Bonsai: Designing for Cultivation in AI Interactive Digital Narratives

Tiffany Wang
cabbages@mit.edu
MIT Media Lab
Cambridge, MA, USA

Max Kreminski
mkremins@cornell.edu
Cornell Tech
New York, NY, USA

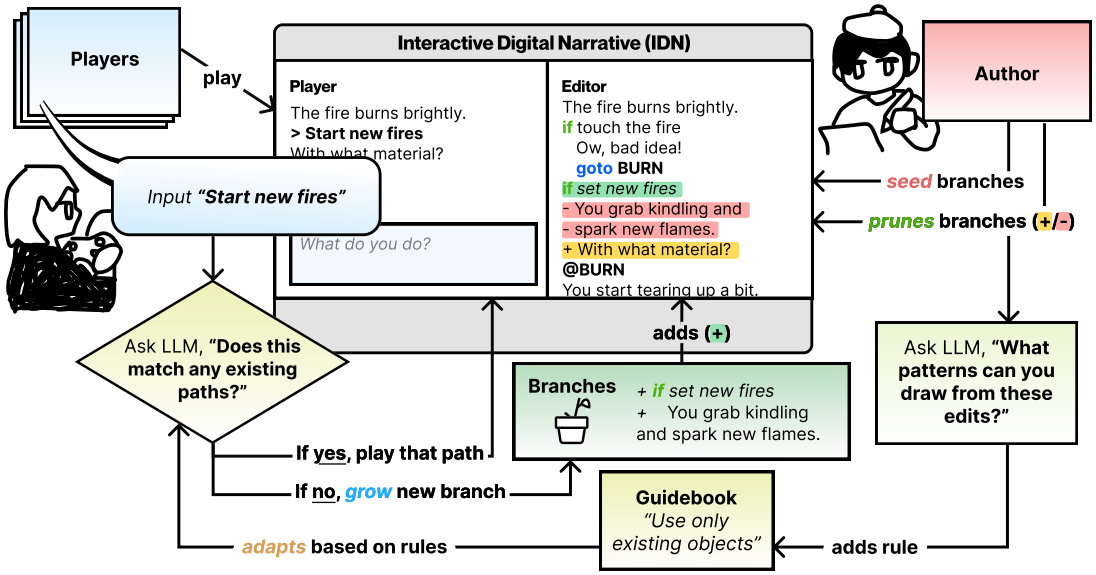

Figure 1: Overview of Bonsai's cultivation cycle. Players input natural language actions (e.g., *"Start new fires"*), triggering similarity matching against existing paths. Novel actions generate new branches guided by the Guidebook. Authors asynchronously review generated content in the Editor, accepting (+) or rejecting (-) branches. Accepted edits create new versions and update the Guidebook with learned patterns (e.g., *"Use only existing objects"*), which inform future generation.

Abstract

Authors of LLM-based interactive digital narratives (IDNs) struggle to preserve creative intent as player choices and real-time generation pull storylines in unpredictable directions. Existing frameworks treat IDNs as static once published, limiting authors' insight into and control over the storylines that emerge from unexpected player input in the wild. We propose *cultivation*: a design metaphor in which LLM-generated branches are stored as persistent material for authors to shape through iterative curation. Authors seed an initial scenario; the system grows new branches in response to player exploration; authors prune and revise what emerges, accumulating preference data that steers future generation. We demonstrate cultivation through *Bonsai*, an IDN authoring tool, and complement this design account with three simulated experiments showing that learned preferences transfer to unseen scenes, are project-specific rather than portable, and improve substantially when extraction is structured around IDN authoring categories. This metaphor reframes human-AI creative collaboration: authors become gardeners, tending ever-growing branches rather than constraining ephemeral outputs.

CCS Concepts

• **Applied computing** → **Media arts**; • **Human-centered computing** → **Interaction design process and methods**; **Interactive systems and tools**.

