InfiniteCATs: Semantic Crafting as a Substrate for Creative Activity Tracing

NICHOLAS JENNINGS*, Stanford University, USA
SHM GARANGANAO ALMEDA*, University of California, Berkeley, USA
SAM PODER, University of California, Berkeley, USA
BJÖRN HARTMANN, University of California, Berkeley, USA

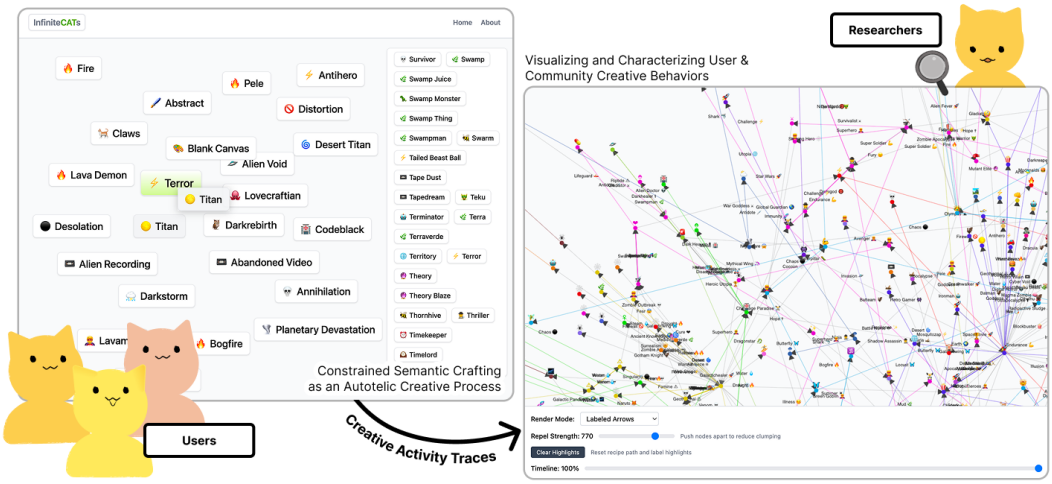

Fig. 1. (Left) the infiniteCATs user interface. (Right) an example of graphing the resource creation space using labeled edges

Large Language Models are increasingly intermediating creative processes, shaping the spaces that users explore, and how they reach them. We lack insight into how LLM-intermediation is implicitly shaping creative exploration processes, and user's strategies for navigating this. Players of semantic crafting games like Infinite Craft merge concepts to craft new ones, from a constrained set of initial options; each combination result is determined by an LLM. This simple, yet expressively open-ended substrate presents a well-scoped environment for observing LLM-mediated, intrinsically motivated creative activity. We present InfiniteCATs, a semantic crafting system, and collect and analyze thousands of actions from 17 users in a preliminary pilot study. We demonstrate InfiniteCATs' potential as a testbed for capturing user and community-level creative interactions in LLM-mediated semantic space. Future work includes further developing these creative activity data collection and analysis tools as open research community resources, and applying them to large-scale creative activity studies.

*Both authors contributed equally to this research.

Authors' Contact Information: Nicholas Jennings, Stanford University, Stanford, CA, USA; Shm Garanganao Almeda, University of California, Berkeley, Berkeley, CA, USA; Sam Poder, University of California, Berkeley, Berkeley, CA, USA; Björn Hartmann, bjoern@eecs.berkeley.edu, University of California, Berkeley, Berkeley, California, USA.

Manuscript submitted to ACM

CCS Concepts: • **Human-centered computing → Visualization techniques**; • **Applied computing → Arts and humanities**.

Additional Key Words and Phrases: Creativity Support Tools, Visualization

ACM Reference Format:
Nicholas Jennings, Shm Garanganao Almeda, Sam Poder, and Björn Hartmann. 2026. InfiniteCATs: Semantic Crafting as a Substrate for Creative Activity Tracing. In *Proceedings of xx.xxxxx.x.x.x. (CHI'26)*. ACM, New York, NY, USA, 6 pages. https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

Large Language Models are increasingly being used to intermediate creative processes. LLMs have been shown to have homogenizing effects on creative behavior [4]; we lack insight into how these black box systems can implicitly influence user's creative exploration processes, shaping the space of possibilities they consider, and how they reach them.

