



ClueCart: Supporting Game Story Interpretation and Narrative Inference from Fragmented Clues

Downloaded from: <https://research.chalmers.se>, 2026-03-14 05:09 UTC

Citation for the original published paper (version of record):

Wang, X., Cao, Y., Xiong, J. et al (2025). ClueCart: Supporting Game Story Interpretation and Narrative Inference from Fragmented Clues. Conference on Human Factors in Computing Systems - Proceedings. <http://dx.doi.org/10.1145/3706598.3713381>

N.B. When citing this work, cite the original published paper.



ClueCart: Supporting Game Story Interpretation and Narrative Inference from Fragmented Clues

Xiyuan Wang
School of Information Science and
Technology
ShanghaiTech University
Shanghai, China
wangxy7@shanghaitech.edu.cn

Sizhe Chen
Chalmers University of Technology
Gothenburg, Sweden
sizhec@student.chalmers.se

Yi-Fan Cao
Computational Media and Arts
Hong Kong University of Science and
Technology
Hong Kong, China
ycaoaw@connect.ust.hk

Wenxuan Li
University of Southern California
Los Angeles, California, USA
wli17702@usc.edu

Junjie Xiong
School of Information Science and
Technology
ShanghaiTech University
Shanghai, China
xiongjj@shanghaitech.edu.cn

Junjie Zhang
Computational Media and Arts -
Information Hub
The Hong Kong University of Science
and Technology (Guangzhou)
Guangzhou, Guangdong, China
jakezhang@hkust-gz.edu.cn

Quan Li*
School of Information Science and
Technology
ShanghaiTech University
Shanghai, Shanghai, China
liquan@shanghaitech.edu.cn

Abstract

Indexical storytelling is gaining popularity in video games, where the narrative unfolds through fragmented clues. This approach fosters player-generated content and discussion, as story interpreters piece together the overarching narrative from these scattered elements. However, the fragmented and non-linear nature of the clues makes systematic categorization and interpretation challenging, potentially hindering efficient story reconstruction and creative engagement. To address these challenges, we first proposed a hierarchical taxonomy to categorize narrative clues, informed by a formative study. Using this taxonomy, we designed *ClueCart*, a creativity support tool aimed at enhancing creators' ability to organize story clues and facilitate intricate story interpretation. We evaluated *ClueCart* through a between-subjects study (N=40), using *Miro* as a baseline. The results showed that *ClueCart* significantly improved creators' efficiency in organizing and retrieving clues, thereby better supporting their creative processes. Additionally, we offer design insights for future studies focused on player-centric narrative analysis.

*Corresponding Author.



This work is licensed under a Creative Commons Attribution 4.0 International License. *CHI '25, Yokohama, Japan*

© 2025 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-1394-1/25/04
<https://doi.org/10.1145/3706598.3713381>

CCS Concepts

• **Human-centered computing** → **Human computer interaction (HCI)**; *Interactive systems and tools*; *User interface toolkits*.

Keywords

Creativity Support Tool, Game Storytelling, Indexical Storytelling, Story Interpretation

ACM Reference Format:

Xiyuan Wang, Yi-Fan Cao, Junjie Xiong, Sizhe Chen, Wenxuan Li, Junjie Zhang, and Quan Li. 2025. ClueCart: Supporting Game Story Interpretation and Narrative Inference from Fragmented Clues. In *CHI Conference on Human Factors in Computing Systems (CHI '25)*, April 26–May 01, 2025, Yokohama, Japan. ACM, New York, NY, USA, 26 pages. <https://doi.org/10.1145/3706598.3713381>

1 Introduction

In video games, the story often serves as a significant selling point, providing players with an immersive experience within the game world [9, 23, 102]. It drives player-generated content and discussions, with interpretations of game stories [16] becoming increasingly popular through videos and forum discussions [3, 15, 55]. On platforms like *YouTube*¹ and *Reddit*², these **players**—who also act as content **creators** by producing and sharing their interpretations through video and forum posts—contribute to a vibrant discourse. Here, content creators refer specifically to those analyzing game narratives, rather than those of user-generated content

¹<https://www.youtube.com/>

²<https://www.reddit.com/>

within games or game worlds. Their content, which garners millions of views (fig. 1(A)) and votes (fig. 1(B)), reflects the depth of player engagement with game stories.

Popular games often employ indexical storytelling [30] to deliver their narratives. This technique uses *indices*—elements within the game world such as items and environments that physically point to events or happenings—as clues to invite players to reconstruct the story. In *Elden Ring* [43], for instance, players might find cryptic messages left by other characters or environment hints about the lore. Without explicit answers from designers, this approach gives players the freedom to develop their own understanding of game events, leading to varying interpretations [4, 30, 57, 83]. The enigmatic nature of these stories and the diversity of viewpoints fuel players’ enthusiasm for discussing game narratives. This, in turn, motivates players, acting as creators, to publish their interpretations, contributing to a rich and diverse content ecosystem.

However, indexical storytelling presents significant challenges for creators of interpretations. The volume and diversity of narrative clues, ranging from environments to item descriptions, make it difficult to **organize** them effectively. The sheer number of these clues can be overwhelming. Additionally, its fragmented structure requires substantial effort to **interpret** the narrative clues holistically. Since the story elements are scattered throughout the game world, creators must invest considerable time and analytical skill to piece together a coherent narrative.

Addressing the challenge of **organization**, current game narrative research primarily classifies narrative elements from a game designer’s perspective [10, 17, 28, 94, 120–122]. Designers focus on creating coherent and immersive narratives, emphasizing structure and progression. In contrast, interpretation creators aim to engage their audience by presenting the story from various perspectives, exploring narrative elements non-linearly and interpretively. The designer-centric classifications do not support this exploratory and interpretive nature, resulting in heavy manual labor for creators when must organize and connect these clues to form a comprehensive understanding.

Concerning the challenge of **interpretation**, current Human-Computer Interaction (HCI) research focuses on the effectiveness of creativity support tools (CSTs) in underpinning user creative practices [39, 50, 93]. These tools, widely used in story creation processes [13, 58, 60], emphasize story analysis and creative thinking. They primarily utilize visualization methods such as storyline [49, 108, 109], timeline [12, 49, 74, 112], and storytelling [29, 70, 113]. However, these tools are unsuitable for non-linear interpretation processes and are designed more for creating new stories rather than interpreting existing narrative clues. Concept maps are another type of CST used to stimulate idea generation and aid creativity [66, 93, 95, 117, 118]. They are effective for note-taking and summarizing, helping users glean key concepts, relationships, and hierarchies from documents and source materials [19]. However, they are not equipped for the dynamic and interpretive processes required in game narrative construction. These gaps underscore the need for specialized tools for game story interpretation, supporting story analysis, creative thinking, and the effective organization and interpretation of narrative clues.

The gaps between existing classifications and tools underscores the urgent need for research that addresses the complexities of

player-driven interpretation. There is a necessity for tools and methodologies that assist creators in effectively organizing, understanding, and interpreting narrative clues. Given these challenges in the current landscape of game story interpretation, we propose investigating the following research questions (**RQs**) to bridge these gaps:

- **RQ1:** What taxonomy of narrative clues is most applicable for interpretation creators?
- **RQ2:** How do creators interact with the proposed tool?
- **RQ3:** To what extent does the proposed tool improve the efficiency of creators’ interpretation processes?

To address these questions, we began by exploring how creators classify and organize game clues through a combination of methods: a literature review, an application survey, and a co-design workshop involving 14 participants. From this, we identified their common classification habits, the challenges they encounter, and distilled the design requirements for our tool. Building on these insights, we derived a clue taxonomy and developed an in-game clue collection method based on it. We then created *ClueCart*, a creativity support tool that helps creators organize, understand, annotate, and explore relationships between various clues, thereby supporting the entire story interpretation workflow.

A between-subjects study with 40 creators compared *ClueCart* with the baseline, *Miro*, a widespread tool in creative fields, especially for story mapping [6, 47]. For quantitative evaluation, we assessed multiple dimensions: perception, including usability and creativity support, specific features of *ClueCart*, and story interpretation outcomes, involving completion time and quality. We also conducted qualitative interviews to gain deeper insights into users’ experiences. The results confirmed the validity of the taxonomy and demonstrated improvements in both the efficiency and quality of story interpretation. In summary, our contributions are:

- We explore narrative clues taxonomy through a literature review, an application survey, and a co-design workshop.
- We develop an open-source creativity support tool, *ClueCart*³, along with a clue collection method, enabling creators to organize, understand, and interpret game stories based on fragmented clues.
- We evaluate *ClueCart* through a between-subjects study (N=40), comparing it a baseline, *Miro*⁴, and offer design implications for future studies on player-centric narrative analysis.

2 Background and Related Work

2.1 Background

2.1.1 Game Story Interpretations. Game story interpretation via videos and forums has surged in popularity. Platforms like *YouTube* host numerous videos analyzing intricate game stories, attracting millions of views (fig. 1(A)). Games like *Elden Ring*, *Dark Souls* [42] and *Hades* [105] have detailed analysis videos demonstrating players’ deep interest in game stories. Similarly, *Reddit threads* allow players to exchange insights, enhancing collective understanding (fig. 1(B)).

³<https://cluecart.github.io/ClueCart/>

⁴<https://miro.com/app/dashboard/>



Figure 1: Examples of game story interpretations on (A) Youtube and (B) Reddit.

Videos play a crucial role in game story interpretation, ranging from simple walkthroughs to in-depth analyses. These videos explore lore, character motivations, and mysterious plots, making complex narratives accessible to a broader audience through engaging visual and auditory content. Community forums and social media platforms further facilitate the collective interpretation of game stories. *Subreddits* dedicated to specific games serve as hubs where players dissect narrative elements and propose interpretations, building on each other’s ideas to form comprehensive understandings. Game story interpretation through videos and community discussions significantly enhance player engagement and immersion [9, 23, 102].

Digital whiteboards, such as *Miro*, are widely employed in creative fields to support collaborative brainstorming and story mapping [6, 47], where visualizing complex narratives is essential. Creators often rely on these tools to manually arrange screenshots, notes, and diagrams that represent fragmented narratives, creating ad-hoc structures to analyze and communicate their interpretations. Research has demonstrated the effectiveness of such tools in organizing intricate ideas and enhancing team collaboration [52, 81], making them ideal for creators to better interpret and communicate multifaceted narratives in video games. However, as the complexity and number of story fragments in modern games grow, these tools often struggle to manage the increasing intricacies of game story analysis. Our work advances such tools by proposing and integrating a taxonomy for the automatic organization and retrieval of game story clues.

2.1.2 Indexical Storytelling. Besides the longstanding academic debate between Ludology and Narratology [37, 38, 53, 71, 116], which explores whether video games should be understood primarily through their mechanics or as narrative mediums, recent research has focused on interactive forms of narrative [31–34, 62, 63]. One notable form is indexical storytelling, introduced by Fernandez et al. [30], who applied a semiotic perspective based on Charles Peirce’s philosophy of language [59]. Building on the approaches of Nitsche [73] and presentations by Rouse [84], Smith and Worch [99] at the Game Developers Conference in San Francisco, Fernandez

et al. conceptualized “indexical storytelling” through meaningful objects, encounters, and traces left by other players.

In games using indexical storytelling, players are not merely passive recipients of a narrative. They are invited to interpret clues—whether physical objects, environmental details, or traces of past player actions—to piece together a cohesive story. The inherently mysterious nature of these narratives creates substantial room for interpretation, motivating players, as creators, to share their unique perspectives and analyses within the gaming community [3, 55]. Game examples include detective works in *Phoenix Wright: Ace Attorney* [14], remains interpretation in *BioShock* [1], and signage and tutorials in *Super Mario 64* [72]. Additionally, it reconstructs player history through traces left by players, as seen in *Demon’s Souls* [41]. Subsequent works have further explored the archaeological aspects of indexical storytelling [57, 100] and expanded its scope. Warpefelt et al. [119] argued that NPCs (non-playable characters) fit within this framework, with their attributes and visual presentations enriching the narrative experience for players. Our work builds on the concept of indexical storytelling, examining how story interpreters understand and classify this form of storytelling, thereby enhancing the narrative experience and engagement in video games.

2.2 Game Narrative Elements Classification

Research on the classifications of game narrative elements, such as *cut-scenes*, *dialogue*, *text*, *voice-over narration*, *narrative artifacts*, and *environment*, includes a wide range of approaches. These elements are the foundational components of a game’s story, shaping how narratives are presented and experienced by players [10, 17, 28, 44, 94, 120, 121].

Most research classifies narrative elements from the perspective of game story designers. For instance, Wei et al. [120] classify elements into horizontal, vertical, and modal embeddings based on the relationship between the embedded narrative and the main narrative. Some studies borrow models and theories from other fields [44, 94]. For example, Silva et al. [94] utilized the literary theory of J. Hillis Miller [67], categorizing elements into situation,

character, and form. Furthermore, some research focuses on the interactive narrative elements [10, 28]. Brandse et al. [10] categorized elements based on their interactivity and traditional aspects, while Domsch et al. [28] classified elements into passive, actively nodal, and dynamic categories.

However, existing research predominantly reflects a designer-centric perspective, focusing on how designers structure and present narrative elements [28, 94, 120, 121]. It neglects the way that players, acting as creators piece together diverse narrative fragments to engage their audiences, which often leads to significant manual effort. Our work addresses this gap by proposing a new taxonomy that emphasizes the interpretation of indexical storytelling clues, offering a novel perspective for understanding and analyzing complex narratives in games.

2.3 Creativity Support Tools

Creativity Support Tools (CSTs) in HCI research aim to improve creative practices by providing innovative ways for generating, organizing, and refining ideas [39, 50, 93]. These tools are widely utilized in story analysis and creative thinking, including storytelling [29, 70, 113], timeline visualization [12, 49, 74, 112], and storyline visualization [27, 49, 75, 76, 108, 109]. They also include other tools aimed at enhancing story creation [13, 58, 60]. For instance, Tang et al. [108] utilized reinforcement learning to optimize storyline layout, aiding creators in plot analysis. Similarly, *TimeSplines* [74], a timeline visualization tool, emphasizes reflection, enabling creators to effectively analyze temporal story flow. Tools like *StoryKit* [8] enhance creative expression by integrating text, images, and sounds into storytelling. Collectively, these tools promote a more immersive storytelling experience. Despite these advancements, existing story analysis and creation tools primarily target the development of new narratives rather than the interpretation and reconstruction of fragmented narrative clues common in complex game storytelling. They do not address the challenges creators face when piecing together diverse narrative fragments into coherent stories.

Concept maps are another type of tool designed to stimulate creativity [56, 66, 93, 95, 107, 117, 118] by organizing thoughts and ideas. Le et al. [56] tested the use of concept map planning to support creativity in photo stories, hypothesizing that organizational skills taught through concept maps would improve creative outcomes. Additionally, Wang et al. [118] integrated the multi-modal framework of learning analytics with the concept mapping approach to enhance students' vocabulary and reading abilities. While concept maps are effective for note-taking, summarizing, and organizing thoughts [19], they are not adequately equipped for the dynamic and interpretive processes required in game narrative construction. Existing tools focus on improving creativity and organization but do not address the specific needs of reconstructing fragmented game narratives.

To fill these gaps, our approach combines concept maps with clues classification and interactions derived from formative studies. This method aids creators in organizing, analyzing, and interpreting narrative clues, facilitating the dynamic synthesis of these elements into cohesive stories, and enriching the creative and interpretive processes.

3 Methodology

With approval from our institution's IRB, we conducted a formative study involving 14 players who are also content creators experienced in sharing their interpretations of game narratives with the public. This study aimed to explore how players, in their role as creators currently interpret game stories and included two main components: (1) a literature review and an application survey (fig. 3 (1)), and (2) a co-design workshop (fig. 3 (2-6)).

3.1 Literature Review and Application Survey

Besides the literature review in section 2.2, we conducted an application survey (fig. 3 (1)) on "video platform commentary" and "forum discussions", utilizing methodologies adapted from previous academic research [25, 98, 125]. We gathered and categorized data from video platforms like *YouTube*, as well as popular forums like *Reddit* and the *Steam Community*⁵. Based on a basic textual analysis of user comments on these popular platforms, most users express their opinions about (1) how decision-making drives the game's plot, (2) the relationships between characters, and (3) how characters express their personalities and develop through interaction.