Keywords

interactive narrative, interactive storytelling, human-AI co-creativity, large language models, authoring tools, research through design, computational creativity, generative AI

C&C '26, London, United Kingdom

ACM ISBN 978-1-4503-XXXX-X/2018/06
https://doi.org/XXXXXXX.XXXXXXX

1 Introduction

Designers of interactive digital narratives (IDNs) must negotiate a three-way tension between author, player, and system control over the narrative's direction [13, 20]. Imagine you've written a sword-and-sorcery fantasy where the player must defeat a dragon—but during playtesting, a player decides they'd rather befriend the dragon than slay it. To stray from the author's pre-defined paths is impossible in conventional IDNs such as *Choose Your Own Adventure*, but actively welcomed in LLM-based IDNs such as *AI Dungeon* [27]. LLMs grant IDNs the ability to improvisationally "play along" to the player's potentially unusual requests—but at a cost to authorial influence. An author who dislikes the new branch finds the LLM hard to constrain [26]; an author who would have embraced it was never consulted.

Pure LLM improvisation instantiates divergent, parallel universes with no shared reality underneath. For authors, the gap between their intent and realized player experience grows wider. LLM stochasticity produces outputs that are neither stable nor predictable—compounded by LLMs drifting, pulled by player tangents and their own improvisation [1, 25]—which makes it difficult for authors to anticipate what may emerge across playthroughs. Steering through prompting alone is cognitively burdensome and relies on brittle heuristics [31]. For players, an IDN that responds differently to the same input every time undermines both the value of replaying [22] and the social practice of sharing experiences through retellings [6, 15]. Stochastic, unpredictable emergent content disposed after each playthrough undermines both author and player trust in the system, including the *choice poetics* [21] the author constructs and the predictability of consequences experienced by the player [29].

Existing authoring frameworks treat IDNs as static artifacts once they are published, limiting authors' insight into and control over the storylines that emerge from unexpected player input in the wild. As Kreminski and Wardrip-Fruin argue, generative systems can be designed around either *mining* dynamics (disposable, on-demand outputs) or *gardening* dynamics (persistent artifacts that evolve through sustained engagement) [16]. Current LLM-based IDNs largely follow mining dynamics, where playthrough content is generated, consumed, and immediately discarded. We propose *cultivation* as a gardening-oriented alternative design metaphor: authors correct generations as they emerge through play, edits accumulate as intent rules, and alignment with authorial intent improves over time—shaping not only individual branches but future improvisations.

We make four contributions:

(1) *Bonsai*, an IDN authoring tool instantiating cultivation through four phases (seeding, growing, pruning, adapting), with analysis of how each addresses IDN authoring challenges;
(2) *Cultivation*, a design metaphor for LLM-based authoring in which narratives remain live after publication, accumulating design knowledge through use rather than calcifying at deployment;
(3) Three simulated experiments probing edit-based preference learning in IDN authoring;
(4) Design implications of the cultivation metaphor for authoring LLM-based interactive systems within and beyond IDN.

2 Related Work

Interactive digital narratives (IDNs) grant players what Murray calls "the satisfying power to take meaningful action and see the results of your decisions and choices" [23]. Yet player agency is fundamentally limited by what different IDN authoring forms allow. For example, conventional branching IDNs (e.g. those implemented via Twine) disallow paths that were not explicitly pre-defined by the author. As narrative complexity expands, it combinatorially expands the number of paths an author must write [13], limiting the degree of comprehensive player agency supported by the system. Modular content structures like storylets enable easier refactoring and emergent interaction design [17, 24], but still leave authors solely responsible for defining the choices that lead to meaningful progression.

LLMs offer a potential solution to these challenges through their ability to improvise coherent responses to arbitrary player input. Early systems like AI Dungeon [7] use LLMs to generate narrative content on demand. Pure prompt-based generation introduces failure modes: models hallucinate [14], misinterpret inputs [18], get derailed by player tangents [8], and lack understanding of narrative structure [2, 25]. Yet even the generated content that works is discarded after play, leaving authors with no mechanism to learn from or act on what actually emerged.

To address consistency issues, recent systems impose additional structure on LLM generation. *Dramamancer* combines modular storylets with authored preconditions that gate critical state changes while generating narration and dialogue in real-time [28]. Orchid uses upfront specifications to constrain narrative progression [30]. These approaches improve coherence, yet still require authors to anticipate player choices and specify all constraints upfront, giving authors no mechanism to correct what actually emerges across playthroughs.

