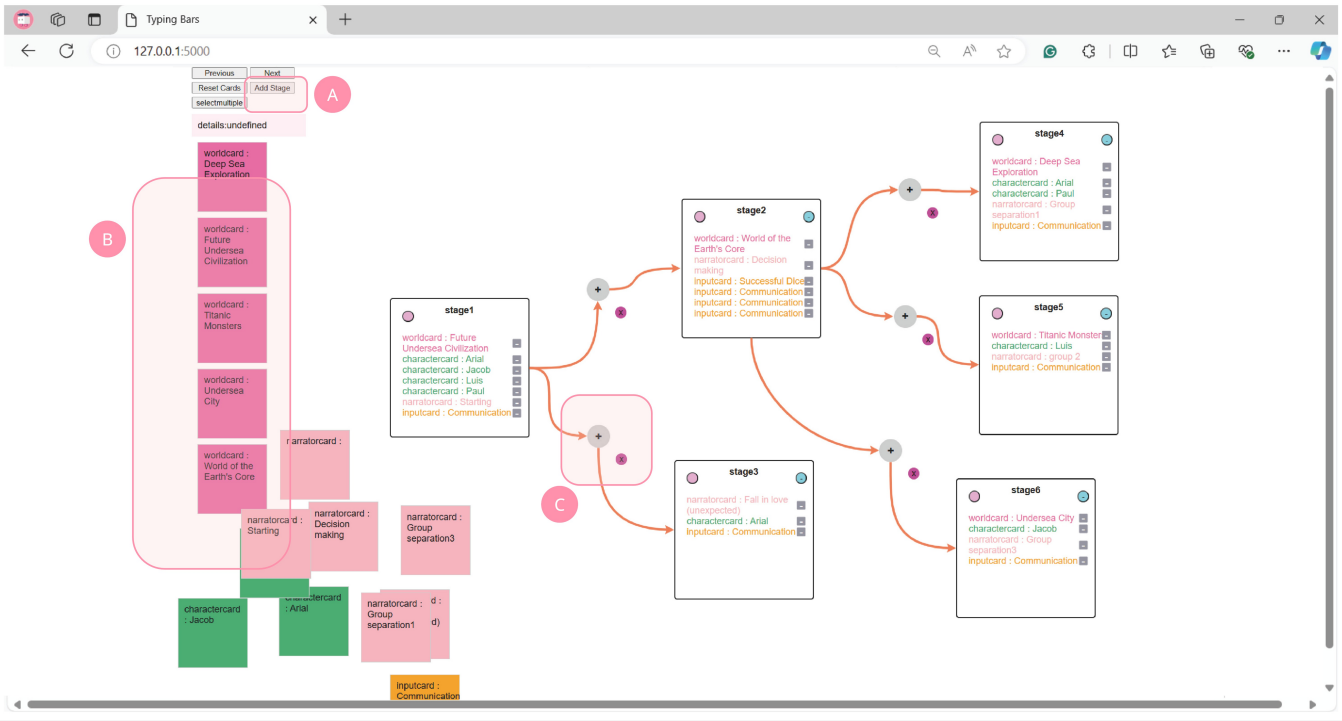

Figure 4: Probe Page 2: On this page, the author can create narrative stage blocks and allocate cards into the different blocks. They can then create connections between each block.

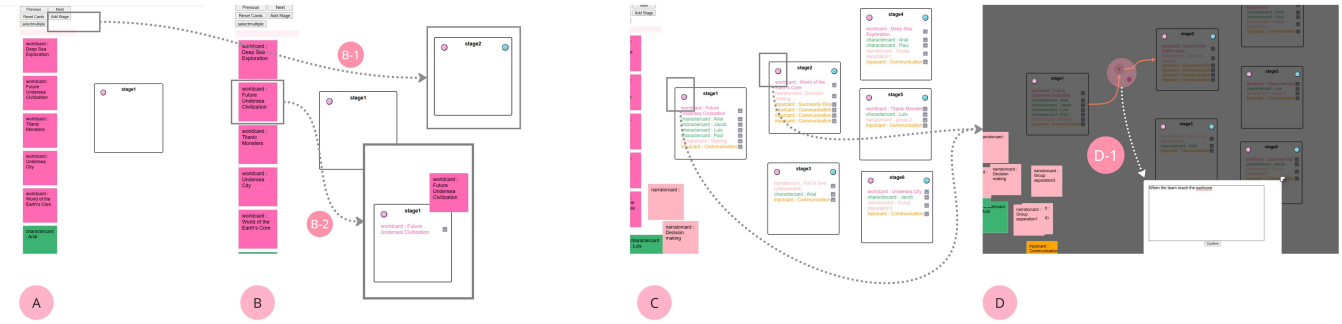

Figure 5: Operation details on Page 2

each input variable influences subsequent generations. *Narrator behaviour* entries directly integrate the system prompt of the assigned node's LLM.

On Page 2, each new node creation results in the instantiation of a corresponding item in the backend dictionary. Cards assigned to the node are appended accordingly. The cards are processed and formatted as detailed in the previous paragraph to provide a system prompt, knowledge base, and list of formatted inputs to the LLM for this node. Similarly, transitions and their descriptions are appended to the node's transition list. A special transition TC0 is introduced to represent the absence of transition, as shown in Figure 7. These transitions are evaluated through a similarity score between the generated narrative and the trigger condition statements in the node's transition list. It returns the name of the node with the highest score. If none of the nodes achieve a high similarity, the narrative may remain in the current stage, which meets TC0.

On Page 3, the input values simulated by the author are formatted into user prompts and sent to the model for generation. Each narrative stage (node) maintains its own history list. This approach prevents inaccuracies in judging trigger conditions. When a node transition happens, the history list is updated accordingly. The prompt engineering settings of the LLM are updated based on its assigned cards in the new node.