We discuss the initial selection of games from a wide range of narratives for further analysis, contrasting the narrative techniques of over 30 well-known games on the market. This resulted in the identification of nine narrative presentation categories: *Relatively Straightforward*, *Visual-Driven*, *Text-Driven*, *Non-Linear*, *Fragmented*, *Lore-Rich*, *Environmental Storytelling*, *Branching*, and *Emergent Narrative*. Two representative games from each category were chosen for further analysis, narrowing the scope to 18 games. After confirming the representativeness and diversity of the aforementioned samples, and using the classical narrative theory of "narrative function analysis"[80], the stories were categorized using the fundamentals of plot, character, and setting. A classification framework that works with our research system was then developed, in which we clearly defined which portions of the story or forum discussions would be analyzed, concentrating on particular character and narrative developments. The next phase is to use the semantic analysis method Latent Dirichlet Allocation (LDA)[7] to thematically model the review data and create a coding framework based on high frequency and centrality. This approach has been validated and summarized in three perspectives: narrative focus based on story progression, character portrayal through gameplay, and interaction patterns characterizing relationships between characters.

While there are likely valid reasons for the various forms of presentation in game commentary, limited research exists on whether particular modes of narration enhance the development of game content. To address this gap, this study presents a unified classification of narrative modalities used across both research and commercial game review platforms. These modalities are categorized into three key areas:

- (1) **Plot Analysis:** This focuses on player choice, environmental storytelling, and non-linear narratives in video games. Narrators often explore key plot points, twists, and how gameplay choices shape the overall narrative.

⁵<https://steamcommunity.com/>

Game	Narrative Style	Main Traits	Examples
 <i>Ori and the Blind Forest</i>	Straightforward, Visual-Driven	Linear story progression, emotional storytelling through environments and cut-scenes	 Environments  Cut-scenes
 <i>Hades</i>	Straightforward, Text-Driven	Dialogues and interactions change each run, revealing more about characters and world	 Dialogues  Character Interactions
 <i>Hollow Knight</i>	Complex, Mixed	Non-linear, exploration-driven narrative with environments, cryptic dialogue and hidden lore	 Environments  Cryptic Dialogues  Hidden Lore
 <i>Elden Ring</i>	Complex, Mixed	Lore-rich, exploration-based narrative through environments, dialogues and item descriptions	 Environments  Dialogues  Item Descriptions

Figure 2: Four games adopted in the workshop using indexical storytelling. *Ori and the Blind Forest* and *Hades* have relatively straightforward narratives, while *Hollow Knight* and *Elden Ring* present more complex and mysterious narratives.

- (2) **Character Relationships:** This examines the dynamic interactions between protagonists, antagonists, and other significant characters. Similar to discussions on emergent narratives in academic research [106], commentators analyze player-driven connections in cooperative settings or intricate character development in story-rich games.
- (3) **Character Portrait:** This delves into how characters evolve and interact with their environment through gameplay [98], assessing how these portrayals contribute to the narrative depth.

These categories represent widely adopted design approaches in game commentary, and we will reference them throughout this study when discussing various narrative design styles.

3.2 Participants and Data

We recruited 14 story interpretation creators (6 females, 8 males; age range 18 – 34) to our workshop through online social platforms. All participants had experience publishing game story interpretations, with 5 specializing in video content, and 9 in both. Each participant was compensated \$10 per hour for their involvement.

Given the absence of systematic classifications, we chose not to provide participants with pre-classified game clues. Instead, we presented them with a list of potential games and asked them to select one to play prior to the workshop. The games included: *Ori and the Blind Forest*, *Hades*, *Hollow Knight*, and *Elden Ring* (fig. 2). Each game invites players to assemble its narrative, albeit in different ways. *Ori and the Blind Forest* [68] and *Hades* present relatively straightforward narratives. *Ori and the Blind Forest* relies on a visual-driven narrative, conveying its story primarily through stunning environments and emotional cut-scenes [65, 86]. *Hades*,

on the other hand, is text-driven, with its narrative unfolding across repeated runs, where character interactions and dialogue evolve as the player progresses [69, 92]. In contrast, *Hollow Knight* [111] and *Elden Ring* offer more mysterious and complex narratives. *Hollow Knight* features a non-linear narrative that encourages exploration and discovery, with its story conveyed through environmental storytelling, cryptic dialogue, and hidden lore, allowing players to piece together the narrative at their own pace [48, 82]. *Elden Ring* employs a fragmented, lore-rich narrative style, where the story is revealed through environmental clues, item descriptions, and brief NPC dialogues [55, 104]. Players are encouraged to explore and interpret the game’s world, piecing together its history and lore through their discoveries.

Participants selected games based on their preferred platforms (e.g., *Windows* and *MacOS*) and personal interest in the gameplay. Their information and selected games are shown in table 1. To ensure a comprehensive narrative experience, participants were required to play their chosen game for at least 4 hours and collect a minimum of 20 narrative-related clues during their gameplay.

3.3 Workshop Setup

To ensure the workshop design was appropriate, two research team members conducted a pilot study prior to the main sessions (fig. 3 (2)). Given the participants’ diverse geographical locations, we organized three virtual workshops using *Zoom*⁶ for video conferencing and *Miro* for collaboration. All necessary materials, including consent forms and pre-workshop guidance, were provided five days before each session, allowing sufficient time for participants to prepare their game clues. (fig. 3 (3)). The guidance included the

⁶<https://zoom.us/>

Table 1: Detailed information of the participants in three workshops.

Workshop	ID	Gender	Age	Platform	Frequency	Game
I	P1	Female	24	Video Platforms & Forums	Monthly	Hades
	P2	Male	22	Video Platforms & Forums	Monthly	Hollow Knight
	P3	Male	20	Video Platforms & Forums	Seasonally	Ori and the Blind Forest
II	P4	Male	22	Video Platforms & Forums	Weekly	Hades
	P5	Female	24	Video Platforms & Forums	Monthly	Hades
	P6	Male	22	Video Platforms & Forums	Monthly	Hades
	P7	Male	21	Video Platforms	Monthly	Hollow Knight
	P8	Female	19	Video Platforms	Monthly	Hollow Knight
III	P9	Female	24	Video Platforms	Seasonally	Ori and the Blind Forest
	P10	Female	24	Video Platforms	Weekly	Hades
	P11	Male	24	Video Platforms & Forums	Seasonally	Hades
	P12	Female	24	Video Platforms	Monthly	Hades
	P13	Male	29	Video Platforms & Forums	Weekly	Elden Ring
	P14	Male	34	Video Platforms & Forums	Weekly	Elden Ring

workshop background, objectives, clue collection instructions, and examples of game story interpretations. To ensure a smooth and efficient workshop experience, we sent the *Zoom* and *Miro* links one day before each session, giving participants time to familiarize themselves with the tools.

3.4 Workshop Procedure

The workshop was divided into three phases: *introduction*, *interpretation*, and *feedback*.

In the *Introduction* phase (fig. 3 (4)), we briefly outlined the workshop structure, provided examples for story interpretation, and explained the tasks for the *Interpretation* phase (fig. 3 (4a)). Since participants had already reviewed the pre-workshop guide, this phase was brief, lasting about 10 minutes. During the *Interpretation* phase (fig. 3 (5)), participants analyzed their narrative clues from three perspectives: *plot analysis*, *character relationships*, and *character portrait*. They could work digitally on *Miro* or sketch on paper and upload their work. Participants followed a think-aloud protocol (fig. 3 (5a)) to articulate their thought processes, enabling us to capture their classification preferences, interpretive strategies, and design decisions (fig. 3 (5b-c)). This phase lasted about 70 minutes. In the *Feedback* phase (fig. 3 (6)), each participant presented their interpretation for 2 minutes to the group (fig. 3 (6a)), followed by 15-minute one-on-one semi-structured interviews (fig. 3 (6b)). The interviews were conducted after all workshops, allowing participants to observe others' approaches and methods, which enriched the discussion. We focused on three main topics: (1) clue classification strategies; (2) interpretation perspective preferences; and (3) desired tool functionalities for classification and interpretation.

3.5 Data Analysis

We employed a systematic approach to analyze both the interview recordings and participant sketches, ensuring the thoroughness and validity of our findings. Two team members were deeply involved in the analysis process.

We conducted an inductive thematic analysis [11] for the interviews. All workshop interviews were recorded, yielding approximately six hours of audio data. The analysis began with automatic transcription using *Otter.ai*⁷, followed by a thorough manual review to ensure accuracy, including nuances such as tone and pauses. Two researchers independently conducted a preliminary coding exercise, identifying and defining codes based on the data. They then convened to discuss their individual codes and collaboratively developed a unified coding framework, which was applied extensively. The researchers independently grouped the codes into potential themes, which were further refined through discussions to establish a clear and coherent set of themes. To complement the interview data, the same process was applied to the sketches: the researchers independently coded the sketches, identifying key visual elements and patterns. After discussing and reconciling their codes, they developed a consistent set of themes for the sketch data.

We then conducted a cross-referencing analysis [2] by comparing the interview themes with those derived from the sketches, examining how visual elements aligned with or expanded upon the participants' verbal accounts. Following the principle of saturation [89], the codebook was developed using data from the first two workshops. The inclusion of sketches helped validate the interview themes, and by the third workshop, no new themes emerged, indicating that saturation had been achieved.

3.6 Findings

Based on the data analysis, we organized our findings into two sections: *Challenges* and *Design Requirements*.

3.6.1 Challenges. We categorized the six challenges into three distinct areas: *Clue Classification and Organization*, *Clue Understanding*, and *Story Interpretation* in table 2.

Classification and Organization. Participants unanimously agreed that the classification and organization of game clues were essential first steps in interpreting game stories. However, they

⁷<http://Otter.ai>

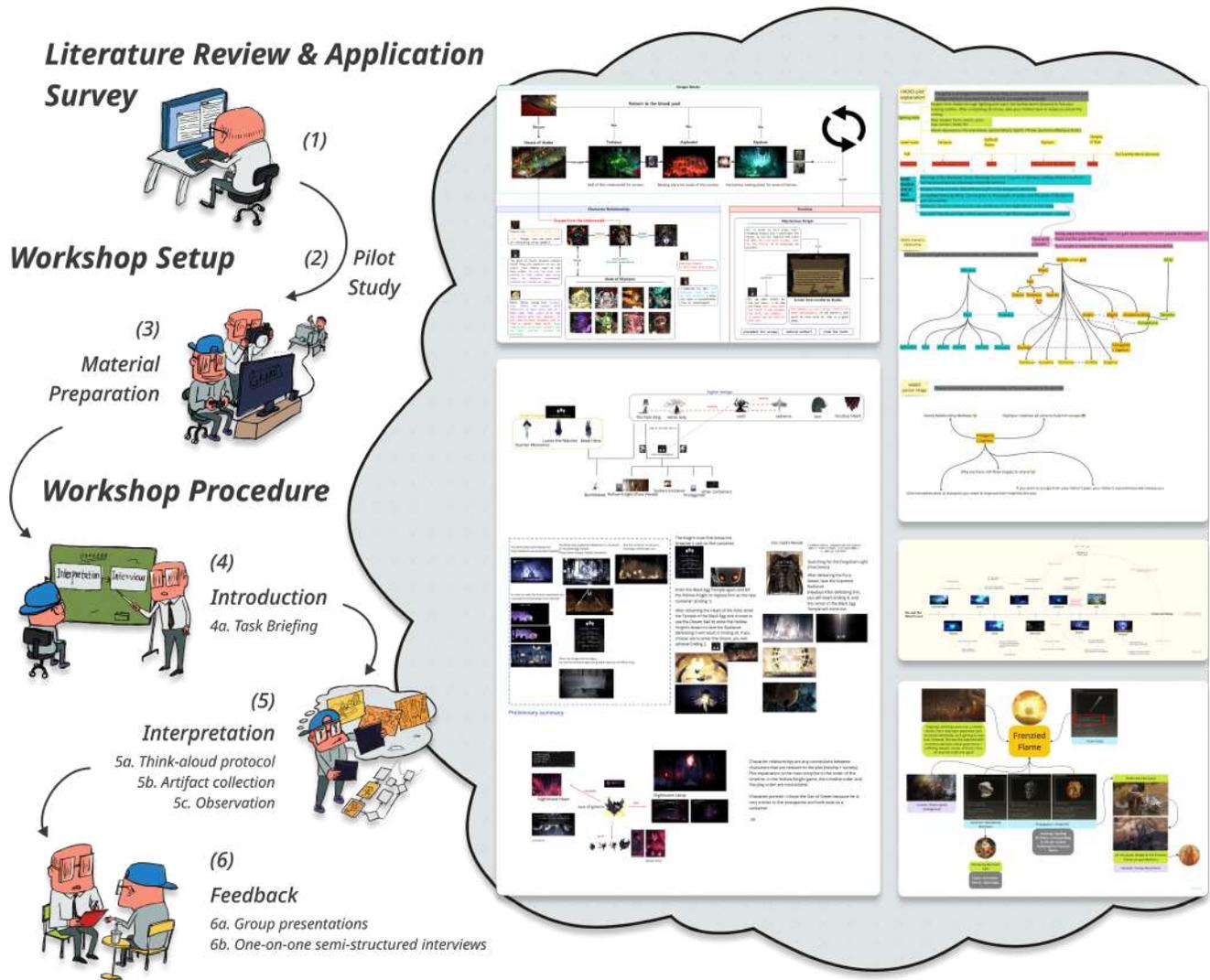


Figure 3: Formative Study procedure. *Literature Review & Application Survey:* (1) Conduct a literature review and application survey to gather three narrative modalities in game story interpretation. *Workshop Setup:* (2) Perform a pilot study to assess appropriateness. (3) Request participants to prepare game clues. *Workshop Procedure:* (4) Introduce the research background and workshop process to participants. (5) Ask participants to complete a game story interpretation task. (6) Interview each participant to gather feedback.

encounter significant challenges in this process. One major difficulty lies in maintaining a structured classification due to the large volume of clues and the incremental nature of their collection (C1). Frequent reclassification after each collection session is time-consuming and tedious, leading to a reluctance to keep things organized. As P1 observed, “When there is only a small amount of clues, not sorting it in advance does not really impact how you find [things] later. But when the volume gets large, it can quickly become overwhelming to manage.”

Additionally, current methods for organizing clues are inadequate, primarily relying on chronological order, which does not align with the non-linear exploration typical in games (C2). This

misalignment hinders effective story interpretation, as chronological organization fails to reflect the logical structure needed. P9 shared, “Sometimes I explore the same place more than once. Like in *Ori and the Blind Forest*, I passed through Hollow Grove twice—once in the main story, and later to explore [more]. If I organize everything chronologically, clues from both visits get mixed with other areas, making it hard to focus on Hollow Grove. That’s why sorting clues by location is key [in cases] like this.”

Clue Understanding. Participants identified significant challenges in understanding the game clues they collected, which hindered their ability to efficiently interpret the stories. C3 highlights

Table 2: Identified challenges and corresponding design requirements.

Categories	Challenges	DRs
Classification & Organization	C1. Maintain structured classification of clues (13/14).	DR1
	C2. Organize clues efficiently (12/14).	DR1
Clue Understanding	C3. Ease the cog. load of comprehending clue content (10/14).	DR2
	C4. Understand the relationship between clues intuitively (9/14).	DR3
Story Interpretation	C5. Select relevant clues (14/14).	DR4
	C6. Integrate clues with external information easily (10/14).	DR5

the difficulty of reducing the cognitive load associated with comprehending clues content. Participants often collect clues in formats like images and videos, requiring extra effort to identify content relevant to the story. This challenge was particularly evident in text-heavy games like *Hades*, where the story is primarily conveyed through dialogue and text. P5 explained, “*Pulling out keywords from such a huge amount of text is really tough for me. It’s like in another text-heavy game, Disco Elysium—there’s just so much [text] that I often feel overwhelmed, and it makes it hard to figure out the main story.*”

Another challenge, C4, involves intuitively understanding the relationship between clues. P13 emphasized the cognitive burden of recalling and connecting them: “*Sometimes I know that two items or pieces of dialogue are connected, but it’s hard to track that [connection] across all the content I’ve collected. This is especially challenging in games where the story elements aren’t directly linked but are scattered around, and you have to piece them together [yourself].*” The participant added, “*In games like Hollow Knight or Elden Ring, where the story is scattered across different clues, figuring out how everything connects often means piecing together subtle hints. It can [get] overwhelming without some kind of system to help you spot those connections.*” This reflects a broader issue in games with indexical storytelling: creators may struggle to synthesize disconnected narrative elements without guidance.