Beyond IDN, creativity support tools have explored author-system alignment through preference learning from edits [9], edit-based writing rewards [3], revision-focused interfaces [32], and multimodal intent capture [4, 5]. These improve creative preference alignment in solo creative writing, but none addresses the triadic challenge of IDN authoring: balancing author intent, player agency, and system behavior simultaneously, across multiple playthroughs. No prior work has applied Kreminski and Wardrip-Fruin's gardening orientation [16] to IDN authoring with generative AI.

3 Bonsai: A Cultivation System

Bonsai is a web-based IDN authoring system built around *cultivation*: an authoring loop in which authorial intent emerges as a byproduct of ordinary use. Bonsai comprises three components (Figure 2): an authoring interface where authors write interactive scripts and curate generated content; a player interface for natural language interaction; and a generation engine that grows the narrative by matching player inputs to existing paths or improvising new branches, adding them to the live artifact. Unlike existing authoring frameworks that treat IDNs as static once published, Bonsai produces an artifact that evolves with author, player, and system.

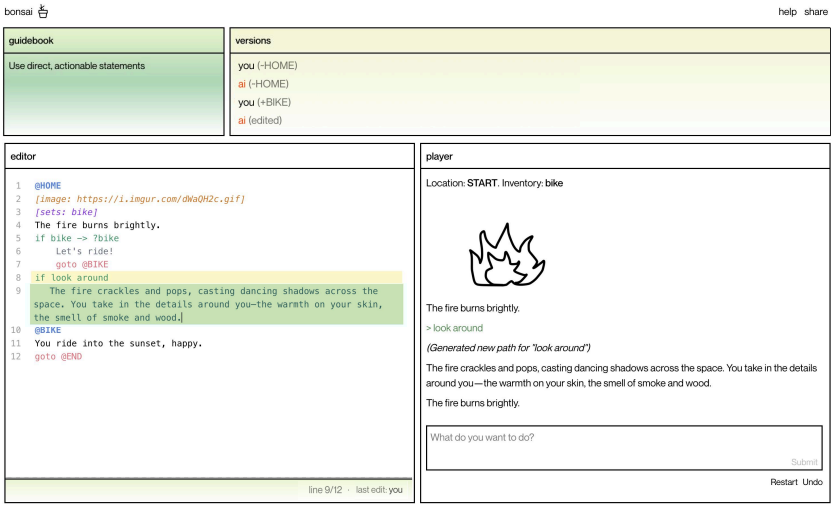

Figure 2: The Bonsai authoring interface showing a generated branch from player action *"look around."* The Editor (left) displays markup with diff highlighting: yellow for LLM-generated lines, green for author edits to generated content. The Player panel (right) shows the rendered story with natural language input. Versions panel (top-right) tracks all author and LLM changes chronologically. The Guidebook (top-left) provides project-wide generation guidance.

The system uses Claude 3.5 Haiku for all generation tasks, selected for its balance of creative writing quality and low latency.

Seeding from initial author content. Authors construct narrative material using a scripting language based on Ink [11], a widely-adopted markup language for narrative games used in numerous award-winning titles [12]. Scripts are composed of scenes and *branches*, which are pairs of conditions and consequences, e.g. "if Alice eats the cake *then* Alice grows larger". The language also supports basic narrative control flow including state and scene changes.

Growing from player interaction. When a player submits an input, an LLM is used to fuzzily match it against existing branch conditions within the current scene. If a strong semantic match is found, the player is routed to the matching branch, experiencing it exactly as written. Otherwise, the LLM generates a new branch conditioned on narrative context and player input, which immediately persists to the artifact. Generations are strictly additive, and existing authored content is never overwritten.

Pruning generated branches. Authors review generated content via a unified version history with attribution ("AI" vs. "author"). In the authoring interface, freshly generated lines are highlighted in yellow until they are edited by the author (highlighted in green) or resolved.