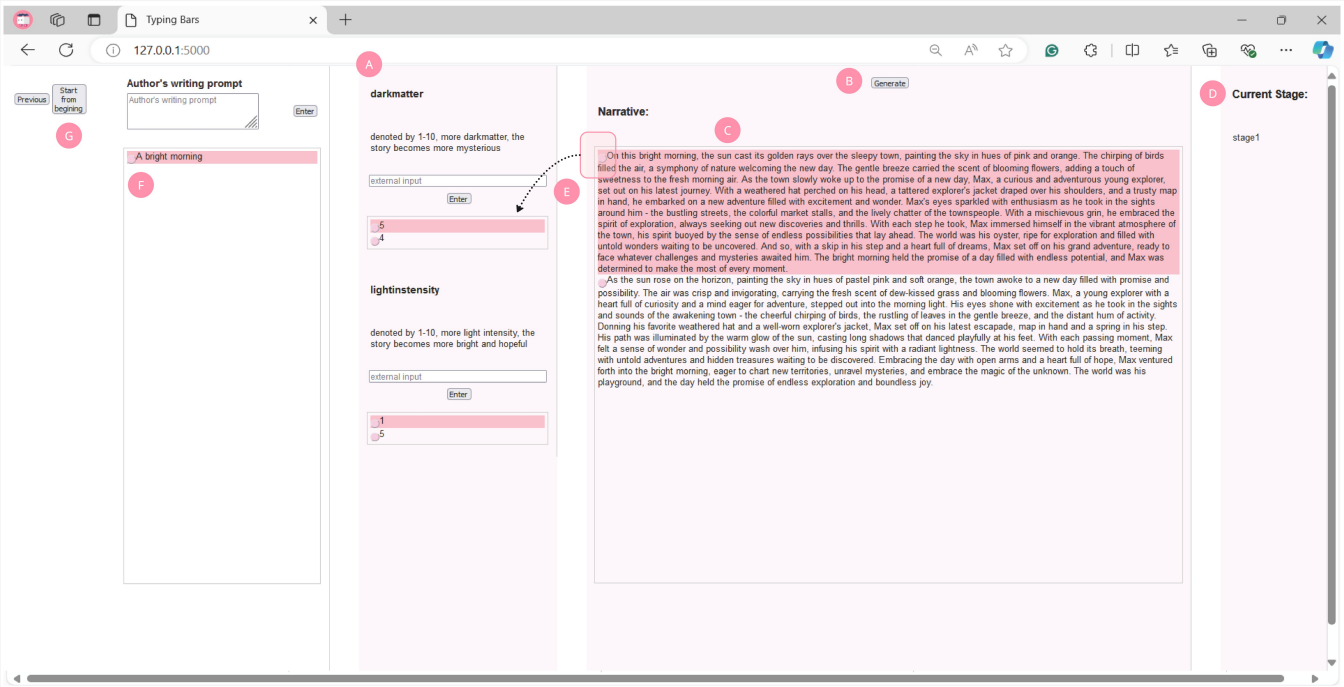

Figure 6: Probe Page 3: Interactive Narrative Testing. On this page, the LLM will operate, and the author can assign values to the input variables they designed for the current stage. The generated narrative content will be displayed step by step. Selecting specific text will highlight the related inputs, and the current stage will be shown in the upper corner. A screenshot from the actual system.

4.3 Example IDN Creation Workflow

We develop an example IDN to demonstrate how authors can use the *Orchid* creative approach with the developed probe. The example IDN is presented in Figure 8. Its design includes in-node narrative settings, the story graph, and transition conditions.

The narrative follows the player and their four friends – Arial, Paul, Jacob, and Luis – as they embark on an adventure guided by an AI narrator. Under the AI narrator's guidance, the group journeys to the Earth's core. Upon reaching the Earth's core, the AI narrator prompts the player to pick a teammate to explore the Earth's core. The player informs the AI narrator of their choice. A random input variable called "Successful Dice" is introduced, which value influences the success rate of that decision. If the player chooses Luis, they will proceed with him to the habitat of the Titanic monster. Jacob leads to an undersea city. Finally, picking Arial and Paul results in further branching adventures as they explore the deep sea together.

To implement this narrative using this probe, the author first enters the descriptions of five world settings and four characters (W1 to W5, C1 to C4 in Figure 8), together with descriptions of the narrator's behaviour for each stage on Page 1. The author also inputs the *Successful Dice* variable, described as *Number 1-10, determining the possibility of success that the player can form a team with their attention. 1 indicates a very low possibility, while 10 indicates a very high possibility.* On Page 2, the author creates the six-node story graph and assigns the in-node interaction cards according to Figure 8. The author also defines transitions between nodes in natural language, for instance, "When the team reaches the Earth Core".

5 Probing Orchid through Design Activities

We conducted a study inviting participants to design IDNs using the probe described in section 4 to explore the creative experience of Orchid from the author's perspective. The study addresses the following research questions:

- **RQ1:** How do authors create IDNs under *Orchid*?
- **RQ2:** How do authors perceive *Orchid* in authoring IDNs, particularly regarding authorial agency and machine contingency?
- **RQ3:** What are the specific needs for future systems supporting IDN creation under *Orchid*'s creative approach?

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. Participants were recruited through social media and personal connections. Participants were selected based on their intention to be future authors of LLM-driven IDNs and their experience in at least one of the following areas: 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