Glow et. al. identify crafting as a useful activity for playful and creative expression [11]. *Semantic* crafting games like Infinite Craft [1] allows players to "craft" new concepts by merging existing ones (e.g., 💧 *Water* + 🔥 *Fire* = 💧*Steam*), starting from a constrained set of initial input options — the result of each combination is determined by an LLM. The game has attracted a growing community of users intrinsically-motivated by the exploration process itself, and different modes of play have emerged. In competitive speedrunning, players attempt to craft a particular "target" before their opponent, racing across the semantic space accessible *through* the LLM [19]. Prolific players demonstrate a well-developed sense of *how* to navigate *through* the model, showcasing creative strategies designed to avoid semantic "pitfalls" to achieve their crafting goals.

LLM-mediated semantic crafting affords flexibly open-ended interaction, but with input-level constraints that add structure not present in natural language prompting. As a constrained, yet expressive substrate [17], semantic crafting presents as a uniquely well-scoped testbed for collecting data on LLM-mediated creative activity.

To this end, we built InfiniteCATs[1], a semantic crafting system instrumented to collect, and analyze, user activity. In a preliminary pilot study with 17 participants, we collected and analyzed thousands of user actions.

We briefly present data visualization and analysis strategies that demonstrate semantic crafting's potential as a testbed for studying user and community-level navigation of LLM-mediated semantic spaces. Future work includes deploying InfiniteCATs in a large-scale study, and further developing its activity tracing and analysis tools as adaptable open-source resources for the research community. We discuss the variety of research questions that InfiniteCATs might be instrumented to investigate, towards insights into *how* different user orientations or interface configurations can shape LLM-mediated exploration processes.

2 RELATED WORK

We draw from *playtracing* in games, and the study of design space exploration (DSE) and creative activity traces (CATs) in HCI. Quantifying design spaces has been used to procedurally generate puzzles [27], and visualize level design strategies [3]. Activity tracing has been used to track player behavior over time in multiplayer settings [20], to explore player motivation [15], and to evaluate collaborative behavior [9]. Analyzing player traces with regular expressions has been shown to be useful for formal verification and finding representative traces [22].

[1] https://infinitecat.vercel.app/

Activity tracing has also been done outside of the context of games, for example, to generate suggestions for creative users exploring software based on other users' traces [18]. CAT and DSE visualization and analysis methods have been used to understand and aid creative processes [12, 14, 25].

Embeddings-based methods are useful in supporting sensemaking open-ended, co-creative interactions with AI [2, 24, 26, 28]. While open-endedness (i.e., natural language prompting) characterizes many users' interactions with LLMs, open-ended creative activity data can be difficult to make sense of; it is difficult to draw meaningful links between different states in an expansive, heterogeneous design space. Embeddings-based approaches like Fuzzy Linkography [26] address this by relaxing the notion of a "link" until graphing becomes *possible*. InfiniteCATs reduces the input space until graph links become *obvious*.

3 SYSTEM

We developed InfiniteCATs[2] atop *OpenCraft* [6], an open source semantic crafting system created to mimic Neal Agarwal's Infinite Craft [1], which itself derives from Jakub Koziol's Little Alchemy [13]. The system is implemented as a Vue.js + TypeScript.js frontend, supported by a Flask server backend written in Python. The Flask server handles requests to store and retrieve information from a database that includes every action taken by a user and every unique resource that has been discovered so far.

Users are presented with a list of *resources*, represented as a card featuring the resource's name and a representative emoji. Users start with only four resources: 💧 *Water*, 🔥 *Fire*, 🌍 *Earth*, and 💨 *Air*. Users can drag cards onto the canvas; dragging one resource onto another merges them, forming a new resource. If those two particular resources have not been merged before, the an LLM is prompted to choose a new resource to result from the merge. All merges are cached, with the older result used for repeated merges. All resources consist of an emoji and an ASCII "name". While there is no character limit on the resource name, the LLM is prompted for "a single word or short phrase". We created multiple graph formats for rendering each user's resource network, with the default (seen in figure 2) being a bipartite graph with each resource represented by a single node, and combinations represented by "combination nodes" with incoming edges from the two recipe components, and and outgoing edge to the result.

4 INITIAL OBSERVATIONS

For our preliminary pilot study, we gave 17 participants a link to InfiniteCATs and invited them to experiment with the game. A total of 12,004 (11,112 unique) combinations were made. A majority of the total combinations came from 4 "power-users" who each contributed 4916, 2179, 1378, and 1191 combinations respectively. (We refer to users as P1, P2, etc. by descending order of their combination count.)