Learning from author edits. When authors edit generated content, an LLM extracts a natural-language author-intent rule from the diff (e.g. softening a violent outcome produces "prefer non-violent resolutions"). Rules accumulate in a visible guidebook and guide future branch generation, so each author edit changes how the artifact grows.

4 Design Journal

We implemented Bonsai through three months of first-person research [19] informed by autobiographical design [10, 33], drawing on our experience as IDN authors with published works each played tens of thousands of times. We iterated on two narrative scenarios with three exploratory testers from AI and creative writing backgrounds (Figure 3).

Our first scenario, `fireplace-expanded`, is a bare-bones narrative where the player sits by a fire and decides what to do next. Its seed content contains two scenes (`HOME` and `BIKE`) and one predefined path (`if bike, then goto BIKE`). Testers responded positively to visible semantic-match confidence (e.g., matching "look around for monsters" to "look around" at 85%). But when a player entered an absurdist provocation, the system generated a flat, deflecting response; rewriting that output into a more dramatic consequence produced a useful guidebook rule ("Embrace consequences: let player choices have real, sometimes painful results").

Our second scenario, `escaperoom`, is a tiny puzzle game seeded with two scenes (`START` and `DESK`). Unlike `fireplace-expanded`, this project needed stronger constraints: generating new scenes could break or distract from the puzzle goal. This tension led to a "freedom" slider in guidebook settings, letting authors trade narrative flexibility for puzzle integrity. The same project exposed "vector similarity hell": one compelling generated branch (a hidden desk compartment) became over-matched to unrelated inputs until authors manually pruned and corrected the surrounding branches.

4.1 Experiments

To complement our first-person design account, we ran three focused formative experiments across three IDN projects. We used

Figure 3: Bonsai authoring and player interfaces showing two projects with different design approaches: `fireplace-expanded` (left) is an open-ended exploration scenario, while `escaperoom` (right) is a constrained puzzle with defined objects and goals. Note the contrast in their Guidebooks (top-left of each): `fireplace-expanded` accumulates open-ended dramatic rules, while `escaperoom` builds constraint-heavy rules that preserve puzzle integrity.

an LLM-simulated author as a proxy for human judgment—a valid stand-in for low-cost directional signal in early-stage design research, though not a substitute for user studies. In each experiment, a set of generated branches (train and held-out scenes) were rated by a second LLM judge on a 1 (no alignment with author intent) to 5 (full alignment) scale under four guidebook conditions.

	PREFERENCE		STYLE	
Condition	train	held-out	train	held-out
none (abs.)	3.17	2.31	3.97	3.22
style	+0.33	+0.39	**+0.42**	**+0.64**
intent	**+1.83**	**+2.56**	−0.61	0.00
wrong intent	−0.86	−0.44	−0.72	−0.28

Table 1: Mean ratings by condition (1 = no alignment with author intent, 5 = full alignment). The *none* row reports raw scores; all other rows report change relative to *none*. Bold marks the best value within each column.

Project specificity. Intent rules from the correct project sharply lift preference alignment; rules from the wrong project actively hurt (−0.44 held-out), confirming that cultivation memory should default to project-scoped histories rather than shared rule pools.

Generalization. Intent rules lifted preference on both construction and held-out scenes, with a larger gain on held-out (+2.56 vs. +1.83), suggesting curation accumulates reusable design knowledge rather than merely patching local branches. Style rules produce the largest style lift but a small preference lift, revealing a voice–intent tradeoff: the seed script is an incomplete proxy for author intent.

Medium-aware extraction. Naive extraction nearly missed world-constraint preferences (8% recovery). Extraction prompted with IDN authoring categories (world, narrator, player interactions) recovered them at 83%. What the system learns from edits depends on the representational vocabulary it is given.

Two worked traces illustrate how a single edit becomes a reusable rule. In `fireplace`, rewriting a flat response to "talk to the fire" into an absurdist escalation yielded: *"escalate dramatically; give elements with agency explicit opinions rather than leaving them open-ended."* In `escaperoom`, deleting one clause from the "pick the lock" response yielded: *"Do not mention the absence of objects—this inadvertently introduces new concepts."* The first captures prose voice; the second captures an IDN-specific interaction constraint, and the two require different extraction vocabulary to surface.