Story Interpretation. Participants faced significant challenges in the story interpretation process, particularly in selecting relevant clues (C5). The large volume of clues and the long time span over which they are collected can easily result in oversights during story interpretation. As P14 noted, “*The scale of Elden Ring is overwhelming, and the scattered story clues are so subtle that I often miss key details when trying to piece the narrative together.*” This challenge is particularly common in expansive, non-linear games where the narrative is fragmented and requires players to actively collect and connect dispersed clues.

Participants also emphasized the importance of integrating clues with external information (C6). As P12 noted, “*When I’m interpreting game stories, finding connections between the game’s clues and real-world history, mythology, or cultural references gives me new insights and helps me think more creatively about how this outside information can enhance the story.*” P7 echoed this sentiment, stating, “*I always pay close attention to place names and architectural styles. These designs aren’t just randomly made up by the game designers; they often pull from real-world references. This kind of information is super helpful for understanding the story.*”

3.6.2 Design Requirements. Based on these challenges, we identified the following design requirements to assist creators in interpreting stories more effectively.

DR1 Implement adaptive and customized classification and organization. Structured classification of clues is crucial for enhancing the efficiency of interpreting and reconstructing game narratives. Systematic categorization, such as grouping narrative elements by type or location (C1), allows for quick access to critical information. Additionally, flexible organization (C2)—whether by time, theme, or relevance—supports the seamless assembly of complex storylines, reducing cognitive load and promoting a more intuitive workflow.

DR2 Provide clue content explanations. Understanding complex or unfamiliar clues can be a significant challenge during the interpretation process (C3). To address this, the tool should offer clear explanations of each clue, including its content and keywords, to convey essential information swiftly. This feature would enable more efficient comprehension and utilization of clues in story interpretation.

DR3 Provide clue relationship explanations. Creators often need to understand how different clues are interconnected to effectively interpret complex narratives (C4). The tool should offer clear explanations of the relationships between various clues, such as how a specific dialogue relates to an item or how different scenes are linked through recurring themes. By highlighting these connections, creators can more easily piece together the broader story and avoid overlooking critical narrative links.

DR4 Enable quick retrieval and personalized recommendation of related clues. Swift access to relevant clues is vital for maintaining the flow of the interpretation process (C5). The tool should support quick retrieval by allowing users to efficiently search for and access related clues. Additionally, recommending clues based on context or user inputs can help uncover connections that might otherwise be missed, ensuring a comprehensive and cohesive interpretation.

DR5 Integrate external information seamlessly. Access to relevant external information can greatly enhance creators’ interpretations of game narratives (C6). The tool should facilitate the seamless integration of external sources, such as historical references, mythological contexts, or cultural details, directly into the interpretation process. Providing this additional context allows creators to deepen their understanding of narrative clues and develop more informed and engaging interpretations.

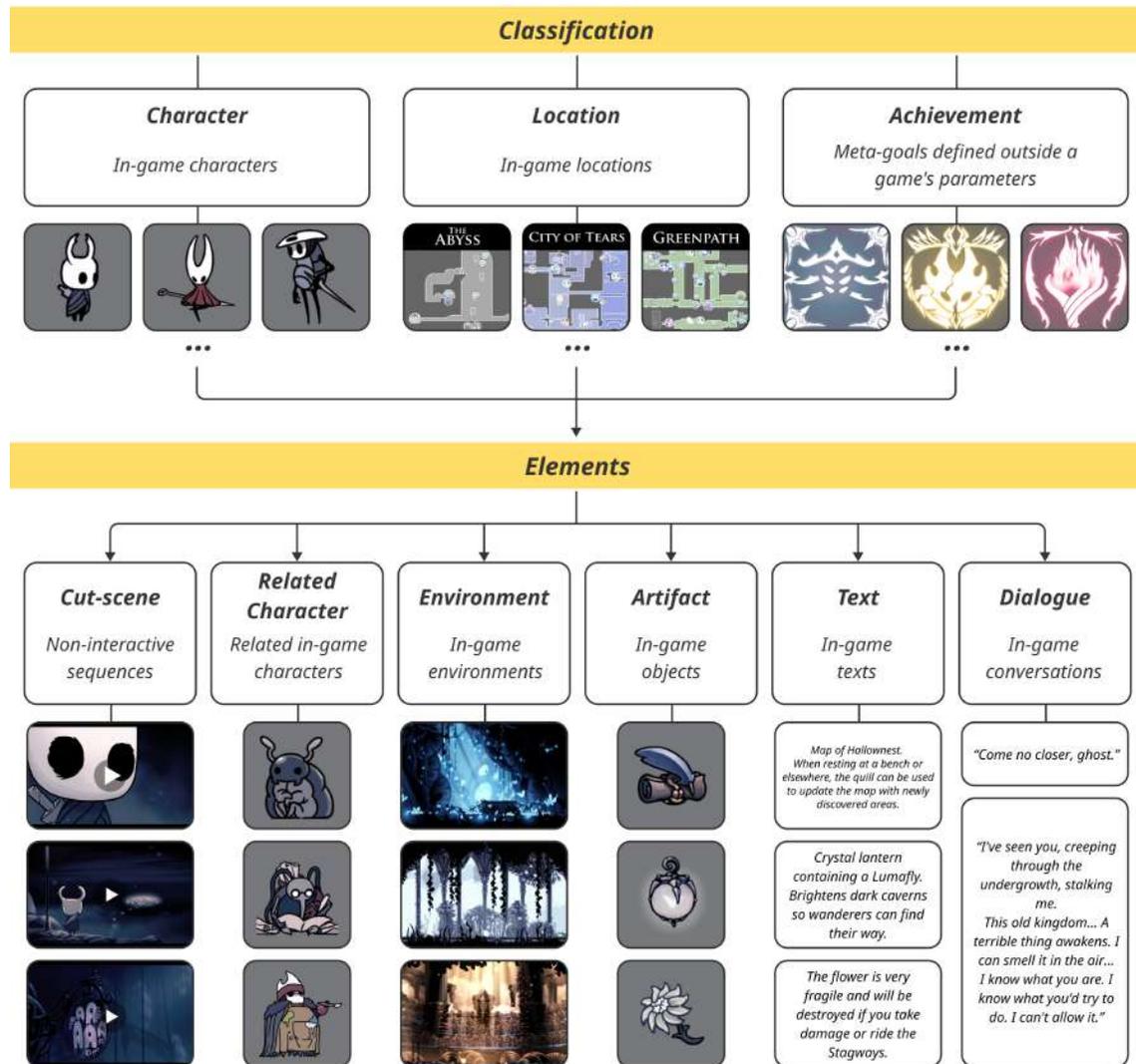


Figure 4: The hierarchical taxonomy, exemplified by the game *Hollow Knight*, is composed of two levels. The first level, *Classification*, categorizes clues into three groups: *Character*, *Location*, and *Achievement*. The second level, *Elements*, further classifies these based on in-game attributes into six types: *Cut-scene*, *Related Character*, *Environment*, *Artifact*, *Text*, and *Dialogue*. Each *Classification* category can include any type of *Elements*.

4 Taxonomy

To address RQ1, we developed a two-level hierarchical taxonomy based on the insights collected from the workshop. This taxonomy was informed by both the clues participants identified during gameplay and the insights from semi-structured interviews conducted in the co-design workshop. The clues provided concrete examples of how participants engaged with the game's narrative, while the interviews shed light on their classification strategies and reasoning. By integrating behavioral data (from the clues) and cognitive insights (from the interviews), the taxonomy offers a user-centered and flexible framework that accommodates diverse interpretations of game stories. This approach captures both what participants did and the rationale behind their interpretive choices.

The process, led by two team members deeply involved in the analyzing process described in section 3.5, involved two key stages. First, we inductively categorized the clues into broad thematic groups based on participants' input through open coding of both clues and interview data. Three major themes emerged from participants' approaches: **Character**, **Location**, and **Achievement**. The two team members independently coded the data, then convened to discuss and develop a unified coding framework. In the second stage, we deductively identified specific narrative elements within each theme, including **Cut-scene**, **Related Character**, **Environment**, **Artifact**, **Text**, and **Dialogue**, based on their role in advancing the plot, conveying emotion, or providing contextual information. The

interviews revealed additional nuances in how participants prioritized or combined these elements, further refining the taxonomy.

After categorizing and defining these elements, we organized the taxonomy into a two-level hierarchy. The first level, **Classification**, outlines the three common ways creators categorize game clues. The second level, **Elements**, identifies six in-game inherent attributes of these clues. It is important to note that clues classified under any category in *Classification* may encompass all types of elements outlined in *Elements*.

4.1 Classification

When interpreting game stories, creators typically categorize clues retrospectively based on their gameplay experiences. We identified three key dimensions defining their categorization habits: (1) Character, (2) Location, and (3) Achievement.

Classification-1: Character. In games, characters include both player-controlled and non-player characters (NPCs). Categorizing clues by character is crucial for retrieving content related to character portraits and relationships. This organization helps track character development, interactions, and their role in the overall narrative, ensuring consistency in portrayal and supporting a cohesive narrative analysis.

Classification-2: Location. Locations refer to specific areas within the game world, often tied to unique storylines or events. Categorizing clues by location helps maintain geographical coherence and allows creators to analyze how the narrative unfolds across different environments, revealing the connection between setting and narrative progression.

Classification-3: Achievement. Achievements are meta-goals linked to key in-game milestones or challenges. Categorizing clues by achievement is valuable because certain achievements relate to critical plot points. This helps map the narrative's progression and analyze how story elements are unlocked through player actions.

4.2 Elements

We identified six fundamental attributes of the assets collected by creators within the game, each contributing uniquely to the narrative structure: (1) Cut-scene, (2) Related Character, (3) Environment, (4) Artifact, (5) Text, and (6) Dialogue.

Element-1: Cut-scene. Cut-scenes are non-interactive sequences that advance the story and provide key plot points and character development within a game. By suspending player control, they deliver the narrative with cinematic precision, ensuring intended emotional and narrative impact.

Element-2: Related Character. Related characters are tied to specific attributes in the game and drive the narrative through interactions with the player, offering quests and revealing critical story elements. They guide or challenge the player, shaping the player's journey through dialogue and interactions.

Element-3: Environment. Environments set the tone, mood, and atmosphere of a game, acting as narrative tools that reflect the story's progression. Environments provide visual storytelling cues, such as remnants of past events or specific objects that hint at the game's lore, influencing how players perceive and engage with the narrative and making them a vital element in narrative immersion.

Element-4: Artifact. Artifacts are objects connected to the game's lore or history, enriching the story through exploration. These objects may be directly tied to the plot, such as an ancient weapon or a diary, or they might offer background information that enriches the world-building. They provide deeper narrative context and reward players with a richer understanding of the game world.

Element-5: Text. In-game texts are a primary source of narrative exposition and world-building. They offer background information, character insights, and historical context, allowing players to explore the narrative at their own pace and adding depth to the story.

Element-6: Dialogue. Dialogues are dynamic narrative components that drive character development, convey emotions, and advance the plot. Players can gain insights into characters' motivations and personalities, as well as the story's underlying conflicts and themes. Dialogue choices can influence the direction of the narrative, offering players agency in how the story unfolds and deepening their connection to the game's characters and events.

5 ClueCart

Guided by the synthesized design requirements and taxonomy, we have developed *ClueCart*, an interactive creativity support tool designed to assist creators in interpreting and understanding game narratives. In the following paragraph, we will first introduce Automatic Classification section, followed by the detail functionalities of *ClueCart* interface.

5.1 Automatic Classification

ClueCart's core functionality relies on the automatic classification of game clues according to the predefined taxonomy (**DR1**). This structured approach enhances creators' ability to analyze and interpret game narratives, revealing the intricate relationships between various game elements. To support this classification, we developed an in-game mod that automatically captures three primary categories of clues (i.e., character, location, and achievement). LLMs are then employed to further classify six additional elements, providing descriptions and extracting keywords from corresponding clues to fulfill **DR2**.

5.1.1 Capturing Three Classifications. Video game modifications (mods) [79] are user-generated enhancements that expand the functionality of existing games, offering players and creators new interaction possibilities. In HCI research, modding is crucial for exploring how collaborative and participatory design within gaming communities fosters innovation and enhances user experience [96, 123]. We developed a custom mod using the *BepInEx*⁸ framework, a widely-used plugin system for *Unity*-based games. This mod integrates seamlessly into the game, enabling creators to efficiently capture and classify game content. We implemented two intuitive methods based on creators' workflows: **1) Screenshots:** Capture with a hotkey; **2) Screen Recording:** Started and stopped with another hotkey.

When a capture function is activated, the mod automatically accesses the game's main and UI camera frustums to check for visible characters or locations. If detected any of these game objects, the mod automatically logs their names for subsequent analysis.

⁸<https://github.com/BepInEx/BepInEx>

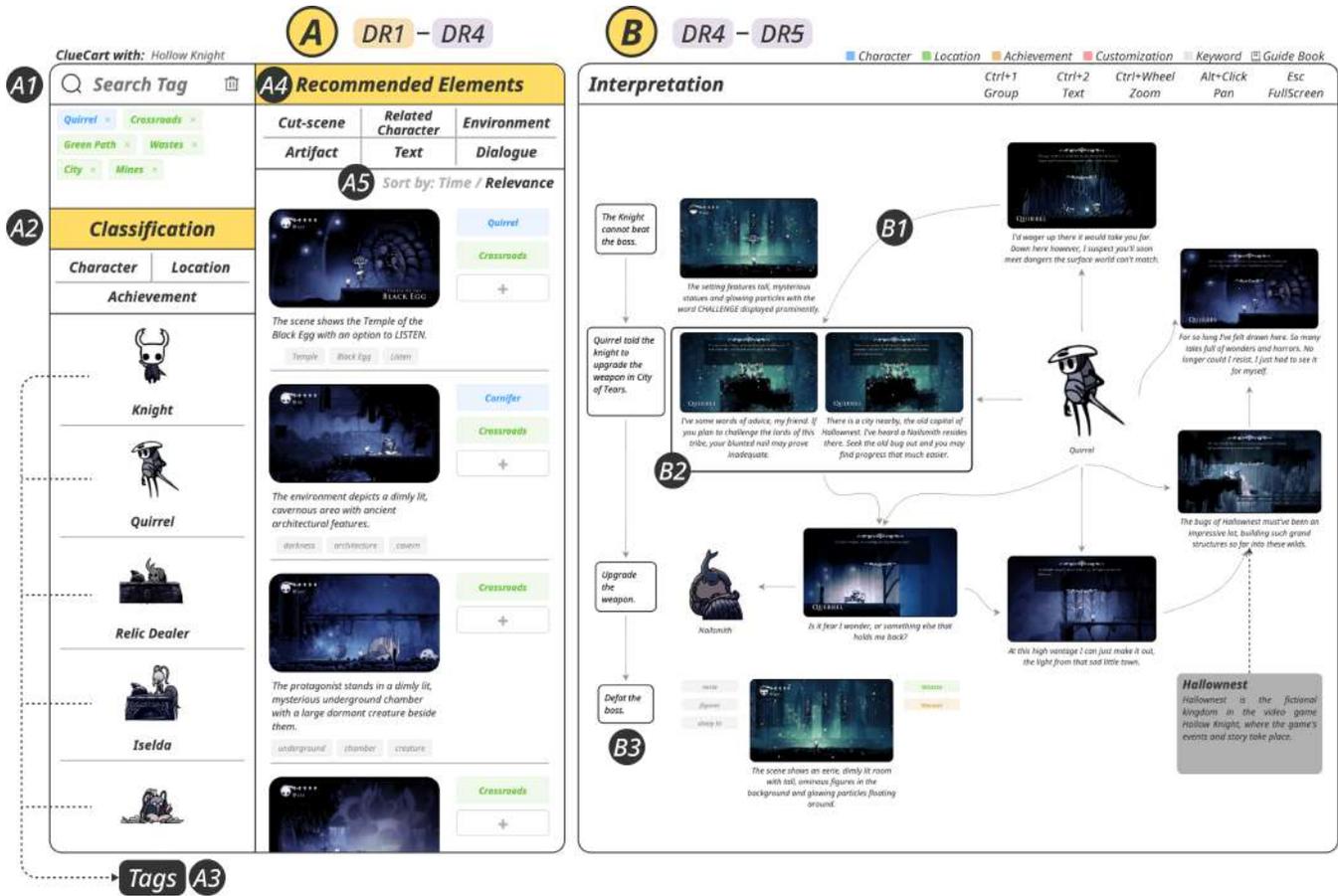


Figure 5: ClueCart Interface. (A) Left panel for Clue Categorization and Retrieval. (B) Right panel for story interpretation.