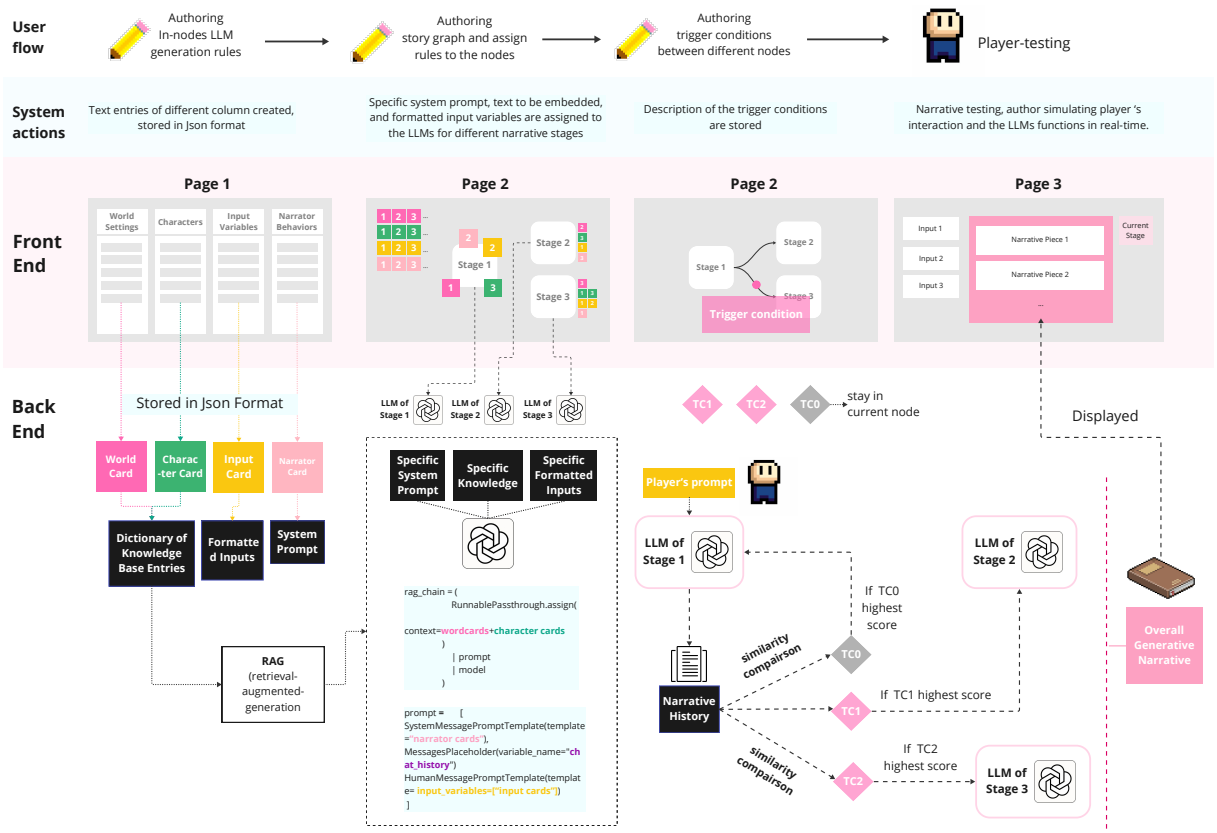

Figure 7: Front end and Back end of the technology probe. Upper: the user flow of the probe, describing how the author creates IDN using the probe through interacting with the three pages. Lower: Abstraction of the backend algorithm that supports the interactions and creation process of each page.

games. 3) Artists who have experience in applying IDNs in their practice. We also encourage participants without the mentioned experiences to apply, provided they have substantial experience in playing IDNs. Table 2 summarises their demographic information. All participants have experience with LLMs. Their backgrounds include IDN designers, AI game researchers, narrative artists, and designers or students in human-computer interaction.

Before the study, participants signed a consent form for the collection of anonymized personal data and interview recording. We created a demonstration video outlining the probe usage process and provided a writing template to help participants prepare their text and story approach, which they could use during the study. These materials were distributed after signing the consent form.

We established a remote link for participants to access the probe online using their personal computers, while we collected their data on a local server. The study was conducted via Zoom and recorded on video. Each session included one participant and two researchers for support and to receive live feedback. We first demonstrated the probe's usage before encouraging participants to think aloud and share their thoughts as they used the probe. After completing the first two pages, participants simulated player interactions through several rounds of narrative generation on the third page. Once participants felt they had tried enough rounds, we conducted a semi-structured interview focused on several themes:

1) *Theme1: Creative Intent and Mindset.* This theme includes questions such as: "Can you introduce the IDN you developed?" and "In what ways does this creative approach within the system influence your narrative ideation?" "How does your creative mindset different from your previous experience of authoring interactive or static narratives?", "How do you perceive the role of AI in this creative process?"

2) *Theme2: Authorial Agency and Machine Contingency*, "Explain how your authorial control is expressed when using this system, and where you feel you share this creative control?", "In what scenarios do you feel the story generation has gone beyond your control, yet you are satisfied with the outcome, and how do you view this machine contingency?", "In what scenarios do you feel unsatisfied with the machine contingency?", "How have your expectations of the agency changed compared to your previous creative approaches, such as static story creation or traditional IDN authoring?" For this theme, the researchers provided further explanations if the