Early testing also revealed limitations in our initial branch-management interface, which treated generated branches as separate from author edits and caused review confusion. We redesigned this into a unified version history with unresolved generated content highlighted in-context, making branch triage and pruning more legible. We also found that learning from approvals and deletions produced noisy guidebook updates; only inline edits consistently surfaced authorial intent.

Across these scenarios and probes, we found cultivation viable but tension-filled. Seeding requires balancing openness and constraint (`fireplace-expanded` vs. `escaperoom`); growing and pruning were intuitive; adapting remained the hardest phase. At its core, cultivation reframes generation errors as expected material for curation rather than as terminal failures.

5 Discussion

Cultivation's central provocation is that an IDN need not be static once published. Rather than treating a shipped narrative as a fixed artifact that authors must fully specify in advance, cultivation keeps the narrative *live*—accumulating design knowledge through use and correction, shaped by the interplay between actual player behavior, authorial curation, and system interpretation. The branching script is not a means to a finished story but an artifact whose form reflects its history of growth and pruning, much as a bonsai's shape is inseparable from the decisions made in tending it.

This reframing suggests three design principles. First, *additive-only generation*: the system only proposes new content, never overwriting existing material, positioning AI as an apprentice learning from authorial judgment. Second, *self-limiting assistance*: as authored and accepted content accumulates, fewer player actions remain undefined, reducing generation burden over time. Third, *distributed authoring over time*: rather than requiring authors to imagine the full possibility space upfront, cultivation makes authorship reactive—authors respond to what players actually do, and the artifact improves through actual play.

Our experiments sharpen these principles. The failure of cross-project rule transfer reinforces distributed authorship as situated: intent is tied to a specific artifact and its evolving constraints. Improvements on unseen scenes support self-limiting assistance by showing that curation produces reusable guidance beyond individual patches. The extraction gap reinforces that representational choices in the authoring tool determine which forms of intent are learnable—preference learning should be treated as an interaction design problem, not only a modeling one. We also observed a voice–intent tradeoff: intent rules occasionally reduced style consistency, because the seed script is an incomplete proxy for author intent. For co-creative authoring, explicit preferences can deliberately move generations beyond the seed's initial voice—a feature, not a failure.

5.1 Limitations

Cultivation redistributes authorial labor across time rather than eliminating it. Persistence introduces fragility at scale: review mechanisms that work at dozens of branches may not hold at hundreds. Preference accumulation faces the same tension: small edits produced noisy Guidebook entries, meaningful patterns only emerged from edits with clear dramatic intent, and accumulated preferences risk calcifying into invisible constraints. Player interaction preferences—governing which actions are possible, redirected, or prevented—proved the hardest to extract and remain the most IDN-specific open challenge. Finally, while advances in model capability may reduce reliance on explicit preference mechanisms, the structural challenge of reactive, cross-playthrough authorial curation is independent of model quality: a more capable model still generates and discards, leaving authors with no mechanism to learn from what players actually did.

6 Conclusion

In this paper we explored *cultivation*: a design metaphor for human-AI collaboration in which AI-generated artifacts remain "live" after publication, gradually accumulating design knowledge through authorial curation rather than calcifying at deployment. Via the LLM-based IDN authoring tool *Bonsai*, we demonstrated that treating generated content as persistent, editable material reframes authorship: shifting effort from anticipating what might *possibly* emerge from player-AI interaction toward curating what *actually* does. Beyond IDNs, cultivation suggests a broader orientation for creative AI tools—one where the effort of correction is never wasted, and the artifact accumulates the shape of its tending.