Additionally, our mod integrates with *Steam's achievement system*⁹, known for its comprehensive tracking capabilities. When an in-game achievement is unlocked, the mod automatically captures a screenshot, logging the current characters, locations, and the specific achievement. This ensures key game moments are documented for further analysis.

It is important to note that our mod is built with rigorous privacy safeguards. It neither interferes with the game's proprietary data nor collects, stores, or transmits any personal player information. It only captures in-game clues relevant to the study, ensuring both game integrity and privacy protection.

5.1.2 Categorizing Six Elements and Providing Descriptions. To enhance classification, we integrated GPT-4o with image interpretation capabilities. After capturing visual data, the LLM categorizes six specific game elements, generates descriptions, and extracts keywords, helping creators understand the clues (DR2). This improves the accuracy and nuance of the classification, aligning with the predefined taxonomy. To facilitate these tasks, we carefully and iteratively crafted prompts to ensure relevance and consistency. Details of the prompt engineering are provided in appendix A.

5.2 Interface

Based on our design requirements and taxonomy, we developed an open-source interactive interface¹⁰ to assist creators in analyzing and reconstructing game narratives using collected game clues. Additionally, the interface includes a built-in system guide¹¹ to help users quickly get started. In response to RQ2, we examine how these interface features support creators throughout their workflows. The interface is divided into two primary components, each tailored to help creators categorize and retrieve game clues and interpret game stories.

5.2.1 Clue Categorization and Retrieval. The left side of the interface (fig. 5(A)) is designed to assist creators in categorizing, understanding, and retrieving game clues based on the hierarchical taxonomy. The interface is divided into three main sections: **Retrieval** (fig. 5(A1)) and **Classification** (fig. 5(A2)) on the left and **Recommended Elements** (fig. 5(A4)) on the right. Each section is specifically tailored to support different aspects of the creators' workflow, aligning with our design requirements (DR1-DR4).

⁹<https://store.steampowered.com>

¹⁰<https://cluecart.github.io/ClueCart/>

¹¹<https://tasty-trick-667.notion.site/cluecart-guide-book>

The **Classification** column on the left presents clues that have been processed and categorized according to the first level of the taxonomy. These are organized into three classifications: *Character*, *Location*, and *Achievement*. Each acts as a tag applied to relevant clues (fig. 5(A3)), providing a structured overview. This tagging system enables creators to quickly identify and group related clues, streamlining organization and retrieval (DR4).

The right column, labeled **Recommended Elements**, further organizes game clues into six elements: *Cut-scene*, *Related Character*, *Environment*, *Artifact*, *Dialogue*, and *Text*. Each group includes relevant clues (fig. 6), accompanied by descriptions, keywords (fig. 6(A1)), and tags (fig. 6(A2)). Descriptions and keywords are automatically generated for each clue to help creators quickly grasp its content and context, aligning with DR2. Tags next to each clue highlight clue relationships, directly aiding in understanding these connections as outlined in DR3.

To enhance usability, we integrated several features that support customized classification and efficient retrieval. The **Retrieval** function in the upper-left corner (fig. 5(A1)) offers robust search capabilities via tags. Creators can drag tags to adjust the priority of recommended node tags, prompting more personalized suggestions (DR4). In the **Recommended Elements** section (fig. 5(A4)), each clue node incorporates two customization features for tags and keywords (DR1). A “+ tag” option (fig. 6(B1)) lets users create and apply custom tags for custom classification. Additionally, as the LLM performance in word extraction can be inconsistent, creators can edit keywords to better summarize the descriptions (fig. 6(B2)). Furthermore, a sorting feature that arranges clues by temporal order and tag relevance (fig. 5(A5)) helps creators view clues within context and enhancing their understanding of clue relationships (DR3).

5.2.2 Story Interpretation. The **Interpretation** panel (fig. 5(B)) provides creators with a flexible workspace for organizing and analyzing collected clues, facilitating the reconstruction of the story’s narrative. Creators can drag clues directly from the **Recommended Elements** column onto the canvas, where they can engage with and manipulate these clues using various features designed to support reasoning and interpretation.

Interactive Node Features. When creators interact with individual clue nodes, several features enhance their analysis. Clicking on a node with the left mouse button displays a bounding box (fig. 7(A)), allowing creators to resize the clue as needed. Right-clicking on a node reveals keywords on the left and corresponding tags on the right (fig. 7(B)), providing quick access to contextual information. Clicking on a keyword (fig. 7(B1)) prompts (see appendix B) *ClueCart* to use an LLM to search for its real-world meaning, placing the result in a new node on the canvas, thus seamlessly integrating external information into their analysis (DR5). Additionally, clicking on a tag (fig. 7(B2)) initiates a search in the **Retrieval** section, and double-clicking a clue node (fig. 7(B3)) adds all its tags to this field, triggering the system to recommend related clues from the **Recommended Elements** list, in line with DR4.

Establishing Relationships Between Nodes. Beyond individual node interactions, the interface facilitates reasoning about the relationships between clues through a node-linking feature (fig. 5(B1)). This functionality enables creators to draw connections between

different nodes, visually mapping the relationships they identify. By double-clicking on a link, creators can add textual annotations, which enhances their analysis of how clues interrelate and contributes to a more thorough narrative reconstruction.

Customization and Organizational Features. To further enhance the reasoning process, the **Interpretation** panel offers customization options such as group nodes and text nodes. Group nodes (fig. 5(B2)) enable creators to cluster related clues, ensuring a clear and organized structure that simplifies the navigation of complex narratives. Text nodes (fig. 5(B3)) allow creators to document their thoughts, hypotheses, and insights directly on the canvas, facilitating ongoing reflection and interpretation as they refine their understanding of the story.

6 Evaluation

We conducted a between-subjects study (N=40) to evaluate *ClueCart* against the baseline tool, *Miro*.

6.1 Baseline and Data Settings

To evaluate the impact of *ClueCart* on creators’ story interpretation efficiency (RQ3), we used *Miro* as the baseline for comparison. *Miro* was selected due to its widespread use among story interpreters, thanks to its flexibility, ease of use, and robust support for visual storytelling [6, 47, 52, 81]. These features made *Miro* an ideal baseline for evaluating interactive story interpretation across different media.

In our study, we selected two games utilizing indexical storytelling: *Hollow Knight* [111] and *Do Not Feed the Monkeys (DNFTM)* [35]. The choice of these games was informed by feedback from our Formative Study (section 3), which indicated that complex gameplay in previous games hindered narrative interpretation. To address this, we included *Hollow Knight* for its rich environmental storytelling and added *DNFTM*, a puzzle-based game with simpler controls, to minimize cognitive load and enhance participants’ focus on story analysis. To maintain consistency and control across participants, we pre-collected all game-related clues instead of asking participants to collect independently. Given the extensive time required to uncover key narrative elements in *Hollow Knight*, gathering clues in advance ensured that all participants had access to the same data, facilitating a fair comparison. The collected clues, adhering to our standard gameplay procedure, included all six elements in the taxonomy, ensuring both input consistency and ecological validity.

6.2 Participants

We recruited 40 participants (18 females, 22 males, aged 20-29) through social media channels (appendix C). All participants had prior experience with discussing and interpreting game stories: 19 through video platforms, 1 on online forums, and 20 through a combination of both. In addition, one participant is a contributor of a game wiki. These players, who are also creators of game story interpretations, were randomly assigned to one of two groups of 20 people each: one group analyzed *Hollow Knight* and the other *DNFTM*. Each group was further split into two subgroups of 10, with one subgroup using *Miro* for story interpretation and the other using *ClueCart*. This between-subjects design was employed

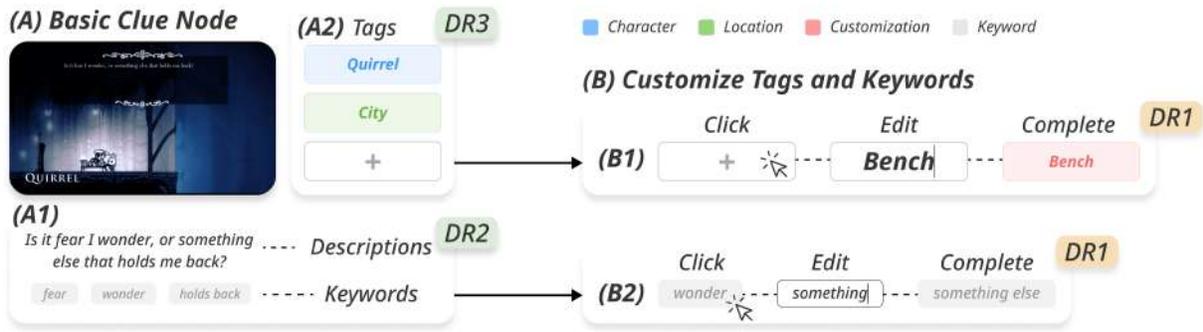


Figure 6: Clue Node Design and Features: (A) Node design with (A1) tag list and (A2) descriptions and extracted keywords. (B) Tag / keyword customization via (B1-2) clicking and editing.

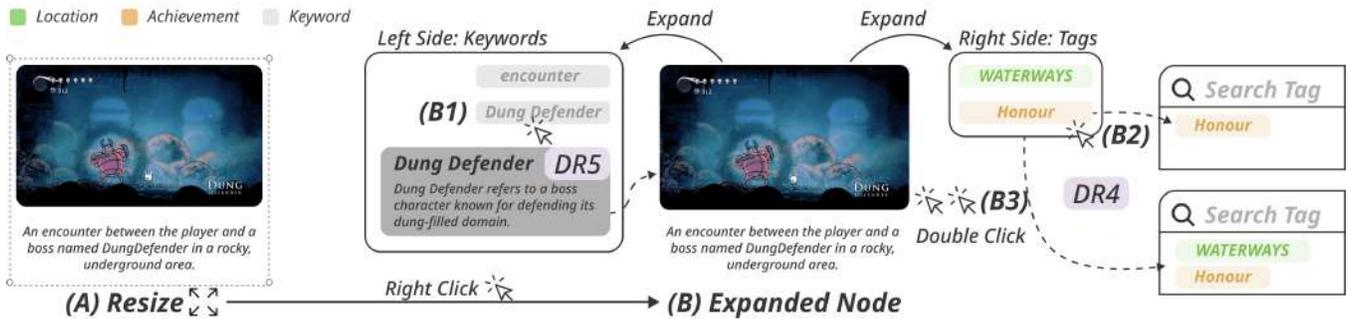


Figure 7: Interactions of a clue node. (A) Drag the bounding box to resize the node. (B) Expand the node by right clicking. (B1) Click on any keyword on the left to search for its real-world meaning. (B2) Click on a tag on the right to include it in the search query. (B3) Double-click on the clue node to add all associated tags to the search.

to address the substantial disparity in gameplay difficulty between *Hollow Knight* and *DNFTM*. To assure familiarity with the game stories and reduce potential learning curve, all participants had previously played through the games before the study. Participants were compensated \$10 per hour.

6.3 Evaluation Dimensions

We adopted three key dimensions for quantitative evaluation: **Perception**, **Features**, and **Story Interpretation Assessment**, as shown in appendix D.

The **Perception** dimension captured participants' subjective experiences, focusing on usability and creativity support. Two widely adopted metrics were used: *System Usability Scale* (SUS) [5] and *Creativity Support Index* (CSI) [22]. The SUS assessed the overall usability of the tools, while the CSI focused on how well the tools supported participants' creative processes. Both of them are commonly used in evaluation of creativity support tools [40, 110], offering reliable insights into participants' subjective impressions and satisfaction with the tool.

The **Features** dimension evaluated *ClueCart*'s specific functionalities, aligning with the categories, challenges (Cs) and design requirements (DRs) outlined in table 2. This dimension was further divided into three aspects: *taxonomy*, corresponding to Classification & Organization (C1, C2; DR1); *functionality*, addressing Clue

Understanding (C3, C4; DR2, DR3) and Story Interpretation (C5, C6; DR4, DR5) and *interaction*, reflecting participants' impressions of *ClueCart*'s overall interactions. This dimension is specifically evaluated by *ClueCart* group.

The **Story Interpretation Assessment** dimension focused on participants' performance in reconstructing narratives, with two key metrics: *task completion time*, measuring the time taken to complete the task, and *peer scoring*, which used a concept map evaluation checklist¹² to assess outputs based on clarity, organization, and logical structure. Although *Miro* outputs may not conform to a traditional concept map format, this framework was effective for assessing key aspects such as clarity, organization, and logical structure. This standardized evaluation ensured that consistent assessment of how accurately participants represented the game's narratives.

To complement these quantitative methods, we conducted **semi-structured interviews** to gather qualitative feedback on participants' experiences, providing deeper insights into strengths, weaknesses, and improvement areas. By integrating quantitative and qualitative methods, this evaluation ensured a rigorous and holistic assessment of *ClueCart*'s usability, functionality, and narrative support capabilities.

¹²<https://www.morgan.edu/Documents/ADMINISTRATION/DIVISIONS/IT/ConceptMap-EvalChecklist.pdf>

6.4 Study Setup

To assess the feasibility and effectiveness of our study procedure, two members of our research team conducted a pilot study, each focusing on one of the games. This process helped refine the study design and ensure that tasks were appropriately structured to align with the narrative style of each game.

Prior to each session, participants were provided with a consent form, background information, and an overview of the research goals to ensure they were well-informed. Additionally, we distributed a pre-collected game material package, which included detailed explanations of the clue collection process, data format, and organizational structure. This preparation facilitated a smooth progression of the study. For participants in the *Miro* group, we provided a tutorial on how to use the tool, while those in the *ClueCart* group received a system guide and instructional video to help them understand the interface.

6.5 Study Procedure

The formal study was divided into four parts: **introduction**, **story interpretation**, **evaluation and peer scoring**, and **feedback**.

6.5.1 Introduction (20 min for Miro group, 25 min for ClueCart group). In the overall introduction part (15 min), we explained the whole study procedure, offering a sequential overview and estimated duration for each segment. We introduced the research background and objectives, illustrating how to analyze the game narratives from three key perspectives: *plot analysis*, *character relationships*, and *character portrait*. We also reviewed the game clues provided the prior day, outlining their collection during gameplay and elucidating the selection criteria to guarantee the inclusion of all relevant narrative elements, so providing participants with a thorough comprehension of the dataset.

Following this, tool-specific orientations were given to each subgroup to ensure participants got only material relevant to their designated tool. The *Miro* group received a brief introduction (5 min) on the tool's visual collaboration features, emphasizing its application in narrative analysis mapping. The *ClueCart* group was given a comprehensive review (10 min) of the two-level taxonomy derived from the formative study, along with a demonstration of how to use the *ClueCart* interface to organize and connect narrative fragments. This preparation ensured all participants were familiar with their respective tools before starting the task.

6.5.2 Story Interpretation (60 min). Following the introduction, participants proceeded to the story interpretation task, where they analyzed game clues from the perspectives of *plot analysis*, *character relationships*, and *character portrait*. The time taken by each participant to complete the task was recorded and subsequently used to assess the efficiency of the tools in supporting narrative interpretation.

6.5.3 Evaluation and Peer Scoring (15 min). After completing the task, participants filled out a 7-point Likert scale questionnaire, assessing the *Perception* dimension, including SUS and CSI, based on their tool experience. In particular, we incorporated questions about specific *Features* of *ClueCart*, focusing on taxonomy, functionality, and interactions for the *ClueCart* group. For participants' *Story Interpretations*, we recorded their completion time and included

peer scoring in the questionnaire to evaluate clarity, organization, and logical structure of the interpretations.

6.5.4 Feedback (20 min). Following the scoring phase, we conducted semi-structured interviews (20 min) with each participant to gather detailed feedback on their experience with the tool. During the interviews, participants were first asked to describe their ideation process throughout the task, followed by their overall experience with the tool they used. We then revisited the taxonomy introduced during the study to gather feedback and suggestions for enhancement. Finally, we examined participants' perspectives on the future of game story interpretation. These qualitative insights complemented the quantitative data from the questionnaire, offering a deeper understanding of the strengths and limitations of both *Miro* and *ClueCart* in facilitating interactive story interpretation.

7 Results

We collected participants' ratings on the three main dimensions mentioned in section 6.3, including **Perception**, **Features** and **Story Interpretation Assessment**. All ratings were collected using a 7-point Likert scale.

Quantitative data for perception and story interpretation assessment was analyzed using the Mann-Whitney U Test [64] to compare *ClueCart* with the baseline tool, *Miro*. Ratings for features of *ClueCart* were summarized descriptively without comparative statistical analysis. For the qualitative analysis, interview transcripts were coded using inductive thematic analysis [11], with two researchers cross-checking to ensure consistency. This mixed-methods approach allowed us to identify key themes and patterns in participants' feedback, complementing the quantitative findings.