Semantic crafting presents units of language as objects with material affordances. A particular resource does not always have a consistent effect in combination; this unpredictability drives users' curiosity towards serendipitous discoveries. We tested using the normalized average of the two input vectors to predict the combination function in embedding space, scoring the cosine similarity between each predicted vs. actual result. Combinations involving *Fungus* (P1, P5) (n=15) had an average score of 0.630, and often created results that used synonyms of the input resources (e.g. 🚽 *Fungus* + 👑 *Empire* = 👑 *Mold Kingdom*). Conversely, combinations involving 😒 *Oxymoron* (P1, P2, P9) (n = 23) had an average score of 0.300, and often resulted in antonyms of the other input word. The heterogeneity of combinations lends towards more complex exploration strategies.

[2]https://github.com/s-almeda/infiniteCATs

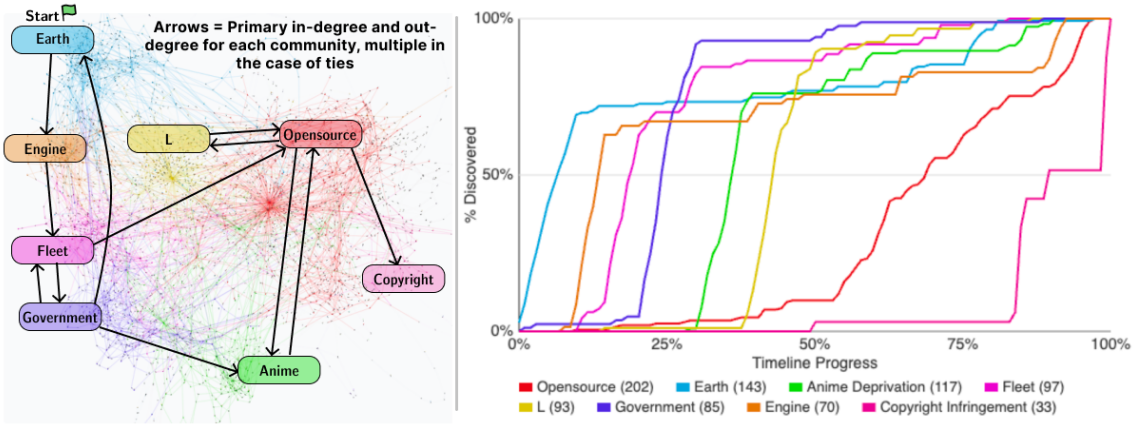

Fig. 2. P2's combination node graph with the top 8 communities highlighted. (Left) The communities highlighted on their network. Arrows between communities represent each community's top incoming and outgoing connections among the top 8. The Fleet community has an effective tie for most incoming edges, so both are shown. (Right) The completion of each community over time.

Our pilot users noticed resources that had more consistent or distinct effects in combination, and began to incorporate them into their exploration strategies. Some words have semantic convergence: of the 134 unique combinations involving 🌎 *Opensource*, 26 resulted in 🖥️ *Linux* (P2). Some words have *orthographic* affordance: 🕵️ *L* and 🤍 *Bl* allowed P2 to "prefix" other resources with those letters; 👠 *Maid* (P1) afforded suffixation (🏴 *Rebel* + 👠 *Maid* = 🤘 *Punk Maid*).

We can surmise a simplified story of a user's activity by partitioning their graph into *communities* (we used the Louvain Method [5, 8]) and exploring which communities were explored when, e.g., by plotting the completion percentage of their largest communities over time. In P2's community completion graph (Fig 2), a series of steep curves followed by gentle curves indicate initially exploring a frequently shifting series of topics, then more deeply exploring an area (*Opensource*) later on. Their movement between different communities resembles a *curiosity zigzag*[26].

5 DISCUSSION AND NEXT STEPS

InfiniteCATs readily facilitates *casual*, autotelic creative activity [7] from users without any particular prerequisite experience, and constrains the possibility space such that data tracing this creative activity remains interpretable. At the same time, its open-ended expressivity can support a wide variety of complex activities and interaction behaviors. This makes InfiniteCATs a promising testbed for exploring a variety of research questions in future work.