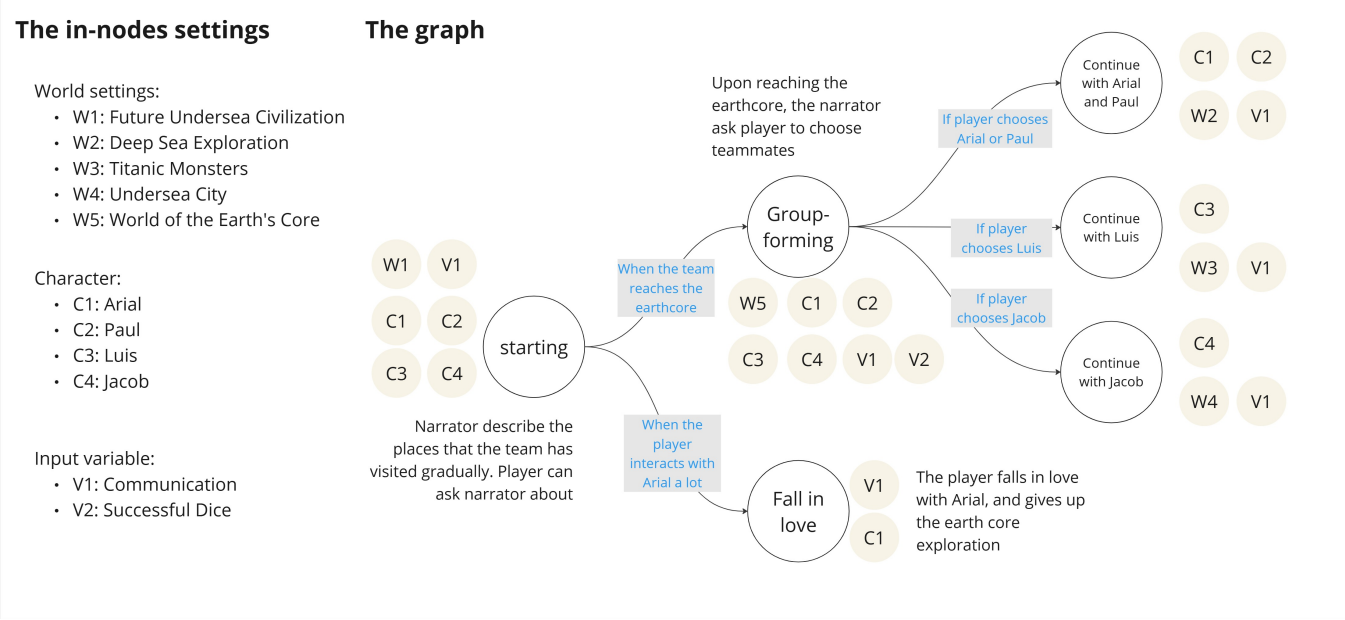

Figure 8: An example of an IDN that can be authored under *Orchid* creative approach and implemented using the technology probe, adapted from the designs created by one participant in the user study.

participants felt confused about the concepts of authorial agency and machine contingency.

3) *Theme3: Perceived difficulties, advantages, and open suggestion*: "How can this creative approach be applied to your current workflows?", "What additional features do you think are necessary to construct playable IDNs out of the system?", "Any open suggestions for the system?"

The entire study lasted 2 hours to 3 hours per participant, depending on the complexity of the IDNs that they constructed, and the interview was conducted in Chinese. 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

The analysis was divided into two parts: First, one researcher conducted a qualitative analysis of the IDNs created by participants during the study, focusing on the video recordings and the participants' prepared templates to address RQ1. Secondly, we conducted a qualitative analysis of participants' think-aloud recordings and interview responses, targeting RQ2 and RQ3. 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. Two researchers performed thematic coding [12] on the text, independently coding the data and identifying common codes, which were organized into axial codes. After refining subcodes, it resulted in a comprehensive codebook. The coding was conducted using Taguette [6]. A total of 5 themes and 105 codes were identified and agreed among the two researchers, demonstrated in Table 1.

[6]https://app.taguette.org

6 Results and Findings

6.1 RQ1: IDN typologies under Orchid

Appendix A summarises the IDNs created by participants, detailing the number of cards and nodes for each. Observations are presented by categories, addressing the three authoring aspects of Orchid:

6.1.1 Buds (nodes). **Worldbuilding:** From the analysis of *world-setting* and *character* cards, we found that some participants provide rich content in character development, with each card containing 150 words on average. Participants who adapted their existing writing material (P3, P10) tended to have more detailed descriptions. Some characters also exhibited strong development across the narrative stages, requiring separate descriptions for the same character across different nodes. For example, in P10's work, the protagonist undergoes an identity transformation due to family conflicts. Her personality shows distance and insecurity by the end of the IDN. P10 achieved this through creating two character cards for the same character.

Some participants provided extensive world-building, detailing political structures that lead to actual narrative scenarios (P6, P8, P9, P10). For example, P8 described a utopian world where citizens work for collective progress, rewarded equally by a central supercomputer.

In contrast, some participants focused more on the story graph design, featuring a large variety of characters and events. These cards thus contained briefer background and character definitions, around 20-30 words (P4, P5, P11, P12).

Table 1: Themes and Subthemes

Theme	Subthemes	Code
Authorial Agency and Contingency	- Perceptions of contingency - When/where authorial agency is needed - Highlighting the 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

Narrator behaviours : In defining narrator behaviours, seven participants mentioned specific styles and genres (P5, P11, P9, P12). For example, P5 described it as "classical sci-fi style: trying to mimic the classical sci-fi style, narrating the story like the book 'Snow Crash.'" Some participants also sought specific narrative perspectives (P1, P2, P4, P8, P10), such as P2 characterising one narrator's behaviour as: "Commentary narrative style: imitating the style of user comments on YouTube, which can include appreciative, kind, declarative, sarcastic, and insulting tones." We found that the requirement for a specific narration genre does not exist in P3, P7, and P10's work, which were more focused on world-setting and character definition.

In addition to narrative style, some participants described how the LLM should engage with players. For example, P7 requested that the LLM should "inform players of their roles when introducing the core conflicts." P3 noted that the LLM should "gradually reveal the truth through the brother's messages and hacking attempts to maintain player interest."

In-node interactions:. We found that user-designed inputs can be categorised into three types:

Natural Language. This is evident in most cases where the player is role-playing the protagonist of the narrative and play from a first-person perspective (P3, P6, P10). Other narrative structures invite players to write instructions in natural language to guide narrative progression. For example, P2's IDN invites players to write "comments" that shape the protagonist's adventure.

Number. Some participants assign specific numerical values to the emotional relationships of the characters and the probabilities of events (P5, P8, P11, P9). They create frameworks in which the course and outcome of the narrative can be parametrically controlled by number. For instance, P8 introduced a mood factor on a scale of 1 to 10. Scores above 8 may trigger confrontations among the different countries in the narrative.

Bool. Boolean variables mostly represents yes/no choices(P1, P7). For instance, in a story designed by P1 about a judge's rebellion, the protagonist's choice to resist or not is determined by the player's yes or no input, affecting the plot's direction.

Some participants assign the same input across different nodes, primarily in first-person role-playing (P3, P7, P10), while others use different inputs in various nodes, mostly using number input variables. We found that some participants linked input values as parameters automatically derived from the generated narrative, instead of active input from the player, which was not supported by the probe. There was also confusion about placing content in incorrect card categories. Some participants mistakenly view narrator behaviour instructions as part of world-building. Additionally, definitions of input variables that influence character behaviour are often placed in character cards. This confusion may arise from a lack of clear explanations about the background logic of different columns and the underlying system architecture (like RAG or system prompts) to the participants.