7.1 Perception of ClueCart

The perception of *ClueCart* was assessed across multiple aspects, including system usability (SUS) and creativity support (CSI), with results compared with baseline *Miro*. Statistical analyses determined the significance of differences between the two tools, with a threshold set at $p < 0.05$. Specific levels of significance are indicated in table 3, where * represents $p < 0.05$, ** represents $p < 0.01$, and *** represents $p < 0.001$. Means (M) and standard deviations (SD) are reported for each metric. The results demonstrated statistically significant improvements over the baseline system across these dimensions. Detailed questions for these aspects are provided in appendix F.

7.1.1 System Usability. The system usability [5] of *ClueCart* was assessed through various measures including *ease of use*, *functionality*, *learning curve*, *frequency of use*, and *user confidence*, as summarized in table 3 and appendix E.

Participants rated *ClueCart* as significantly easier to use ($M = 6.10$, $SD = 0.59$) than the baseline system ($M = 5.15$, $SD = 0.70$). Similarly, functionality was evaluated as having more comprehensive and better-integrated features ($M = 6.15$, $SD = 0.64$) than the baseline system ($M = 4.07$, $SD = 1.07$). Learnability was also rated significantly higher for *ClueCart* ($M = 6.12$, $SD = 0.83$) compared to the baseline system ($M = 4.87$, $SD = 0.92$). Additionally, *ClueCart* was perceived to have a significantly lower learning curve

Table 3: The statistical user feedback with *Baseline* and *ClueCart*, where the *p*-values (*: $p < .050$, **: $p < .010$, *: $p < .001$) is reported.**

Categories	Factors	Baseline Mean/S.D.	ClueCart Mean/S.D.	<i>P</i> -value	Sig.
System Usability Scale	Easy to use	5.15/0.70	6.10/0.59	1.79×10^{-4}	***
	Functions	4.07/1.07	6.15/0.64	6.92×10^{-7}	***
	Quick to learn	4.87/0.92	6.12/0.83	2.80×10^{-4}	***
	Learning curve	3.20/1.00	1.70/0.80	8.00×10^{-5}	***
	Frequency	3.65/1.24	6.07/0.65	7.95×10^{-7}	***
	Confidence	4.57/1.20	6.18/0.64	2.45×10^{-5}	***
Creativity Support Index	Enjoyment	4.05/1.32	6.10/0.86	1.50×10^{-5}	***
	Exploration	3.93/1.29	6.28/0.63	4.26×10^{-6}	***
	Results Worth Effort	4.12/1.63	6.02/1.04	4.20×10^{-4}	***
	Expressiveness	3.67/1.18	5.92/1.02	1.50×10^{-5}	***
	Immersion	3.55/1.19	6.12/0.86	6.92×10^{-7}	***

($M = 1.70$, $SD = 0.80$) than the baseline ($M = 3.20$, $SD = 1.00$), indicating its ease of adoption and user-friendliness. Participants also indicated that they would use *ClueCart* more frequently ($M = 6.07$, $SD = 0.65$) than the baseline system ($M = 3.65$, $SD = 1.24$) and felt greater confidence while using it ($M = 6.18$, $SD = 0.64$) compared to the baseline system ($M = 4.57$, $SD = 1.20$). Detailed *p*-values for these comparisons are presented in table 3.

7.1.2 Creativity Support. As shown in table 3 and appendix E, the *Creativity Support Index* (CSI) [22] for *ClueCart* was evaluated across five factors: *enjoyment*, *exploration*, *perceived effort*, *expressiveness*, and *immersion*. We excluded the inapplicable *collaboration* dimension.

Participants rated *ClueCart* significantly higher in enjoyment ($M = 6.10$, $SD = 0.86$) compared to the baseline system ($M = 4.05$, $SD = 1.32$). The system also encouraged greater exploration of creative avenues ($M = 6.28$, $SD = 0.63$) compared to the baseline system ($M = 3.93$, $SD = 1.29$). The perceived value of the results achieved with *ClueCart* was also higher ($M = 6.02$, $SD = 1.04$) than the baseline ($M = 4.12$, $SD = 1.63$). In terms of expressiveness, *ClueCart* received a higher rating ($M = 5.92$, $SD = 1.02$) than the baseline system ($M = 3.67$, $SD = 1.18$). Additionally, participants reported a significantly greater sense of immersion in the creative process ($M = 6.12$, $SD = 0.86$) with *ClueCart* compared to the baseline system ($M = 3.55$, $SD = 1.19$). Overall, *ClueCart* demonstrated significant improvements in creativity support across all the evaluated dimensions. Detailed *p*-values are presented in table 3.

7.2 Features of *ClueCart*

The features of *ClueCart* were evaluated based on three key aspects: *taxonomy*, *functionality*, and *interaction*. Each aspect focused on a specific feature of how participants engaged with the tool’s capabilities during story interpretation tasks. Unlike perception, features were assessed descriptively rather than comparatively with baseline *Miro*. The results are shown in table 4 while detailed questions corresponding to each feature aspect are provided in appendix G.

7.2.1 Taxonomy: Classification & Organization. The taxonomy feature in *ClueCart*, particularly its clue categorization function, received high praise from participants, with an average score of 6.35 ($SD = 0.73$). This aligns with **DR1**, suggesting that the system effectively mirrors users’ natural organization of story clues. Most participants (37 out of 40) expressed strong approval of the taxonomy. P1 mentioned, “*This categorization aligns with how I typically organize story clues in games.*” However, not all feedback was positive. Further analysis revealed that while the taxonomy was appreciated by most, some participants found it challenging to use. P23 noted that “*understanding and adapting to this taxonomy takes some time—there’s a bit of a learning curve*” and suggested integrating a detailed tutorial to ease adoption. P32 mentioned that while the taxonomy was helpful for structured analysis, it sometimes hindered creative exploration, stating, “*I feel constrained by the classification when trying to think outside the box.*” P3 and P10 recommended that future iterations explore quest-based clue categorization to align with built-in quest structures in some games, further enhancing the experience. These mixed responses highlight that while the taxonomy provides a useful framework, it can be difficult to adapt when game narratives or clues do not neatly fit predefined categories. This underscores the need for a more flexible, user-customizable taxonomy, particularly at the classification level, which remains a key limitation of the current design.

7.2.2 Functionality: Clue Understanding. The average score for content understanding was 6.30 ($SD = 0.78$), showing that participants found the system highly effective in helping them grasp story clues, aligning with **DR2**. Participants noted that summarizing visual clues or extracting keywords significantly improved their understanding of clues. P15 noted, “*Interpreting images can be exhausting, especially with large amounts of clues.*” P38, who worked on the *DNFTM* project, mentioned, “*In a text-heavy game like DNFTM, manually reviewing each clue is laborious. Summaries and keyword extraction make it easier to quickly grasp each clue’s content.*” Similarly, P18 added, “*When working on the presidential election story in DNFTM, the information was often scattered across newspapers*

Table 4: User ratings on *ClueCart* features.

Categories	DR	Features	Mean	S.D.
Taxonomy	DR1	Clue Categorization	6.35	0.73
Functionality	DR2	Content Understanding	6.30	0.78
	DR3	Relationship Understanding	6.10	0.77
	DR4	Retrieving Clues	6.30	0.71
	DR5	New Insights from External Information	6.35	0.79
Interaction		Overall System Interaction	6.20	0.75

and dialogues. Without the system summarizing the content, it would have been overwhelming.”

The average score for relationship understanding was 6.10 ($SD = 0.77$), indicating that *ClueCart* effectively aids users in identifying connections between clues, consistent with DR3. P15 noted, “The tag system makes it easier to track clue connections. Visualizing relationships helped me see links I hadn’t noticed before” However, some participants offered constructive feedback, with six (P13, P15, P14, P19, P32, P39) suggesting room for improvement. For instance, P39 said, “The relationship view feels a bit manual... I first look for clues with similar tags, then check for new ones layer by layer. It works for about 80% of cases, but more automation could help. If the system suggests relationships based on visuals, text, and tags after selecting two clues, that’ll really boost my workflow and give me new insights.” Their feedback points to opportunities for improving understanding efficiency.

7.2.3 Functionality: Story Interpretation. In terms of retrieving clues, *ClueCart* scored 6.30 ($SD = 0.71$), demonstrating its effectiveness in helping users quickly find necessary ones, aligning with DR4. Most participants in the *ClueCart* group (19 out of 20) found the filtering and retrieval features highly beneficial. P17 noted, “It’s really convenient to pull up what I need.” Improvements still remained. P16 suggested improving the sorting and recommendation by incorporating techniques like cosine similarity. Furthermore, P36, along with P12, P34, and P37, recommended adding a keyword search function within the retrieval process to further boost efficiency. This suggestion reflects a strong user demand for more advanced retrieval options to streamline content discovery and make the process even more intuitive.

The new insights from external information dimension achieved an average score of 6.35 ($SD = 0.80$), showing that participants found the external search feature valuable for gaining new insights, supporting DR5. Those who used the feature extensively provided positive feedback. P17 highlighted how the feature enriched her experience with *Hollow Knight*, explaining, “Sly’s name—‘cunning’ or ‘clever’—matches his character in the game. That deepened my understanding of him.” Similarly, P35 found the feature immersive while working on *DNFTM*, remarking, “Walker’s slogan, ‘make the country glorious again’, was clearly a nod to Trump. The system picked up on this and linked it to real-world context, making the narrative more immersive for me.” However, some participants (6 out of 20) explored this feature only briefly, suggesting that it might not have been central to their workflow. P39 remarked, “Most keywords were descriptive... I didn’t really get the key points.” He suggested that

the feature could be further enhanced by evolving from manual exploration to an automated system that identifies and highlights key insights for the user.

7.2.4 Overall System Interaction. The overall system interaction received an average score of 6.20 ($SD = 0.75$), indicating that participants found the system effective for organizing clues and facilitating story interpretation tasks. Many participants praised the interpretation panel, particularly its node grouping and arrow functionality, which supported their reasoning process. The drag-and-drop feature, highlighted by P12, P17, P32, P35 for its engaging and interactive nature. P17 and P31 appreciated the intuitiveness and efficiency of the keyword search function, while the tagging system in the clue library was valued for its quick organization capabilities.

Despite these strengths, participants suggested areas for improvement. Arrows were difficult to select and delete individually, and P31 recommended a bulk selection feature. Text node interactions posed issues for P33 and P12, who found resizing and moving text unclear and limited to corner adjustments. P32 pointed out the need for better layering control for text and images, as text often overlapped visuals, affecting organization. P19 also noted that clue retrieval became challenging with large datasets, suggesting improvements for better search functionality. Overall, participants found the system intuitive and helpful, particularly the interpretation panel and search features, though improvements in arrow usage, text box interaction, layering, and clue retrieval were recommended to enhance workflow efficiency.

7.3 Story Interpretation Assessment

The assessment of story interpretation using *ClueCart*, compared to *Miro*, focused on three main aspects: completion time, quality, and result comparison. They were analyzed to evaluate how *ClueCart* influenced the efficiency and effectiveness of the story interpretation process. Detailed questions are provided in appendix H.

7.3.1 Completion Time. In terms of completion time (fig. 8 (A)), *ClueCart* facilitated faster creative outputs, achieving a median time of 53 minutes compared to 62.5 minutes for the baseline system. Additionally, *ClueCart* exhibited more consistent performance, with first and third quartile times of 47.25 and 65.25 minutes, respectively, compared to 49.5 and 68.5 minutes for the baseline. These results suggest that *ClueCart* not only sped up task completion but also delivered more reliable performance across users.

7.3.2 Quality. In terms of quality (fig. 8 (B)), users rated their peers’ work more favorably across all five dimensions—Clarity and Focus,

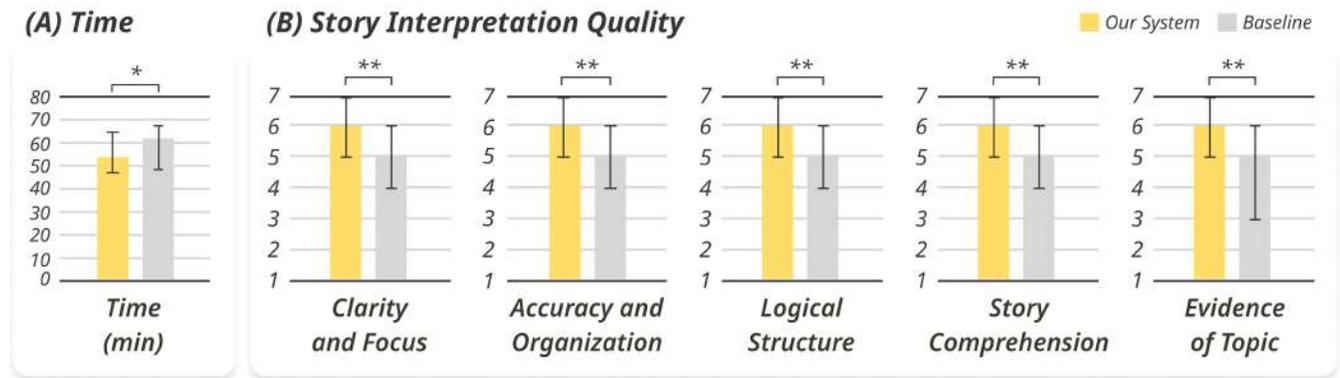


Figure 8: Story Interpretation Assessment. (A) Completion Time of *ClueCart* is shorter than baseline. (B) Story Interpretation Quality using *ClueCart* shows significant improvement compared to the baseline.

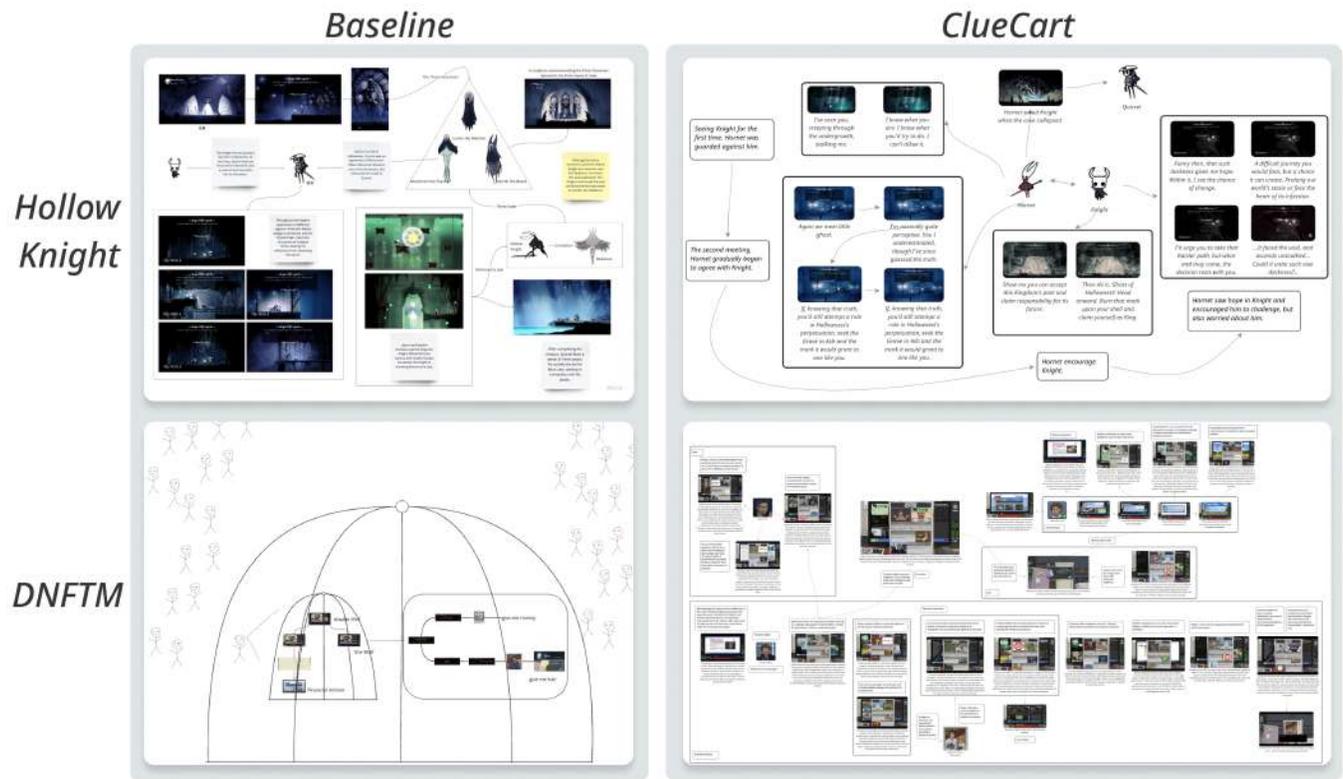


Figure 9: Story interpretation results of high average scores. The presentations from the baseline group were more stylistically varied, while outputs of *ClueCart* emphasized logical relationships between clues.