Experiments designed to direct participants' objectives while collecting qualitative (e.g., think-aloud) data, and activity traces annotated with a user's objective or strategy at that time, could provide useful insights into **how different objectives shape process**, e.g. for mixed-initiative creative interface design [10, 16] and user modeling [23]. User strategies may change depending on their objective (e.g., crafting a particular resource vs. freely exploring.) Users seeking *novel* discoveries may push the "edges" of global space, a behavior characterized elsewhere as *curious* creation [14, 21]. The CAT exploration tools developed so far focus on analyzing the resource network. We are interested in using embedding space analysis methods (e.g., Wizmap [29], Fuzzy Linkography [26]) to **analyze how per-user trajectories and the global exploration graphs move across semantic embedding space**. Altering the generation system (e.g. preventing duplicate combination recipes) can alter network structure. By isolating certain network changes, we can investigate **how system constraints or guardrails shape user strategies**. We designed InfiniteCATs to

display a dynamically updating graph of users' exploration activity as they craft, but chose to hide it for the pilot. **How do strategies change when users can see their own, or their community's, creative activity?**

6 CONCLUSION

We present InfiniteCATs, an LLM-mediated semantic crafting system, as a testbed for tracing and analyzing autotelic creative exploration activity. We use InfiniteCATs' trace analysis tools on findings from a pilot study with 17 participants, demonstrating its potential as a configurable experimental substrate. These preliminary results suggest semantic crafting as a well-constrained, yet expressively flexible space for capturing and studying user and community-level creative explorations of LLM-mediated semantic space.

REFERENCES

[1] Neal Agarwal. 2024. Infinite Craft. https://neal.fun/infinite-craft/

[2] Shm Garanganao Almeda, John Joon Young Chung, Bjoern Hartmann, Sophia Liu, Brett Halperin, Yuwen Lu, and Max Kreminski. 2025. Artographer: a Curatorial Interface for Art Space Exploration. https://doi.org/10.48550/arXiv.2512.02288 arXiv:2512.02288 [cs].

[3] Alberto Alvarez, Jose Font, and Julian Togelius. 2022. Toward Designer Modeling Through Design Style Clustering. *IEEE Transactions on Games* 14, 4 (2022), 676–686. https://doi.org/10.1109/TG.2022.3143800

[4] Barrett R Anderson, Jash Hemant Shah, and Max Kreminski. 2024. Homogenization Effects of Large Language Models on Human Creative Ideation. In *Proceedings of the 16th Conference on Creativity & Cognition (C&C '24)*. Association for Computing Machinery, New York, NY, USA, 413–425. https://doi.org/10.1145/3635636.3656204

[5] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. 2008. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment* 2008, 10 (2008), P10008.

[6] @bufferhead code. 2024. OpenCraft. https://github.com/bufferhead-code/opencraft

[7] Katherine E. Compton. 2019. *Casual Creators: Defining a Genre of Autotelic Creativity Support Systems.* Ph. D. Dissertation. UC Santa Cruz. https://escholarship.org/uc/item/4kg8g9gd

[8] Corneliu[s]. 2013. Upphiminn/jLouvain: Louvain community detection for Javascript. https://github.com/upphiminn/jLouvain

[9] Damian Debkowski, Andrew Marrero, Nicole Yson, Li Yin, Yichen Yue, Seth Frey, and Mubbasir Kapadia. 2016. Contained: Using multiplayer online games to quantify success of collaborative group behavior. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, Vol. 12. 149–154.

[10] Sebastian Deterding, Jonathan Hook, Rebecca Fiebrink, Marco Gillies, Jeremy Gow, Memo Akten, Gillian Smith, Antonios Liapis, and Kate Compton. 2017. Mixed-Initiative Creative Interfaces. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '17)*. Association for Computing Machinery, New York, NY, USA, 628–635. https://doi.org/10.1145/3027063.3027072

[11] April Grow, Melanie Dickinson, Johnathan Pagnutti, Noah Wardrip-Fruin, and Michael Mateas. 2017. Crafting in games. *Digital Humanities Quarterly* 11, 4 (2017).

[12] Noor Hammad. 2025. Towards Trail-Aware Digital Content Creation Tools. In *Companion Proceedings of the Annual Symposium on Computer-Human Interaction in Play (CHI PLAY Companion '25)*. Association for Computing Machinery, New York, NY, USA, 273–275. https://doi.org/10.1145/3744736.3749328

[13] Jakub Koziol. 2010. Little Alchemy. https://littlealchemy.com

[14] Max Kreminski, Isaac Karth, Michael Mateas, and Noah Wardrip-Fruin. 2022. Evaluating Mixed-Initiative Creative Interfaces via Expressive Range Coverage Analysis. In *Joint Proceedings of the ACM IUI Workshops 2022*. ACM, Helsinki, FInland. https://mkremins.github.io/publications/ERaCA_HAI-GEN2022.pdf

[15] Joomi Lee and Allison Eden. 2023. How Motivation and Digital Affordances Shape User Behavior in a Virtual World. *Media Psychology* 26, 5 (2023), 551–578.