6.1.2 Stems (graph). Three types of narrative structures arose from the study.

Linear progressive. P5, P7, P9, P10 designed this type of narrative structure. This narrative structure features progressions from one stage to the next, aligning with traditional plot development and linear storytelling found in film and drama, but allows for precise scene segmentation. For example, P7's IDN features three stages akin to an interactive drama: the player witnesses a couple arguing on a train and can intervene (Stage 1), learns about the female protagonist Selina and her backstory (Stage 2), and finally meets the male protagonist Jessie for a conversation when the train stops (Stage 3).

Branching. Under this structure, the narrative begins with a clear start, which then branches into multiple paths. Distinct branches often lead to significant shifts in the world-settings and narrator behaviours, while the characters' definition, especially the protagonist, may remain unchanged. For example, in P2's design, while the protagonist Neo's character cards were assigned throughout the three nodes, negative comments by the player will send Neo to a desperate desert with distinct world settings and narrator behaviours.

Looping. This narrative structure features shifting perspectives, going back and forth, for a dynamic exploration of the story's development. In this graph type, each narrative stage features distinct narrator behaviours. For example, in P4's story, the protagonist

navigates three narrative stages -— Omniscient Mode (comprehensive details), Multiple Perspectives Mode (varied viewpoints), and Negative Mode (ironic contrasts) -— to uncover their true identity.

6.1.3 Transitions. Transitions are one of the most critical aspects of the created narratives. The current probe did not address all aspects of the transitions designed by participants. Some participants based their transitions on time in the narrative, like P9's "six months later." This is hard to recognise, as the LLM-generated narrative may not include time information. Additionally, some participants considered character changes, e.g., P10's "when Hua becomes corrupted". Such descriptions were difficult to assess through the LLM similarity comparison due to their complexity. These scenarios led to triggering transitions based on incorrect or incomplete data. Issues also arose when participants tried to link transitions to input variables. For instance, in P5's original design, a stage transition was to occur when tension exceeded 5, which is the input variable for controlling narrative generation.

However, the transition judgment worked well where participants defined the transition based on whether specific events or plot points happened. For example, in P1's work, "They were found by the big brother".

6.2 RQ2: Perceived authorial agency and machine contingency in Orchid

Interview codes from the first three themes in Table 1 provide insights for RQ2.

6.2.1 Agency addressed by Orchid. Participants noted that the graph design effectively supports authorial agency by allowing complete initial settings rather than relying on iterative conversations to train LLMs, as in their previous experiences. P1 stated, "Instead of loosely defining AI initially and making major changes later through chatting, it's better to establish a complete initial structure. A vague definition can lead to extensive revisions and reorganisation."

Participants also noted a shift in authorial agency compared to traditional script-based narratives. The first aspect is the willingness to share authorial agency with LLM. P5 highlighted that, compared to past experiences, Orchid offers a "meta" authoring experience, where he intends to retain control over the main plot while delegating secondary storylines to the LLM. P7 highlighted, "I'm open to the outcome, even if AI generates something completely beyond my control, I'm still curious about what will happen next." P1 described the collaboration with AI as "the entire story feels collaborative, striking a balance between the unexpected and the logical, unexpected yet reasonable." The second aspect is agency sharing with players. Some participants described Orchid as creating the dynamic that invites players to co-create. P6 noted that some agency is now shared with players, allowing them to create narratives within author-defined structures. P8 noted that *Orchid* fosters an open-ended creative model that encourages player participation and co-creation, a view supported by P12, who remarked this change in creative intent was "transformative".

6.2.2 Unexpected agency. We also observed some agency aspects that we did not anticipate when putting forward *Orchid*. Firstly, participants expressed a strong desire to control the narrative pacing, particularly the richness of content generated in each turn. Most participants consider the narrative progression to be too fast, with LLMs often presenting all the information in a single generation. Some details that were intended to be revealed gradually were instead disclosed all at once. For instance, P6 simulated the player with a single question, "Where am I now?". The LLM was supposed to introduce the first location, but the narrator immediately revealed all the locations at once.

Secondly, inaccuracies in transition judgments can cause the narrative to jump to the next stage before the previous one has finished. This led to fast transitions between stages, causing the loss of many important details within each stage. P5 reported that, "the transitions in the story are the aspects where we, as authors, need to retain our thinking. This is something I believe should not be left to AI." P8 designed transitions based on earlier events, such as characters completing several tasks. However, due to the probe limitations, each transition updated the LLM's memory, losing historical task completion data in previous nodes, thus preventing the expected transitions.

6.2.3 Contingency addressed by Orchid. Some participants praised the LLM's contributions in providing contingency. In a scenario featuring the protagonist passively observing a couple arguing, P7 was surprised when the LLM led one of the other characters to rush in and join the argument. P7 noted, "It's like creating a parallel universe; you have a sense of what will happen, but you don't know the exact details." P8 reported that he was surprised because he initially mentioned only that the protagonist had dual nationality. However, the LLM unexpectedly added that the protagonist was from South America and the Sahara Desert, along with a detailed depiction of the desert world that he hadn't authored. P9 was surprised by the inclusion of emojis in the outcome that she did not put any direct instruction towards that, which made her feel creative and aligned with her initial vision of a children's storytelling book (Figure 10)

6.2.4 Unexpected contingency. However, there were unexpected contingencies. The LLMs sometimes misinterpreted participants' intent, leading to generated content perceived as inaccurate. This includes misunderstandings of character names and mixing up gender and age (P3, P5, P10). For instance, P10, who paid much attention to the character details, mentioned that the LLM fabricated a non-existent character, creating a younger sister for the protagonists, despite their defined birth order. Additionally, the LLM changed a character's name and altered her role within the family. This was unacceptable to the P10.