References

[1] Nina Beguš. 2024. Experimental narratives: A comparison of human crowdsourced storytelling and AI storytelling. *Humanities and Social Sciences Communications* 11, 1 (Oct. 2024). doi:10.1057/s41599-024-03868-8

[2] Tuhin Chakrabarty, Philippe Laban, Divyansh Agarwal, Smaranda Muresan, and Chien-Sheng Wu. 2024. Art or Artifice? Large Language Models and the False Promise of Creativity. arXiv:2309.14556 [cs.CL] https://arxiv.org/abs/2309.14556

[3] Tuhin Chakrabarty, Philippe Laban, and Chien-Sheng Wu. 2025. AI-Slop to AI-Polish? Aligning Language Models through Edit-Based Writing Rewards and Test-time Computation. arXiv:2504.07532 [cs.CL] https://arxiv.org/abs/2504.07532

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

[5] John Joon Young Chung, Melissa Roemmele, and Max Kreminski. 2025. Toyteller: AI-powered Visual Storytelling Through Toy-Playing with Character Symbols. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. ACM, 1–23. doi:10.1145/3706598.3713435

[6] Mirjam Palosaari Eladhari. 2018. Re-tellings: the fourth layer of narrative as an instrument for critique. In *International Conference on Interactive Digital Storytelling*. Springer, 65–78.

[7] Cátia Ferreira. 2025. Genre, Bias, and Narrative Logic in AI Dungeon: Generative AI as a Game-Based Storytelling Engine. *Hipertext.net* 31 (Nov. 2025), 77–89. doi:10.31009/hipertext.net.2025.i31.08

[8] Roberto Gallotta, Graham Todd, Marvin Zammit, Sam Earle, Antonios Liapis, Julian Togelius, and Georgios N. Yannakakis. 2024. Large Language Models and Games: A Survey and Roadmap. *IEEE Transactions on Games* (2024), 1–18. doi:10.1109/tg.2024.3461510

[9] Ge Gao, Alexey Taymanov, Eduardo Salinas, Paul Mineiro, and Dipendra Misra. 2024. Aligning LLM Agents by Learning Latent Preference from User Edits. arXiv:2404.15269 [cs.CL] https://arxiv.org/abs/2404.15269

[10] William Gaver. 2012. What should we expect from research through design?. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 937–946.

[11] Inkle Studios. 2016. Ink: Inkle's Narrative Scripting Language. https://www.inklestudios.com/ink/. Accessed: February 2, 2026.

[12] Inkle Studios. 2026. inkle/ink-library: A collection of ink samples, tools and a list of projects that use ink. https://github.com/inkle/ink-library#ink-games-and-non-games. Accessed: February 3, 2026.

[13] Joey Donald Jones and David Millard. 2024. Experiencing The Authorial Burden. In *Proceedings of the 35th ACM Conference on Hypertext and Social Media* (Poznan, Poland) *(HT '24)*. Association for Computing Machinery, New York, NY, USA, 78–87. doi:10.1145/3648188.3675134

[14] Adam Tauman Kalai, Ofir Nachum, Santosh S. Vempala, and Edwin Zhang. 2025. Why Language Models Hallucinate. arXiv:2509.04664 [cs.CL] https://arxiv.org/abs/2509.04664

[15] Max Kreminski, Ben Samuel, Edward Melcer, and Noah Wardrip-Fruin. 2019. Evaluating AI-based games through retellings. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, Vol. 15. 45–51.

[16] Max Kreminski and Noah Wardrip-Fruin. 2018. Gardening Games : An Alternative Philosophy of PCG in Games. In *FDG Workshop on Procedural Content Generation in Games*.

[17] Max Kreminski and Noah Wardrip-Fruin. 2018. *Sketching a Map of the Storylets Design Space*. 160–164. doi:10.1007/978-3-030-04028-4_14

[18] Yiheng Liu, Tianle Han, Siyuan Ma, Jiayue Zhang, Yuanyuan Yang, Jiaming Tian, Hao He, Antong Li, Mengshen He, Zhengliang Liu, Zihao Wu, Lin Zhao, Dajiang Zhu, Xiang Li, Ning Qiang, Dingang Shen, Tianming Liu, and Bao Ge. 2023. Summary of ChatGPT-Related research and perspective towards the future of large language models. *Meta-Radiology* 1, 2 (2023), 100017. doi:10.1016/j.metrad.2023.100017

[19] Andrés Lucero, Audrey Desjardins, Carman Neustaedter, Kristina Höök, Marc Hassenzahl, and Marta E Cecchinato. 2019. A sample of one: First-person research methods in HCI. In *Companion Publication of the 2019 Designing Interactive Systems Conference*. 385–388.