[16] Antonios Liapis, Georgios N. Yannakakis, Constantine Alexopoulos, and Phil Lopes. 2016. Can Computers Foster Human Users' Creativity? Theory and Praxis of Mixed-Initiative Co-Creativity. *Digital Culture & Education (DCE)* 8, 2 (2016), 136–152.

[17] Wendy E. Mackay and Michel Beaudouin-Lafon. 2025. Interaction Substrates: Combining Power and Simplicity in Interactive Systems. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, 1–16. https://doi.org/10.1145/3706598.3714006

[18] Justin Matejka, Wei Li, Tovi Grossman, and George Fitzmaurice. 2009. CommunityCommands: command recommendations for software applications. In *Proceedings of the 22nd Annual ACM Symposium on User Interface Software and Technology* (Victoria, BC, Canada) *(UIST '09)*. Association for Computing Machinery, New York, NY, USA, 193–202. https://doi.org/10.1145/1622176.1622214

[19] Tanner Charles "SmallAnt" Minecraft. 2025. Speedrunning Words in InfiniteCraft - REMATCH vs PointCrow. https://www.youtube.com/watch?v=UdAwHKkQnAE

[20] Stephan Mueller, Barbara Solenthaler, Mubbasir Kapadia, Seth Frey, Severin Klingler, Richard P. Mann, Robert W. Sumner, and Markus Gross. 2015. HeapCraft: interactive data exploration and visualization tools for understanding and influencing player behavior in Minecraft. In *Proceedings of the 8th ACM SIGGRAPH Conference on Motion in Games* (Paris, France) *(MIG '15)*. Association for Computing Machinery, New York, NY, USA, 237–241. https://doi.org/10.1145/2822013.2822033

[21] Mark J. Nelson, Swen E. Gaudl, Simon Colton, and Sebastian Deterding. 2018. Curious users of casual creators. In *Proceedings of the 13th International Conference on the Foundations of Digital Games (FDG '18)*. Association for Computing Machinery, New York, NY, USA, 1–6. https://doi.org/10.1145/3235765.3235826

[22] J. Osborn, B. Samuel, Michael Mateas, and Noah Wardrip-Fruin. 2021. Playspecs: Regular Expressions for Game Play Traces. In *Artificial Intelligence and Interactive Digital Entertainment Conference*. https://doi.org/10.1609/aiide.v11i1.12803

[23] Omar Shaikh, Shardul Sapkota, Shan Rizvi, Eric Horvitz, Joon Sung Park, Diyi Yang, and Michael S. Bernstein. 2025. Creating General User Models from Computer Use. In *Proceedings of the 38th Annual ACM Symposium on User Interface Software and Technology (UIST '25)*. Association for Computing Machinery, New York, NY, USA, 1–23. https://doi.org/10.1145/3746059.3747722

[24] Hanshu Shen, Lyukesheng Shen, Wenqi Wu, and Kejun Zhang. 2025. IdeationWeb: Tracking the Evolution of Design Ideas in Human-AI Co-Creation. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, 1–19. https://doi.org/10.1145/3706598.3713375

[25] Evan Shimizu, Matthew Fisher, Sylvain Paris, James McCann, and Kayvon Fatahalian. 2020. Design Adjectives: A Framework for Interactive Model-Guided Exploration of Parameterized Design Spaces. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) *(UIST '20)*. Association for Computing Machinery, New York, NY, USA, 261–278. https://doi.org/10.1145/3379337.3415866

[26] Amy Smith, Barrett R Anderson, Jasmine Tan Otto, Isaac Karth, Yuqian Sun, John Joon Young Chung, Melissa Roemmele, and Max Kreminski. 2025. Fuzzy Linkography: Automatic Graphical Summarization of Creative Activity Traces. In *Proceedings of the 2025 Conference on Creativity and Cognition*. 637–650.

[27] Adam M Smith, Eric Butler, and Zoran Popovic. 2013. Quantifying over play: Constraining undesirable solutions in puzzle design.. In *FDG*. 221–228.

[28] Sangho Suh, Meng Chen, Bryan Min, Toby Jia-Jun Li, and Haijun Xia. 2024. Luminate: Structured Generation and Exploration of Design Space with Large Language Models for Human-AI Co-Creation. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 644, 26 pages. https://doi.org/10.1145/3613904.3642400

[29] Zijie J. Wang, Fred Hohman, and Duen Horng Chau. 2023. WizMap: Scalable Interactive Visualization for Exploring Large Machine Learning Embeddings. arXiv:2306.09328 [cs.LG] https://arxiv.org/abs/2306.09328

Received 22 January 2026