Participants also reported that openness to contingency relates to whether the author focuses on playful mechanics or a specific theme they wish to express (P1, P3, and P7). For authors with strong intentions regarding structure, discovering that the generated content deviates from their established framework may significantly diminish their openness to contingency, and even regard "LLM [as] stupid" (P9). For participants with strong authorial intent, discrepancies in style and genre can lead to significant disappointment. For instance, P10 complained, "The LLM generated many perplexing rhetorical sentences in Western descriptive methods that do not reflect the characteristics of my Eastern novel." However, some participants expressed an understanding of the uncontrollable aspects of AI generation. They noted that, while the generated text is

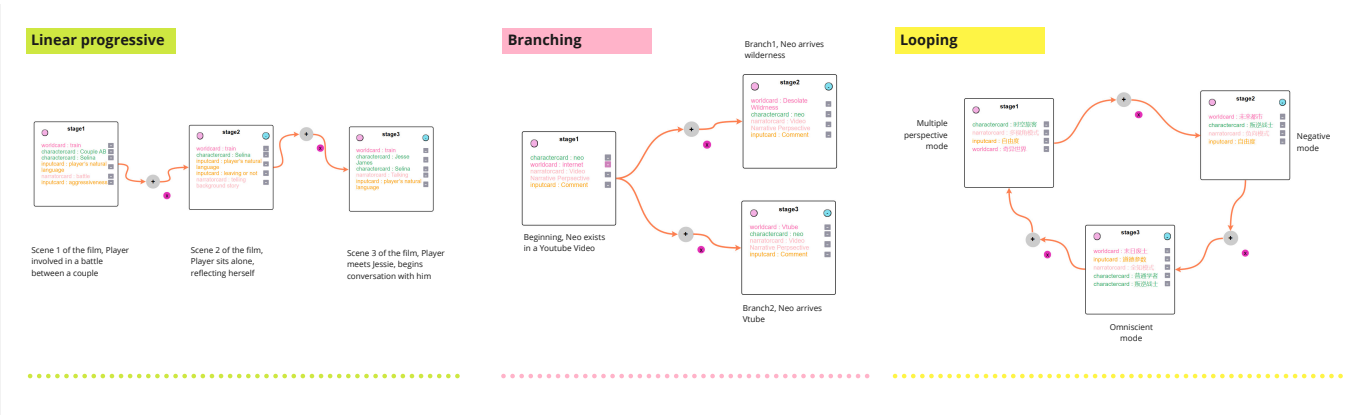

Figure 9: The three types of narrative structures created by participants using the probe, along with screenshots of the example story graphs. The linear progressive structure is from P7, the branching structure is from P2, and the looping structure is from P4.

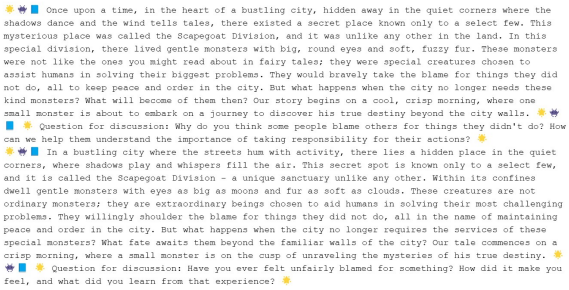
🌟 👾 📘 Once upon a time, in the heart of a bustling city, hidden away in the quiet corners where the shadows dance and the wind tells tales, there existed a secret place known only to a select few. This mysterious place was called the Scapegoat Division, and it was unlike any other in the land. In this special division, there lived gentle monsters with big, round eyes and soft, fuzzy fur. These monsters were not like the ones you might read about in fairy tales; they were special creatures chosen to assist humans in solving their biggest problems. They would bravely take the blame for things they did not do, all to keep peace and order in the city. But what happens when the city no longer needs these kind monsters? What will become of them then? Our story begins on a cool, crisp morning, where one small monster is about to embark on a journey to discover his true destiny beyond the city walls. 🌟 👾 📘 🌟 Question for discussion: Why do you think some people blame others for things they didn't do? How can we help them understand the importance of taking responsibility for their actions? 🌟 🌟 👾 📘 In a bustling city where the streets hum with activity, there lies a hidden place in the quiet corners, where shadows play and whispers fill the air. This secret spot is known only to a select few, and it is called the Scapegoat Division - a unique sanctuary unlike any other. Within its confines dwell gentle monsters with eyes as big as moons and fur as soft as clouds. These creatures are not ordinary monsters; they are extraordinary beings chosen to aid humans in solving their most challenging problems. They willingly shoulder the blame for things they did not do, all in the name of maintaining peace and order in the city. But what happens when the city no longer requires the services of these special monsters? What fate awaits them beyond the familiar walls of the city? Our tale commences on a crisp morning, where a small monster is on the cusp of unraveling the mysteries of his true destiny. 🌟 👾 📘 🌟 Question for discussion: Have you ever felt unfairly blamed for something? How did it make you feel, and what did you learn from that experience? 🌟

Figure 10: The generated outcome from P9's created IDN included emojis, giving it a children's book style, which surprised P9.

difficult to manage, its unpredictability mirrors the complexities of social issues and human problems, which can be fascinating (P7). P12 indicated that, due to her limited creativity in constructing stories, this lack of control could help her create content they could not have imagined otherwise.

6.2.5 Perception of collaborating with AI for creativity. The involvement of AI also influenced participants' creative mindset. Firstly, participants expressed concerns about AI's understanding, prompting them to change their writing habits. For instance, P2 worried that AI's comprehension would force her to alter her original style, leading her to add more detail, nuances, and genre constraints that she previously omitted. P5 reported that he initially wanted to incorporate numerical calculations but abandoned the idea due to concerns about potential AI errors. Conversely, some participants felt that AI's involvement made them want to challenge its creativity. P10 expressed that her purpose in using AI was to step outside her perspective and see if it could generate something innovative. P8 aimed to create complex character setups to test whether the AI could successfully understand the intricate relationships he designed. Five participants viewed AI as an asset rather than just an assistant. For instance, P9 explained that AI should not merely be seen as a tool for extending the author's thought but as an integral part of the overall creative experience.