Accuracy and Organization, Logical Structure, Story Comprehension, and Evidence of Topic –when using *ClueCart*. *ClueCart* consistently achieved higher scores at all quartile levels, with a median score of 6, compared to 5 for the baseline. Notably, in the Evidence of Topic category, *ClueCart* showed significant improvement, with a first quartile score of 5 compared to 3 for the baseline. These results suggest that *ClueCart* not only expedited task completion

but also delivered superior-quality creative outputs compared to the baseline.

7.3.3 Result Comparison. Story interpretation results with high average scores are presented in fig. 9. In terms of flexibility, the Miro group’s presentations were more stylistically varied. For example, in *Hollow Knight*, P5 represented the relationships between the three *Dream Keepers* using a triangle, while in *DNFTM*, P21 used a

“cage” metaphor to convey the layers surrounding the game world, highlighting its unique atmosphere. In contrast, *ClueCart* outputs focused on logical relationships and completing clues, with users identifying more clues than the *Miro* group. This suggests that *ClueCart* encourages a more structured, systematic approach to story interpretation, prioritizing thoroughness and logical coherence over creative expression.

8 Discussion

In this section, we begin by comparing data collection methods, focusing on game mods and multi-modal classification models. We then explore *ClueCart*'s impact on the *Fan Creation Community*. Finally, we summarize design implications to guide future efforts in enhancing the efficiency, completeness, and creativity of story interpretation for creators.

8.1 Exploration of Data Collection Techniques

This study is grounded in the collection of raw materials from a variety of games. During the data collection process, we explored two methods: multi-modal classification models and in-game mods.

Multi-modal models can automate categorization using diverse data inputs, but they require large, labeled datasets for effective training. The variation in character appearances, locations, and achievements across games makes generalization difficult. Moreover, retraining these models for each game is costly and time-consuming, limiting their practicality in dynamic creative workflows. In contrast, in-game mods provide a more practical solution by enabling real-time extraction and categorization of game elements. This method integrates smoothly into workflows, capturing game-specific data during interaction for highly accurate classification. While mod development faces challenges like engine compatibility, our adaptable design principles mitigate these issues, allowing the mod to scale efficiently across different gaming environments.

8.2 Fostering Fan Creation Community

8.2.1 Enhancing Content Appeal and Engagement. For video creators, maintaining content relevancy and fostering audience engagement are essential [36, 127]. Content development often involves complex tasks such as evaluating clues and constructing narratives, which can be daunting (32 out of 40). *ClueCart* offers features like clue retrieval, recommendations, and linking in-game elements to real-world meanings, enabling creators to discover unique insights (P17, P35). This helps craft unexpected yet rational narrative twists that captivate viewers, boosting engagement and content virality (P32). Additionally, the concept maps produced by *ClueCart* can serve as supplementary resources for videos and forum discussion [126], helping audiences better understand complex story connections (P12, P18, P36). These maps not only enhance content sophistication but also provide a structured way for audiences to follow along. Furthermore, *ClueCart* empowers creators to present varied interpretations of a single story, encouraging vigorous discourse among viewers in the comments, enhancing interaction and promoting deeper engagement (P39, P40). This exchange of ideas enriches the content's depth and fosters the video's popularity and community discourse.

8.2.2 Lowering Knowledge Barrier Through Story Interpretation. As a new kind of art form [101], video games not only serve as a medium of entertainment but also possess significant cultural dissemination potential [20]. For instance, the *God of War* [87] series introduces players to Norse mythology [114], while *Black Myth: Wukong* [88] is based on China's *Journey to the West* [21], showcasing rich cultural heritage [61]. Similarly, *Assassin's Creed* [115] incorporates diverse cultural and historical backgrounds across its series [78, 90, 91], making games a bridge for global cultural exchange. By embedding real-world culture and history into game narratives, creators can lower the knowledge barriers for players. Those familiar with specific cultural references often form deeper emotional connections with the game content, while others can acquire new cultural knowledge and historical insights through the narrative. This approach not only enhances the immersive experience for players but also positions video games as a potential platform for education and cross-cultural exchange [24, 103].

8.2.3 Expanding Storytelling Beyond Games. *ClueCart* is not limited to game content but also has the potential to support other storytelling media, such as novels and films. This multi-media compatibility provides creators with a broader range of applications. For instance, creators can automatically categorize and manage key plot points, character relationships from novels, or visual symbols and narrative clues from films through *ClueCart*. This integration allows different narrative clues to be organized and presented on a unified platform [26], improving content management efficiency and offering powerful support for cross-media creation and analysis. This flexibility not only expands *ClueCart*'s application scope but also facilitates seamless transitions and in-depth interpretation across various storytelling formats.

8.3 Design Implications

8.3.1 Comparative Analysis of Game Clue Collection Tools. When comparing our mod with widely-used capture tools (P1, P27) like *OBS*¹³ or *Fraps*¹⁴, important design insights emerge. These tools, while popular for their simplicity, lack automatic in-game object detection and categorization, requiring creators to manually review and organize captured content. This can be particularly challenging for managing large volumes of assets (34 out of 40). Our mod fills this gap by automatically detecting and classifying game objects during capture, significantly reducing manual effort and enhancing workflow efficiency. This suggests a need for integrating automated classification tools into future capture systems to streamline processes for creators. However, mod installation can be more complex than standalone tools, which highlights the importance of simplifying mod integration for less experienced users. Additionally, privacy concerns (P14, P39, P40) arise from automated data capture, making it crucial for developers to design privacy-first frameworks that secure user and game data. Looking ahead, integrating automated asset capture directly into the game's core design could enhance platform compatibility across systems like

¹³<https://obsproject.com/>

¹⁴<https://fraps.com/>

*Nintendo Switch*¹⁵, *Xbox*¹⁶, and *PlayStation*¹⁷, offering a seamless, secure experience while complying with platform regulations.

8.3.2 Creator-Centric Clue Taxonomy. By applying the proposed taxonomy to categorize game story clues, participants (37 out of 40) reported significant improvements in their ability to comprehend and retrieve narrative elements. This taxonomy aligns closely with participants' intuitive workflows, indicating its potential for broader generalizability. For other developers and researchers, this approach provides a practical, adaptable framework that can be leveraged in diverse narrative-driven games. Its flexibility suggests that it could be extended beyond the contexts we studied, offering potential benefits for various story-rich games and even transmedia storytelling. While many games include built-in menu systems (e.g., in *The Witcher 3: Wild Hunt* [77]), these tend to focus on gameplay mechanics or achievements rather than supporting deeper narrative interpretation [51]. By contrast, our taxonomy addresses a gap by providing creators with a dedicated tool for analyzing and organizing story elements, independent of the developer's original intent. This aligns with other work advocating for more story-centric design approaches in interactive media [54, 85].

8.3.3 Bridging Game Worlds with Real-World Context. Integrating real-world contexts into game story interpretation has profound implications for content creation tools, particularly in their ability to support creators in producing deeper, more nuanced story interpretations [18]. System designed to facilitate this process — such as LLM-powered contextualization tools — automatically surface symbolic meanings and cultural references, reducing the need for creators to have extensive prior knowledge in specific areas like history or mythology. For example, games like *Dark Souls* [42] are filled with complex references to medieval history, religious symbolism, and esoteric lore [97]. Traditionally, creators had to conduct deep research to interpret these clues (P11, P15, P19, P37). By leveraging LLMs, creators can access contextual explanations automatically, allowing them to focus on crafting richer, more multidimensional interpretations with out being hindered by the need for specialized knowledge. This design approach lowers barriers to content creation, enabling a wider range of creators to engage with and interpret complex game stories. By simplifying the process, these tools make game narratives more accessible and open to diverse interpretations, enhancing both creator workflows and the storytelling experience for audiences.

8.3.4 Balancing Structured Thinking and Creative Expression. *ClueCart* encourages creators to adopt a structured, systematic approach to story interpretation, promoting both creative and critical thinking. By linking clues and discovering connections between in-game clues and real-world concepts, creators can uncover narrative threads more comprehensively, enhancing creativity (P15, P18, P33, P38) [45]. On the critical thinking side, *ClueCart*'s tagging system and relationship visualization tools lead creators to assess how different clues contribute to the overarching story (P13, P14, P16, P31, P39). This approach helps spot gaps, inconsistencies, or

overlooked connections, ultimately supporting a more coherent and logical interpretation of the narrative.

However, the concept map's structure introduces a trade-off between logical coherence and creative expression. While the tool supports clear, systematic thinking [46], it limits the flexibility available in more freeform tools like *Miro*, which allow creators to visually stretch, move, or reconfigure their ideas in more abstract, creative ways [124]. In contrast, *ClueCart*'s emphasis on organized, interconnected concepts, while valuable for ensuring logical clarity, may restrict the ability to visually explore the more organic, dynamic relationships that could emerge in less structured formats (P5, P7, P21). This trade-off highlights how a more structured, concept map approach, while effective in promoting logical analysis, can restrict the ability to visually represent creative, fluid interpretations. Future iterations could explore increasing flexibility in the concept map layout, enabling creators to blend the strengths of systematic thinking with greater room for creative expression.

8.4 Limitation and Future Work

This work has several limitations. First, although we isolated three pivotal analytical perspectives—plot analysis, character relationships, and character portraits—creators did not express a strong preference for these categories in practice. Future work could explore more flexible, adaptive methods to narrative analysis. Second, the system's reliance on modding limits its accessibility on closed platforms like the *Nintendo Switch*, *Xbox* and *PlayStation*. Future development should aim to integrate automatic clue collection directly into games for broader platform compatibility. Third, the manual exploration of clue relationships proved time-consuming for participants, suggesting future iterations could benefit from automated relationship detection to enhance efficiency. Fourth, another limitation lies in the heavy reliance on LLM technology for tasks like keyword extraction and clue summarization. While these features streamline workflows, they may shift too much interpretive responsibility to the tool, potentially reducing user engagement. Future work could balance automation with user control by introducing customization options, such as adjustable extraction pro or customizable outputs, ensuring creators remain actively involved in the interpretive process.

Regarding the taxonomy, while it was designed to support narrative analysis, some participants found it slightly difficult to adapt to, indicating both a learning curve and limited flexibility across diverse game genres. Specially, the inability to add high-level categories further constrained users' ability to tailor the taxonomy to their needs, increasing the learning curve. Future work could focus on enhancing the taxonomy's intuitiveness and expanding its customization capabilities. By introducing a short tutorial to explain the taxonomy and integrating category customization functions into the modding, *ClueCart* could better support flexible and adaptive narrative analysis while reducing user workload. This would allow users to create and manage additional high-level categories more efficiently, addressing the diverse needs of different game genres and improving the overall usability.

Looking forward, *ClueCart* could evolve into a dynamic, collaborative platform where creators share and organize clues in a wiki-like format. This would provide a resource pool for extensive

¹⁵<https://www.nintendo.com/jp/hardware/switch/index.html>

¹⁶<https://www.xbox.com>

¹⁷<https://www.playstation.com>

game clues and allow creators to showcase their narrative interpretations. *ClueCart* promotes creator-audience engagement by making interactions visible. As such, it fosters a more dynamic, involved community. Additionally, future iterations could explore features supporting fan fiction creation, such as cross-genre character pairings and integration of assets from multiple game narrative universes, enabling innovative storytelling. As the content grows, integrating machine learning could further personalize the experience by recommending clues based on each creator's unique style. This would enable innovative exploration of new concepts and enhance creative expression.

9 Conclusion

This study introduced *ClueCart*, a tool aimed at enhancing creativity by supporting game story interpretation through clue collection, classification, understanding, and inference. Insights gained from our workshop led to the creation of a creator-centric clue taxonomy, the development of an in-game mod for automatic clue collection, and the implementation of *ClueCart* as a platform where creators can interpret game narratives. Results from the between-subjects study confirmed that *ClueCart* improves the efficiency and quality of creators' story organization processes. It also validated the comprehensiveness and acceptance of the clue taxonomy. For future work, we plan to expand *ClueCart*'s application, transforming it into an online platform that integrates multiple game story clues and reasoning resources, fostering the growth of the fan creation community.

Acknowledgments

We thank anonymous reviewers for their valuable feedback. This work is supported by grants from the National Natural Science Foundation of China (No. 62372298), Shanghai Engineering Research Center of Intelligent Vision and Imaging, Shanghai Frontiers Science Center of Human-centered Artificial Intelligence (ShangHAI), and MoE Key Laboratory of Intelligent Perception and Human-Machine Collaboration (KLIP-HuMaCo).