[20] Stacey Mason. 2021. *Responsiveness in Narrative Systems*. Ph. D. Dissertation. UC Santa Cruz. https://escholarship.org/uc/item/9gq229h4 ProQuest ID: Mason_ucsc_0036E_12374.

[21] Peter Mawhorter, Michael Mateas, Noah Wardrip-Fruin, and Arnav Jhala. 2014. Towards a theory of choice poetics. In *In Proceedings of the 9th International Conference on the Foundations of Digital Games*. http://fdg2014.org/papers/fdg2014_paper_19.pdf

[22] Alex Mitchell. 2025. Understanding the Kaleidoscopic Nature of Interactive Digital Narratives Through Repeat Experience. In *International Conference on Interactive Digital Storytelling*. Springer, 1–14.

[23] Janet H. Murray. 1997. *Hamlet on the Holodeck: The Future of Narrative in Cyberspace.* The Free Press, New York.

[24] Emily Short. 2016. Beyond Branching: Quality-Based and Salience-Based Narrative Structures. https://emshort.blog/2016/04/12/beyond-branching-quality-based-and-salience-based-narrative-structures/. Accessed: 2026-01-15.

[25] Yufei Tian, Tenghao Huang, Miri Liu, Derek Jiang, Alexander Spangher, Muhao Chen, Jonathan May, and Nanyun Peng. 2024. Are Large Language Models Capable of Generating Human-Level Narratives? arXiv:2407.13248 [cs.CL] https://arxiv.org/abs/2407.13248

[26] Sam Toyer, Olivia Watkins, Ethan Adrian Mendes, Justin Svegliato, Luke Bailey, Tiffany Wang, Isaac Ong, Karim Elmaaroufi, Pieter Abbeel, Trevor Darrell, Alan Ritter, and Stuart Russell. 2023. Tensor Trust: Interpretable Prompt Injection Attacks from an Online Game. https://arxiv.org/pdf/2311.01011.pdf

[27] Nick Walton. 2019. AI Dungeon 2. https://aidungeon.io/ Accessed: February 2, 2026.

[28] Tiffany Wang, Yuqian Sun, Yi Wang, Melissa Roemmele, John Joon Young Chung, and Max Kreminski. 2025. Dramamancer: Interactive Narratives with LLM-powered Storylets. In *Adjunct Proceedings of the 38th Annual ACM Symposium on User Interface Software and Technology (UIST Adjunct '25)*. Association for Computing Machinery, New York, NY, USA, Article 33, 5 pages. doi:10.1145/3746058.3758995

[29] Noah Wardrip-Fruin, Michael Mateas, Steven Dow, and Serdar Sali. 2009. Agency reconsidered. In *Proceedings of DiGRA 2009 Conference: Breaking New Ground: Innovation in Games, Play, Practice and Theory.*

[30] Zhen Wu, Serkan Kumyol, Shing Yin Wong, Xiaozhu Hu, Xin Tong, and Tristan Braud. 2025. Orchid: A Creative Approach for Authoring LLM-Driven Interactive Narratives. In *Proceedings of the 2025 Conference on Creativity and Cognition (C&C '25)*. Association for Computing Machinery, New York, NY, USA, 774–791. doi:10.1145/3698061.3726906

[31] J.D. Zamfirescu-Pereira, Richmond Y. Wong, Bjoern Hartmann, and Qian Yang. 2023. Why Johnny Can't Prompt: How Non-AI Experts Try (and Fail) to Design LLM Prompts. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 437, 21 pages. doi:10.1145/3544548.3581388

[32] David Zhou and Sarah Sterman. 2024. Ai.llude: Investigating Rewriting AI-Generated Text to Support Creative Expression. In *Creativity and Cognition*. ACM, 241–254. doi:10.1145/3635636.3656187

[33] John Zimmerman, Jodi Forlizzi, and Shelley Evenson. 2007. Research through design as a method for interaction design research in HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '07)*. Association for Computing Machinery, New York, NY, USA, 493–502. doi:10.1145/1240624.1240704