6.3 RQ3: Needs and suggested functions

The elements of RQ3 are related to the last two themes from Table 1. User needs were identified from interview records regarding desired IDN outcomes and functions, as well as the highlighted advantages of existing probes. These insights inform design implications for future work to develop *Orchid* into an actual system.

6.3.1 Ideal generative outcome. Based on the participants' comments on their ideal generated outcome, there are two key expectations, which is currently not achieved by Orchid:

1) *Coherence:* Four participants emphasised the importance of coherence and 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. This issue is closely related to the LLM memory management in the current probe.

2) *Narrative Progression:* Similarly, four participants expressed concern about whether the plot was advancing and identified a lack of progression as a significant issue. For example, in P11's generated outcome, he complained that it kept describing the world-setting, instead of introducing the core event.

6.3.2 Suggested functions. **Stricter Trigger Judgment.** Most participants stressed that trigger judgments should be strict and clear, rather than depending on a "black box" mechanism based on similarity comparison (P1, P4, P6, P7). P5 suggested using deterministic checks on values, similar to traditional games, as opposed to LLM-driven similarity checking to help enforce the transition conditions.

Providing tutorials and narrative templates. Six participants expressed the need for guidance on defining input variables, which are unfamiliar to the participants (P1, P4, P11, P12). P4 noted that certain narrative styles, such as circular narratives, may confuse users and suggested offering various structural graph templates.

Explainability and real-time guidance. Participants expressed a need for real-time guidance and explanation by an AI

agent if *Orchid* is developed as an actual system. For instance, it was unclear how much content should be included in the cards to ensure high-quality output. As P11 suggested, an AI assistant could provide guidance on optimising card definitions. Additionally, they anticipated a function showing how each generated segment relates to earlier definitions (P1, P11, P12).

Dynamic text modification and retention. Participants reported the intention of the ability to directly modify the generated text to correct the LLM if the generated outcome is out of their intention. P1 and P5 suggested that the model should seamlessly incorporate these author adjustments into its understanding in the next round of testing.

Integration with Interactive Input/Output. Participants expressed a strong desire for the tool to support connections to real interactive input and output systems. For interactive input, participants envision direct data integration with commonly used game engines, such as Unity for utility in dynamic environments(P2, P3, P5). On the output side, there is a call for visual development tools like TouchDesigner and generative AI platforms focused on visuals, such as Stable Diffusion and Midjourney.

7 Discussion

7.1 Addressing the Two Hypothesis

Orchid's first hypothesis focuses on its capacity to facilitate open-ended collaboration among authors, machine creativity, and player interactivity to generate emergent textual content. Participants reported a successful experience of authorial agency enabled by the node and graph structure, expressing openness to share a portion of their agency with AI while. Additionally, some participants noted the beneficial contingency introduced by LLMs. Engaging with Orchid led most participants to perceive a notable difference in their authoring experience compared to their past experiences.

While participants are open to the contingency brought by LLMs, there are varying requirements for this contingency among different authors. Those who prioritize literary details are more concerned with whether the narratives produced by Orchid accurately reflect the specifics defined by the author. Conversely, authors inclined toward structural experimentation focus on the accuracy of conditional transitions. This raises questions about whether "literary-focused" or "structurally rich"—is more suitable for LLM driven IDNs.

Orchid's second hypothesis focuses on supporting the creation of richer LLM-driven IDN's structures. This is supported by the work produced by our participants, which showcases structures not typically seen in the IDN field.

7.2 Design Implications for Orchid

This study reveals that AI's understanding of character relationships and the authors' world settings still has considerable room for improvement for effective application in IDNs. More rigorous research is needed, focusing on prompt engineering strategies and iterative evaluations based on authors' assessments. Additionally, advanced computational approaches, such as applying knowledge graphs, can facilitate understanding of character relationships. Explainability mechanisms should also be established. For instance, an additional testing interface could be implemented to confirm the LLM's understanding before proceeding to narrative generation.

Given the limited context token capacity of LLMs, some narrative content can easily lose its previous context after long generation rounds, harming coherence. Implementing a dynamic memory management system is crucial for improvement. This memory should abstractly represent narrative progression while covering the key facts that help the AI maintain understanding for future generations.

Although *Orchid* provides a graph to structure the IDN, the LLM itself can not understand the relationships between nodes, and the hierarchy of the structure. This may lead to poor narrative cohesion and requires future attention. Moreover, trigger conditions are a key aspect where the current probe falls short. Future efforts need to consider the accuracy of trigger conditions as an integral element of authorial agency. It also requires further investigation into whether trigger conditions should be managed by AI or handled using numerical comparisons or hard-coded methods.

7.3 Limitation and Future work

While we gathered significant qualitative insights from the user evaluation, our work does not collect quantitative metrics such as measures of human-AI collaboration and authoring experience. Moreover, the creative approach of *Orchid* involves individual components supporting the human-AI co-creation process. Future studies will address how significantly each component affects the user's authoring experience and the final output, involving testing by ablating individual components. We also did not evaluate the quality of the generated output, which requires more rigorous assessments. Finally, since this paper mainly focuses on the perspective of authorship, evaluation from the player's perspective on the IDN outcomes will be the focus of future studies.

8 Conclusion

This study introduces *Orchid*, a new approach to facilitate collaborative narrative creation between human authors and LLMs. Orchid emphasises 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] 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

[6] 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

[7] 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

[8] 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.

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

[10] 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.

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

[12] 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

[13] 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

[14] 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

[15] 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.

[16] 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

[17] 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

[18] Elizabeth Clark, Anne Spencer Ross, Chenhao Tan, Yangfeng Ji, and Noah A. Smith. 2018. Creative Writing with a Machine in the Loop: Case Studies on Slogans and Stories. In *Proceedings of the 23rd International Conference on Intelligent User Interfaces* (Tokyo, Japan) *(IUI '18)*. Association for Computing Machinery, New York, NY, USA, 329–340. https://doi.org/10.1145/3172944.3172983

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

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

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

[22] 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.

[23] 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

[24] 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

[25] 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).

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

[27] 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.

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

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

[30] 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/.