References

- [1] 2K Boston and 2K Australia. 2007. *BioShock*. Video Game. A first-person shooter game developed by 2K Boston and 2K Australia, and published by 2K: <https://2k.com/en-US/game/bioshock-the-collection/>.
- [2] Mohammed Ibrahim Alhojailan. 2012. Thematic analysis: a critical review of its process and evaluation. In *WEI international European academic conference proceedings, Zagreb, Croatia*. Citeseer.
- [3] Sky LaRell Anderson. 2018. Extradigital narratives: Online communities and video games. *Transformative Works and Cultures* 28 (2018).
- [4] Angelo Maria Andriano. 2024. Enjoying the Uncertainty: How Dark Souls Performs Incompleteness Through Narrative, Level Design and Gameplay. *Games and Culture* (2024), 15554120241226837.
- [5] Aaron Bangor, Philip T Kortum, and James T Miller. 2008. An empirical evaluation of the system usability scale. *Intl. Journal of Human-Computer Interaction* 24, 6 (2008), 574–594.
- [6] Margo E Berendsen, Jeffrey D Hamerlinck, and Gerald R Webster. 2018. Digital story mapping to advance educational atlas design and enable student engagement. *ISPRS International Journal of Geo-Information* 7, 3 (2018), 125.
- [7] David M Blei, Andrew Y Ng, and Michael I Jordan. 2003. Latent dirichlet allocation. *Journal of machine Learning research* 3, Jan (2003), 993–1022.
- [8] Elizabeth Bonsignore, Alexander J Quinn, Allison Druin, and Benjamin B Bederson. 2013. Sharing stories “in the wild” a mobile storytelling case study using storykit. *ACM Transactions on Computer-Human Interaction (TOCHI)* 20, 3 (2013), 1–38.
- [9] Daniel Bormann and Tobias Greitemeyer. 2015. Immersed in virtual worlds and minds: effects of in-game storytelling on immersion, need satisfaction, and affective theory of mind. *Social Psychological and Personality Science* 6, 6 (2015), 646–652.
- [10] Michael Brandse. 2022. Conveyance of Narrative Through Digital Game Environments. In *International Conference on Human-Computer Interaction*. Springer, 141–159.
- [11] Virginia Braun and Victoria Clarke. 2012. *Thematic analysis*. American Psychological Association.
- [12] Matthew Brehmer, Bongshin Lee, Benjamin Bach, Nathalie Henry Riche, and Tamara Munzner. 2016. Timelines revisited: A design space and considerations for expressive storytelling. *IEEE transactions on visualization and computer graphics* 23, 9 (2016), 2151–2164.
- [13] Xiang Cao, Siàn E Lindley, John Helmes, and Abigail Sellen. 2010. Telling the whole story: anticipation, inspiration and reputation in a field deployment of TellTable. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work*, 251–260.
- [14] Capcom. 2001. *Phoenix Wright: Ace Attorney*. Video Game. A visual novel adventure game developed and published by Capcom: <https://www.ace-attorney.com/>.
- [15] Marco Caracciolo. 2024. *On Soulsring Worlds: Narrative Complexity, Digital Communities, and Interpretation in Dark Souls and Elden Ring*. Taylor & Francis.
- [16] Rogelio E Cardona-Rivera, José P Zagal, and Michael S Debus. 2020. Gfi: A formal approach to narrative design and game research. In *Interactive Storytelling: 13th International Conference on Interactive Digital Storytelling, ICIDS 2020, Bournemouth, UK, November 3–6, 2020, Proceedings 13*. Springer, 133–148.
- [17] Marc Cavazza and R Michael Young. 2017. Introduction to Interactive Storytelling 17. *Handbook of digital games and entertainment technologies* (2017), 377.
- [18] Vanessa Cesário, Mariana Ribeiro, and António Coelho. 2023. Exploring the intersection of storytelling, localisation, and immersion in video games—a case study of the *Witcher III: wild hunt*. In *International Conference on Human-Computer Interaction*. Springer, 546–552.
- [19] Kuo-En Chang, Yao-Ting Sung, and Ine-Dai Chen. 2002. The effect of concept mapping to enhance text comprehension and summarization. *The Journal of Experimental Education* 71, 1 (2002), 5–23.
- [20] Chi-Ying Chen. 2013. Is the video game a cultural vehicle? *Games and Culture* 8, 6 (2013), 408–427.
- [21] Wu Cheng'en. 2015. *Journey to the West*. In *Voices of East Asia*. Routledge, 68–80.
- [22] Erin Cherry and Celine Latulipe. 2014. Quantifying the creativity support of digital tools through the creativity support index. *ACM Transactions on Computer-Human Interaction (TOCHI)* 21, 4 (2014), 1–25.
- [23] Georgios Christou. 2014. The interplay between immersion and appeal in video games. *Computers in human behavior* 32 (2014), 92–100.
- [24] Malgorzata Cwil and William T Howe. 2020. Cross-cultural analysis of gamer identity: A comparison of the United States and Poland. *Simulation & Gaming* 51, 6 (2020), 785–801.
- [25] Rowan Daneels, Maarten Denoo, Alexander Vandewalle, Bruno Dupont, and Steven Malliet. 2022. The Digital Game Analysis Protocol (DiGAP): Introducing a guide for reflexive and transparent game analyses. *Game Studies* 22, 2 (2022).
- [26] Drew Davidson. 2010. *Cross-media communications: An introduction to the art of creating integrated media experiences*. Lulu.com.
- [27] Emilio Di Giacomo, Walter Didimo, Giuseppe Liotta, Fabrizio Montecchiani, and Alessandra Tappini. 2020. Storyline visualizations with ubiquitous actors. In *International Symposium on Graph Drawing and Network Visualization*. Springer, 324–332.
- [28] Sebastian Domsch. 2013. *Storyplaying: Agency and narrative in video games*. De Gruyter.
- [29] Nicola Dusi, Ilaria Ferretti, and Marco Furini. 2016. PlayTheCityRE: A visual storytelling system that transforms recorded film memories into visual history. In *2016 IEEE Symposium on Computers and Communication (ISCC)*. IEEE, 85–90.
- [30] Clara Fernández-Vara. 2011. Game spaces speak volumes: Indexical storytelling. (2011).
- [31] Gabriele Ferri. 2007. Making sense of a game: a preliminary sketch for a semantic approach to games. In *Proceedings of the international conference on Advances in computer entertainment technology*. 226–227.
- [32] Gabriele Ferri. 2007. Narrating machines and interactive matrices: a semiotic common ground for game studies. In *Proceedings of DiGRA 2007 Conference: Situated Play*.
- [33] Gabriele Ferri. 2013. Satire, propaganda, play, storytelling, notes on critical interactive digital narratives. In *Interactive Storytelling: 6th International Conference, ICIDS 2013, Istanbul, Turkey, November 6–9, 2013, Proceedings 6*. Springer, 174–179.
- [34] Gabriele Ferri. 2015. *Narrative structures in IDN authoring and analysis*. In *Interactive Digital Narrative*. Routledge, 77–90.
- [35] Fictiorama Studios. 2018. *Do Not Feed The Monkeys*. Video Game. A 2018 adventure and life simulation video game developed by the Fictiorama Studios and published by Alawar Premium: <https://store.steampowered.com/app/658850/>

- Do_Not_Feed_the_Monkeys/.
- [36] Flavio Figueiredo, Fabricio Benevenuto, and Jussara M Almeida. 2011. The tube over time: characterizing popularity growth of youtube videos. In *Proceedings of the fourth ACM international conference on Web search and data mining*. 745–754.
- [37] Gonzalo Frasca. 1999. Ludology meets narratology: Similitude and differences between (video) games and narrative. *Ludology.org* (1999).
- [38] Gonzalo Frasca. 2013. Simulation versus narrative: Introduction to ludology. In *The video game theory reader*. Routledge, 221–235.
- [39] Jonas Frich, Michael Mose Biskjaer, and Peter Dalsgaard. 2018. Why HCI and creativity research must collaborate to develop new creativity support tools. In *Proceedings of the Technology, Mind, and Society*. 1–6.
- [40] Jonas Frich, Peter Dalsgaard, Mihaela Taranu, and Michael Mose Biskjaer. 2024. How Are Measures of Usability and Creativity Support Correlated?. In *Proceedings of the European Conference on Cognitive Ergonomics 2024*. 1–7.
- [41] FromSoftware. 2009. Demon's Souls. Video Game. An action role-playing game developed by FromSoftware: <https://www.fromsoftware.jp/ww/detail.html?csm=070>.
- [42] FromSoftware. 2011. Dark Souls. Video Game. A dark fantasy action role-playing game series developed by FromSoftware and published by Bandai Namco Entertainment: <https://www.darksouls.jp/>.
- [43] FromSoftware. 2022. Elden Ring. Video Game. A game developed by FromSoftware, known for its rich lore and hidden story elements: <https://en.bandainamcoent.eu/elden-ring/elden-ring>.
- [44] Elizabeth Goins. 2018. Structuring Digital Game Stories. In *Interactive Storytelling*, Rebecca Rouse, Hartmut Koenitz, and Mads Haahr (Eds.). Springer International Publishing, Cham, 265–269.
- [45] Julie Goldstein. 2001. Concept mapping, mind mapping and creativity: documenting the creative process for computer animators. *ACM SIGGRAPH Computer Graphics* 35, 2 (2001), 32–35.
- [46] James Gorman and Jane Heinze-Fry. 2015. Conceptual Mapping Facilitates Coherence and Critical Thinking in the Science Education System. In *STEM Education: Concepts, Methodologies, Tools, and Applications*. IGI Global, 1227–1258.
- [47] Garth Groshans, Elena Mikhailova, Christopher Post, Mark Schlautman, Patricia Carbajales-Dale, and Kayla Payne. 2019. Digital story map learning for STEM disciplines. *Education Sciences* 9, 2 (2019), 75.
- [48] Alex Grunberg. 2024. The Ones Who Walk Away from Hollownest: Hollow Knight's Radical Response to the Omelas Dilemma. *Games and Culture* 19, 4 (2024), 433–448.
- [49] Golina Hulstein, Vanessa Peña-Araya, and Anastasia Bezerianos. 2022. Geo-Storylines: Integrating maps into storyline visualizations. *IEEE Transactions on Visualization and Computer Graphics* 29, 1 (2022), 994–1004.
- [50] Angel Hsing-Chi Hwang. 2022. Too late to be creative? AI-empowered tools in creative processes. In *CHI conference on human factors in computing systems extended abstracts*. 1–9.
- [51] Martti Kandelin. 2021. Creating immersive menu systems in video games: menu systems in an early access game. (2021).
- [52] Philipp Klein. 2023. *Reimagining design: Exploring the potential of the infinite canvas for collaborative interactive articles*. Ph. D. Dissertation. Technische Universität Wien.
- [53] Hartmut Koenitz. 2024. Narrative in video games. In *Encyclopedia of computer graphics and games*. Springer, 1230–1238.
- [54] Hannu Kukka, Johanna Ylipulli, Jorge Goncalves, Timo Ojala, Matias Kukka, and Mirja Syrjälä. 2017. Creator-centric study of digital art exhibitions on interactive public displays. In *Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia*. 37–48.
- [55] Bjarke Alexander Larsen, Nic Junius, and Elin Carstensdottir. 2023. Communal Ritual Play: Repetition and Interpretation of Game Narratives Across Communities. In *International Conference on Interactive Digital Storytelling*. Springer, 472–488.
- [56] Pierre Le Bras, David A Robb, Thomas S Methven, Stefano Padilla, and Mike J Chantler. 2018. Improving user confidence in concept maps: Exploring data driven explanations. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–13.
- [57] Daniel Livingstone, Sandy Jean-Jacques Louchart, and Stuart Jeffrey. 2016. Archaeological storytelling in games. (2016).
- [58] Fei Lu, Feng Tian, Yingying Jiang, Xiang Cao, Wencan Luo, Guang Li, Xiaolong Zhang, Guozhong Dai, and Hongan Wang. 2011. ShadowStory: creative and collaborative digital storytelling inspired by cultural heritage. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1919–1928.
- [59] William G Lycan. 2018. *Philosophy of language: A contemporary introduction*. Routledge.
- [60] Neil Maiden, Konstantinos Zachos, Amanda Brown, George Brock, Lars Nyre, Aleksander Nygård Tonheim, Dimitris Apsotholou, and Jeremy Evans. 2018. Making the news: Digital creativity support for journalists. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–11.
- [61] Hengbo Mao. 2024. A Study on Cross-Cultural Communication Factors of Black Myth Wukong. In *2024 3rd International Conference on Social Sciences and Humanities and Arts (SSHA 2024)*. Atlantis Press, 700–708.
- [62] Michael Mateas. 2001. A preliminary poetics for interactive drama and games. *Digital Creativity* 12, 3 (2001), 140–152.
- [63] Michael Mateas and Andrew Stern. 2005. Structuring content in the Façade interactive drama architecture. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, Vol. 1. 93–98.
- [64] Patrick E McKnight and Julius Najab. 2010. Mann-Whitney U Test. *The Corsini encyclopedia of psychology* (2010), 1–1.
- [65] Bria Mears and Jichen Zhu. 2017. Design patterns for silent player characters in narrative-driven games. In *Proceedings of the 12th International Conference on the Foundations of Digital Games*. 1–4.
- [66] Xiaojun Meng, Shengdong Zhao, and Darren Edge. 2016. HyNote: integrated concept mapping and notetaking. In *Proceedings of the International Working Conference on Advanced Visual Interfaces*. 236–239.
- [67] Joseph Hillis Miller and Julian Wolfreys. 2005. *The J. Hillis Miller Reader*. Stanford University Press.
- [68] Moon Studios. 2015. Ori and the Blind Forest. Video Game. A platform-adventure Metrodvania video game developed by Moon Studios and published by Microsoft Studios: <https://www.xbox.com/en-US/games/store/ori-and-the-blind-forest-definitive-edition/BW85KQB8Q31M>.
- [69] Ian Morgan. 2022. "A Little Jest that Maybe We Could Try"—Playing with Myth in *Supergiant Games' Hades (2020)*. Master's thesis. The University of Arizona.
- [70] Eric Mörth, Stefan Bruckner, and Noeska N Smit. 2022. ScrollyVis: Interactive visual authoring of guided dynamic narratives for scientific scrollytelling. *IEEE Transactions on Visualization and Computer Graphics* 29, 12 (2022), 5165–5177.
- [71] Janet Murray. 2005. The Last Word on Ludology v Narratology in Game Studies (2005)". *Janet H. Murray: Humanistic Design for an Emerging Medium* (2005).
- [72] Nintendo. 1996. Super Mario 64. Video Game. A platform game developed and published by Nintendo: <https://www.nintendo.com/en-gb/Games/Nintendo-64/Super-Mario-64-269745.html>.
- [73] Michael Nitsche. 2008. *Video game spaces: image, play, and structure in 3D worlds*. MIT Press.
- [74] Anna Offenwanger, Matthew Brehmer, Fanny Chevalier, and Theophanis Tsandilas. 2023. Timesplines: Sketch-based authoring of flexible and idiosyncratic timelines. *IEEE Transactions on Visualization and Computer Graphics* (2023).
- [75] Kalpesh Padia, Kaveen Herath Bandara, and Christopher G Healey. 2018. Yarn: Generating Storyline Visualizations Using HTN Planning. In *Graphics Interface*. 26–33.
- [76] Kalpesh Padia, Kaveen Herath Bandara, and Christopher G Healey. 2019. A system for generating storyline visualizations using hierarchical task network planning. *Computers & Graphics* 78 (2019), 64–75.
- [77] Polish studio CD Projekt. 2015. The Witcher 3: Wild Hunt. Video Game. A 2015 action role-playing game developed and published by the Polish studio CD Projekt: <https://www.thewitcher.com/us/en/witcher3>.
- [78] Aris Politopoulos, Angus AA Mol, Krijn HJ Boom, and Csilla E Ariese. 2019. "History is our playground": action and authenticity in assassin's creed: Odyssey. *Advances in Archaeological Practice* 7, 3 (2019), 317–323.
- [79] Nathaniel Poor. 2014. Computer game modders' motivations and sense of community: A mixed-methods approach. *New media & society* 16, 8 (2014), 1249–1267.
- [80] Vladimir Iakovlevich Propp. 1968. *Morphology of the folktale*. University of Texas Press. <http://books.google.com/books?isbn=0292783760>
- [81] R Radics. 2021. Methodological processing of gastronomic short stories using MIRO. COM platform. In *ICER2021 Proceedings*. IATED, 8713–8717.
- [82] Anthony Raduazzo and Ashley Lear. 2024. Isolation Without Despair: Hollow Knight in a Post-Pandemic Landscape. (2024).
- [83] Tine Rassalle. 2021. Archaeogaming: When archaeology and video games come together. *Near Eastern Archaeology* 84, 1 (2021), 4–11.
- [84] Richard Rouse III. 2010. Environmental Narrative: Your World is Your Story. In *Game Developers Conference, San Francisco CA*.
- [85] Asmalina Saleh, Cindy E Hmelo-Silver, Krista D Glazewski, Bradford Mott, Yuxin Chen, Jonathan P Rowe, and James C Lester. 2019. Collaborative inquiry play: A design case to frame integration of collaborative problem solving with story-centric games. *Information and Learning Sciences* 120, 9/10 (2019), 547–566.
- [86] Adam Salonen and Ara Mohammad. 2017. Practicing Octalysis: The comparison of an Octalysis analysis versus Metacritic's reviews.
- [87] Santa Monica Studio. 2005. God of War. Video Game. a series of video games centered around Greek Mythology developed by Santa Monica Studio and published by Sony Computer Entertainment: <https://www.playstation.com/en-us/god-of-war/>.
- [88] Santa Monica Studio. 2024. Black Myth: Wukong. Video Game. A 2024 action role-playing game developed and published by Game Science: <https://www.heishenhua.com/>.
- [89] Benjamin Saunders, Julius Sim, Tom Kingstone, Shula Baker, Jackie Waterfield, Bernadette Bartlam, Heather Burroughs, and Clare Jinks. 2018. Saturation in qualitative research: exploring its conceptualization and operationalization. *Quality & quantity* 52 (2018), 1893–1907.