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

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

[33] 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/.

[34] 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

[35] 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

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

[37] 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.

[38] 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

[39] 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

[40] 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

[41] 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

[42] 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.

[43] 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.

[44] 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.

[45] 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.

[46] 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.

[47] 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

[48] 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.

[49] 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).

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

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

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

[53] 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

[54] 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.

[55] 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

[56] Jeba Rezwana and Mary Lou Maher. 2023. User Perspectives on Ethical Challenges in Human-AI Co-Creativity: A Design Fiction Study. In *Proceedings of the 15th Conference on Creativity and Cognition* (Virtual Event, USA) *(C&C '23)*. Association for Computing Machinery, New York, NY, USA, 62–74. https://doi.org/10.1145/3591196.3593364

[57] 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.

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

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

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

[61] 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.

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

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

[64] Nikhil Singh, Guillermo Bernal, Daria Savchenko, and Elena L. Glassman. 2023. Where to Hide a Stolen Elephant: Leaps in Creative Writing with Multimodal Machine Intelligence. *ACM Trans. Comput.-Hum. Interact.* 30, 5, Article 68 (Sept. 2023), 57 pages. https://doi.org/10.1145/3511599

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

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

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

[68] 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.

[69] 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

[70] 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

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

[72] 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.

[73] Find the Lamp Games. 2024. Turlock Holmes. https://store.steampowered.com/app/2142160/_/?l=schinese Video game.

[74] 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

[75] 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

[76] 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

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

[78] Qian Wan, Siying Hu, Yu Zhang, Piaohong Wang, Bo Wen, and Zhicong Lu. 2024. "It Felt Like Having a Second Mind": Investigating Human-AI Co-creativity in Prewriting with Large Language Models. *Proc. ACM Hum.-Comput. Interact.* 8, CSCW1, Article 84 (April 2024), 26 pages. https://doi.org/10.1145/3637361

[79] 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

[80] 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

[81] Zihan Yan, Chunxu Yang, Qihao Liang, and Xiang 'Anthony' Chen. 2023. XCreation: A Graph-based Crossmodal Generative Creativity Support Tool. 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 48, 15 pages. https://doi.org/10.1145/3586183.3606826

[82] 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

[83] 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

[84] Zheng Zhang, Jie Gao, Ranjodh Singh Dhaliwal, and Toby Jia-Jun Li. 2023. VISAR: A Human-AI Argumentative Writing Assistant with Visual Programming and Rapid Draft Prototyping. 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 5, 30 pages. https://doi.org/10.1145/3586183.3606800

A Summary of participant created IDNs during the study

Table 2: Demographic background of participants and their created IDNs using the technology probe.

Participant No	Career and Related IDN Experience	IDN typologies	Summary of created IDN
P1	Interaction Designer.	Branching, few cards with rich descriptions	The player commits crimes and faces police capture, with narrative generated based on choices of resistance or compliance. (6 cards, 5 nodes)
P2	Game designer with LLM experience.	Branching, rich cards with short description	Neo starts his video account, documenting an adventure in an abandoned wilderness. Players' comments on his livestream will guide his exploration choices. Malicious comments can harm Neo's mindset, reducing his chances of success, while his fans will also comment to steer him forward. (16 cards, 3 nodes)
P3	Table Top RPG writer.	Branching, few cards with rich descriptions	In a future Earth ruled by evil aliens, the player, as the protagonist, seeks to reunite with their missing family in a reality-distorting dream world, while navigating a relationship with a senior female student whose conversations impact their paths. (11 cards, 3 nodes)
P4	Virtual Interactive Film Production.	Looping, rich story graph with short descriptions	Generative narratives with a jump mechanism, including omniscient and unreliable narrators, allowing players to assign values that influence the narrative. (14 cards, 3 nodes)
P5	Researcher in AI and Games.	Linear progressive, rich story graph with short description	In a VR Metaverse, a narrator tells stories about Jeffery and K, with the plot evolving based on players' controlled variables including trust, love and tension. (10 cards, 3 nodes)
P6	Interaction Design student.	Branching, rich story graph with short description	A journey to the Earth's core where the player communicates with NPCs and teams up with one of three characters, leading to varied experiences. (12 cards, 7 nodes)
P7	Curator and interaction designer.	Linear progressive, cards with rich descriptions	Inspired by the film "Before Sunrise," players control a female protagonist who meets Jesse and chooses whether to join his exploration of Vienna. The rounds of conversation between the player and other characters will lead to scene transitions. (13 cards, 2 nodes)
P8	Social and Interaction Design student.	Looping, many cards with rather concise descriptions	Players build economies and explore space tech, influenced by Quantum (mood scale 1-10) and Hierarchy (staff levels 1-3). Experience follows a loop of daily routine: Morning: Work in the Office. Lunch: Chat in the Canteen. Afternoon: Return to the Office. After Work: Relax at the bar. (18 cards, 4 nodes)
P9	Narrative Art student with GenAI experience.	Linear progressive, many cards with rich descriptions	A three-stage IDN designed for kids, the narrator shares stories about scapegoats, ending with questions, allowing kids to determine the scapegoat's fate using tangible wood blocks. (15 cards, 3 nodes)
P10	Creative Writer of traditional novels.	Linear progressive, few cards with very rich descriptions	In an Eastern world of royal struggles, players control YUE, the crown prince, making choices that impact relationships and family drama across three scenes: 1. Hua is named heir apparent and betrothed to a humble man. 2. During a confrontation, Hua learns Yue plans to ally with the largest noble family. 3. Resentful of her mother's favoritism, Hua decides to pursue her own power struggle. (12 cards, 3 nodes)
P11	Web3 Designer specialized in character design.	Branching, rich story graph with short description	Different fighting scenes about the protagonist, Sara, with players defining explosions' intensity. Encountering monster Qiyu will shift the narrative to different paths. (10 cards, 5 nodes)
P12	Narrative writer in children's storytelling.	Branching, few cards, few descriptions and nodes	A simple linear narrative inspired by "Little Women." (11 cards, 2 nodes)