- [90] Magy Seif El-Nasr, Maha Al-Saati, Simon Niedenthal, and David Milam. 2008. Assassin's creed: A multi-cultural read. *Loading...* 2, 3 (2008).
- [91] Adrienne Shaw et al. 2015. Tyranny of realism: Historical accuracy and the politics of representation in Assassin's Creed III. *Faculty/Researcher Works* (2015).
- [92] Yotam Shibolet. 2023. The Harmony of Escaping Hell: Ludonarrative Interplay in Hades. In *Abstract Proceedings of DiGRA 2023 Conference: Limits and Margins of Games*.
- [93] Ben Shneiderman. 2007. Creativity support tools: accelerating discovery and innovation. *Commun. ACM* 50, 12 (2007), 20–32.
- [94] Inês Silva, Pedro Cardoso, and Eva Oliveira. 2019. Narrative and gameplay: the balanced and imbalanced relationship between dramatic tension and gameplay tension. In *Proceedings of the 9th International Conference on Digital and Interactive Arts*. 1–8.
- [95] Natalie Simper, Richard Reeve, and JR Kirby. 2016. Effects of concept mapping on creativity in photo stories. *Creativity Research Journal* 28, 1 (2016), 46–51.
- [96] Rebekah Small. 2019. *How Mods Empower Co-Creation and Remix*. Michigan State University.
- [97] Tobi Smethurst. 2016. *Mapping myth: an analysis of the hero myth in from-software's dark souls series*. Ph.D. Dissertation. PhD thesis, Ghent University, Accessed 19 Sept 2023. <https://lib.ugent.be>
- [98] Evan T Smith and Chandramallika Basak. 2023. A game-factors approach to cognitive benefits from video-game training: A meta-analysis. *PLoS One* 18, 8 (2023), e0285925.
- [99] Harvey Smith and Matthias Worch. 2010. What happened here? Environmental storytelling. In *Game Developers Conference*. 9–13.
- [100] Florence Smith Nicholls and Michael Cook. 2023. "That Darned Sandstorm": A Study of Procedural Generation through Archaeological Storytelling. In *Proceedings of the 18th International Conference on the Foundations of Digital Games*. 1–8.
- [101] Aaron Smuts. 2005. Are video games art? *Contemporary Aesthetics (Journal Archive)* 3, 1 (2005), 6.
- [102] Melissa Somerdin. 2016. The game debate: Video games as innovative storytelling. *The Oswald Review: An International Journal of Undergraduate Research and Criticism in the Discipline of English* 18, 1 (2016), 7.
- [103] Ali Soyoof. 2018. Video game and culture: A case study of EFL student players' views on their acquisition of cultural knowledge and sensitivity. *International Journal of Pedagogies and Learning* 13, 2 (2018), 91–102.
- [104] CARLO SPEZZAFERRI. [n. d.]. Soulslike video games—a corpus linguistic analysis of the language of "Elden Ring". ([n. d.]).
- [105] Supergiant Games. 2020. Hades. Video Game. A roguelike action role-playing game developed and published by Supergiant Games: <https://store.privatedivision.com/game/buy-hades>.
- [106] Karen Tanenbaum and Theresa J Tanenbaum. 2009. Commitment to meaning: A reframing of agency in games. (2009).
- [107] Chien-Lin Tang, Jingxian Liao, Hao-Chuan Wang, Ching-Ying Sung, and Wen-Chieh Lin. 2021. Conceptguide: Supporting online video learning with concept map-based recommendation of learning path. In *Proceedings of the Web Conference 2021*. 2757–2768.
- [108] Tan Tang, Renzhong Li, Xinke Wu, Shuhan Liu, Johannes Knittel, Steffen Koch, Thomas Ertl, Lingyun Yu, Peiran Ren, and Yingcai Wu. 2020. Plotthread: Creating expressive storyline visualizations using reinforcement learning. *IEEE Transactions on Visualization and Computer Graphics* 27, 2 (2020), 294–303.
- [109] Tan Tang, Sadia Rubab, Jiewen Lai, Weiwei Cui, Lingyun Yu, and Yingcai Wu. 2018. iStoryline: Effective convergence to hand-drawn storylines. *IEEE transactions on visualization and computer graphics* 25, 1 (2018), 769–778.
- [110] Renaud Bougueng Tchemeube, Jeffrey Ens, Cale Plut, Philippe Pasquier, Maryam Safi, Yvan Grabit, and Jean-Baptiste Rolland. 2023. Evaluating Human-AI Interaction via Usability, User Experience and Acceptance Measures for MMM-C: A Creative AI System for Music Composition. In *IJCAL*. 5769–5778.
- [111] Team Cherry. 2017. Hollow Knight. Video Game. A 2017 Metroidvania video game developed and published by independent developer Team Cherry: <https://www.hollowknight.com/>.
- [112] Elizabeth Thiry, Siân Lindley, Richard Banks, and Tim Regan. 2013. Authoring personal histories: Exploring the timeline as a framework for meaning making. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1619–1628.
- [113] Chao Tong, Richard Roberts, Rita Borgo, Sean Walton, Robert S Laramee, Kodzo Wegba, Aidong Lu, Yun Wang, Huamin Qu, Qiong Luo, et al. 2018. Storytelling and visualization: An extended survey. *Information* 9, 3 (2018), 65.
- [114] Camilla Tosi. 2022. Classical world and game learning. The case of Athena in "God of War". *Una Kov-Rivista di studi sul classico e sulla sua ricezione nella letteratura italiana moderna e contemporanea* 3, 3 (2022), 184–213.
- [115] Ubisoft. 2007. Black Myth: Wukong. Video Game. A historical action-adventure video game series and media franchise published by Ubisoft and developed mainly by its studio Ubisoft Montreal: <https://www.ubisoft.com/en-gb/game/assassins-creed>.
- [116] Robert Veloya. 2023. An Analysis of the Debate and How It Changed Everything: Narratology vs. Ludology. (2023).
- [117] Jorge J Villalon and Rafael A Calvo. 2008. Concept map mining: A definition and a framework for its evaluation. In *2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, Vol. 3. IEEE, 357–360.
- [118] Shih-Ping Wang and Yih-Lan Chen. 2018. Effects of multimodal learning analytics with concept maps on college students' vocabulary and reading performance. *Journal of Educational Technology & Society* 21, 4 (2018), 12–25.
- [119] Henrik Warpefelt. 2016. *The Non-Player Character: Exploring the believability of NPC presentation and behavior*. Ph.D. Dissertation. Department of Computer and Systems Sciences, Stockholm University.
- [120] Huaxin Wei. 2010. Embedded narrative in game design. In *Proceedings of the international academic conference on the future of game design and technology*. 247–250.
- [121] Huaxin Wei. 2011. Analyzing the game narrative: Structure and technique. (2011).
- [122] Huaxin Wei, Jim Bizzocchi, and Tom Calvert. 2010. Time and space in digital game storytelling. *International Journal of Computer Games Technology* 2010, 1 (2010), 897217.
- [123] Matthew Wells. 2018. Game not over: End-user programming and game system modding as models for extending community engagement. In *Journal d'Interaction Personne-Système*.
- [124] Christina K White, Kristin L Wood, and Dan Jensen. 2012. From brainstorming to C-sketch to principles of historical innovators: ideation techniques to enhance student creativity. *Journal of STEM Education: Innovations and Research* 13, 5 (2012).
- [125] Tim Wulf, Daniel Possler, and Johannes Breuer. 2022. Content Analysis in the Research Field of Video Games. In *Standardisierte Inhaltsanalyse in der Kommunikationswissenschaft—Standardized Content Analysis in Communication Research: Ein Handbuch-A Handbook*. Springer Fachmedien Wiesbaden Wiesbaden, 287–297.
- [126] Shan Zhang, Xiaojun Meng, Can Liu, Shengdong Zhao, Vibhor Sehgal, and Morten Fjeld. 2019. ScaffoMapping: Assisting Concept Mapping for Video Learners. Springer-Verlag, Berlin, Heidelberg, 314–328. doi:10.1007/978-3-030-29384-0_20
- [127] Renjie Zhou, Samamon Khemmarat, Lixin Gao, Jian Wan, and Jilin Zhang. 2016. How YouTube videos are discovered and its impact on video views. *Multimedia Tools and Applications* 75 (2016), 6035–6058.

A Prompt for Categorization, Description Generation, and Keyword Extraction

```

**Task:** Analyze screenshots from a game and categorize their content. For each category, perform specific analysis and generate structured output.
1. **Text-Only Screenshots**
  - **Task:** Extract all text visible in the screenshot.
  - **Output:** Summarize the main content in 30 words or fewer, focusing on keywords (up to three).
2. **NPC Dialogue Screenshots**
  - **Task:** Extract the full dialogue from NPC interactions.
  - **Output:** Summarize the conversation in 30 words or fewer and highlight up to three main keywords.
3. **Artifact-Related Screenshots**
  - Screenshots with text containing ``inventory``:
    - **Task:** Identify as an inventory page. Extract item names and descriptions.
    - **Output:** Provide a concise summary of the inventory content, emphasizing up to three main keywords.
  - Other artifact-related screenshots:
    - **Task:** Extract and list artifact names and descriptions.
    - **Output:** Summarize findings, with up to three main keywords.
4. **Environment Screenshots**
  - Applicable for screenshots without text or those containing the keywords: ``ENTER``, ``CHALLENGE``, or ``LISTEN``.
  - **Task:** Provide a one-sentence summary of the depicted environment.
  - **Output:** Include up to three main keywords relevant to the scene.

**General Guidelines:**
- Ignore any health indicators or character statuses shown in the top-left corner of screenshots.
- Summaries must not exceed 30 words per screenshot.

**Expected Output Format:**
- **Element:** Specify the type (e.g., ``text-only``, ``dialogue``, ``artifact``, or ``environment``).
- **Description:** Include extracted text or summaries as per category.
- **Keywords:** List up to three keywords related to the content.

```

B Prompt for Real-World Context

```

Define the term /${keyword} as it relates to the game /${game}. Provide a concise explanation in 20 words or fewer.

```

C Participants in the user study

Table 5: Detailed information of the participants in two user studies.

Game	Tool	ID	Gender	Age	Platform	Frequency
Hollow Knight	Miro	P1	Male	22	Video Platforms	Seasonally
		P2	Male	23	Video Platforms	Monthly
		P3	Female	21	Video Platforms & Forums & Game Wiki	Seasonally
		P4	Male	21	Video Platforms	Monthly
		P5	Male	20	Video Platforms	Weekly
		P6	Female	21	Video Platforms	Daily
		P7	Male	23	Video Platforms	Weekly
		P8	Male	22	Video Platforms	Weekly
		P9	Female	24	Video Platforms & Forums	Weekly
		P10	Female	22	Video Platforms & Forums	Weekly
	ClueCart	P11	Male	24	Video Platforms & Forums	Monthly
		P12	Female	20	Video Platforms & Forums	Seasonally
		P13	Female	24	Video Platforms	Weekly
		P14	Male	23	Video Platforms & Forums	Monthly
		P15	Male	24	Video Platforms	Seasonally
		P16	Female	22	Video Platforms	Monthly
		P17	Male	22	Video Platforms & Forums	Weekly
		P18	Male	20	Video Platforms	Seasonally
		P19	Female	21	Video Platforms & Forums	Monthly
		P20	Male	22	Video Platforms & Forums	Monthly
DNFTM	Miro	P21	Male	28	Video Platforms	Seasonally
		P22	Female	21	Forums	Weekly
		P23	Female	22	Video Platforms	Weekly
		P24	Male	21	Video Platforms & Forums	Weekly
		P25	Male	20	Video Platforms	Weekly
		P26	Male	22	Video Platforms & Forums	Monthly
		P27	Male	22	Video Platforms & Forums	Weekly
		P28	Female	24	Video Platforms & Forums	Monthly
	ClueCart	P29	Male	24	Video Platforms	Weekly
		P30	Female	24	Video Platforms	Monthly
		P31	Female	24	Video Platforms & Forums	Seasonally
		P32	Male	25	Video Platforms & Forums	Monthly
		P33	Female	22	Video Platforms	Seasonally
		P34	Male	21	Video Platforms & Forums	Weekly
ClueCart	P35	Female	23	Video Platforms	Seasonally	
	P36	Female	23	Video Platforms & Forums	Weekly	
	P37	Male	21	Video Platforms & Forums	Weekly	
	P38	Female	24	Video Platforms	Monthly	
	P39	Female	20	Video Platforms	Seasonally	
	P40	Male	29	Video Platforms & Forums	Daily	

D Evaluation Dimensions

Table 6: Three adopted dimensions, along with their categories and corresponding aspects.

Dimensions	Categories	Aspect
Perception	System Usability Scale	Gives a global view of subjective assessments of tool usability.
	Creativity Support Index	Assesses how effectively the tool supports creative processes during task completion.
Features	Taxonomy	Facilitates Classification & Organization , addressing C1 and C2 (DR1).
	Functionality	Facilitates Clue Understanding , addressing C3 and C4 (DR2, DR3).
		Facilitates Story Interpretation , addressing C5 and C6 (DR4, DR5).
Interaction	Assesses the overall system interaction design.	
Story Interpretation Assessment	Completion Time	Measures the time taken by participants to complete the task.
	Quality	Evaluates clarity, organization, and logical structure using a concept map evaluation checklist.

E System Usability and Creativity Support Results

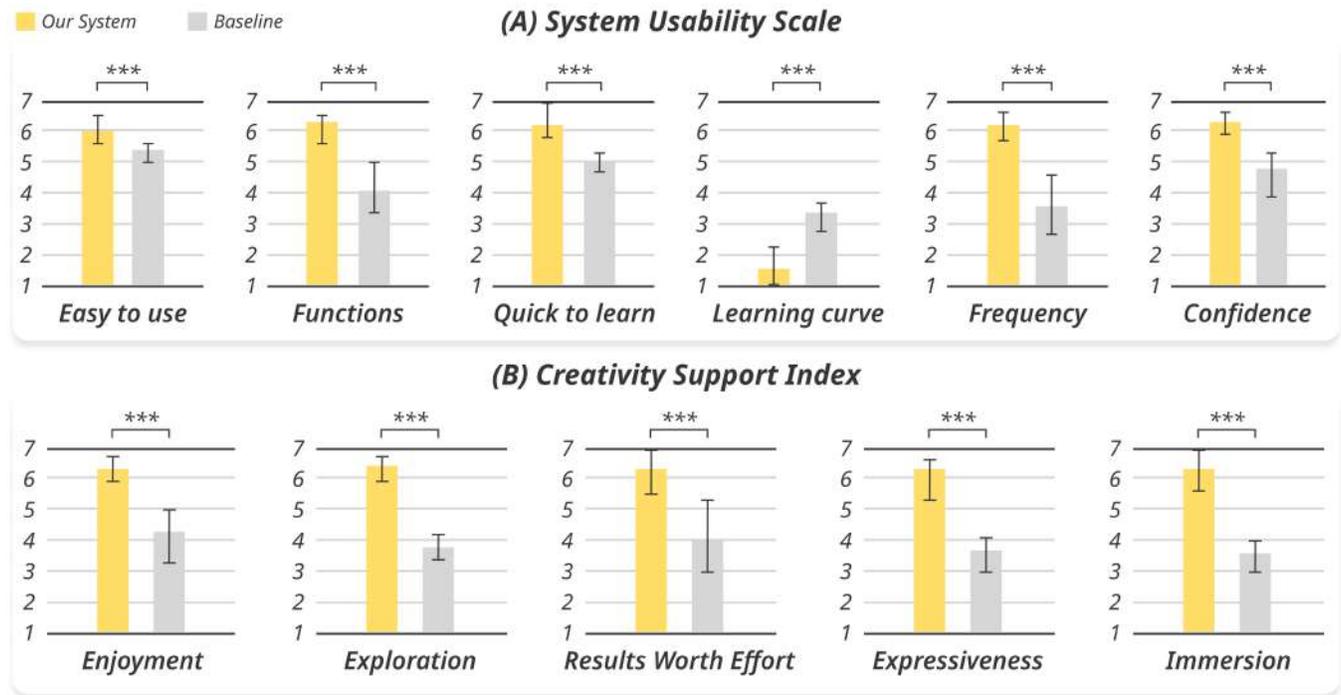


Figure 10: System Usability Scale and Creativity Support Index. All the metrics show significant improvement compared with the baseline.

F Questions about Perception of *ClueCart*.

Table 7: Questions about perception of *ClueCart*.

Categories	Factors	Questions
System Usability Scale	Easy to use	I thought the system was easy to use.
	Functions	I found the various functions in this system were well integrated.
	Quick to learn	I would imagine that most people would learn to use this system very quickly.
	Learning curve	I needed to learn a lot of things before I could get going with this system.
	Frequency	I think that I would like to use this system frequently.
	Confidence	I felt very confident using the system.
Creativity Support Index	Enjoyment	I enjoyed using this system.
	Exploration	The system helped me to track different ideas, outcomes, or possibilities.
	Results Worth Effort	What I was able to produce was worth the effort I had to exert to produce it.
	Expressiveness	The system allowed me to be very expressive.
	Immersion	I became absorbed in the activity and forgot about the system that I was using.

G Questions about Features of *ClueCart*.

Table 8: Questions about features of *ClueCart*.

Categories	Features	Questions
Taxonomy	Clue Categorization	<i>ClueCart</i> 's clue categorization feature fits my usage habits.
Functionality	Content Understanding	<i>ClueCart</i> helps me quickly understand the content of each clue.
	Relationship Understanding	<i>ClueCart</i> helps me quickly understand the relationship between clues.
	Retrieving Clues	<i>ClueCart</i> helps me quickly find the materials I need.
	New Insights from External Information	The external information feature in <i>ClueCart</i> 's Interpretation section helps me gain new insights.
Interaction	Overall System Interaction	<i>ClueCart</i> 's interactions help me organize clues and reason through the story.

H Questions about Story Interpretation Assessment of *ClueCart*.

Table 9: Questions about story interpretation assessment of *ClueCart*.

Factors	Questions
Clarity and Focus	The result is clear, legible, and focused.
Accuracy and Organization	Information is clear, accurate, and well organized.
Logical Structure	Concepts are linked logically and reflect the essential information about the story.
Story Comprehension	Content is logically arranged to facilitate comprehension.
Evidence of Topic	The result shows evidence of the topic.