

# Visual theatrical improvisation alongside Artificial Intelligence image generators

**Piotr Mirowski**

Improbobotics  
London, UK  
piotr@improbobotics.org

**Boyd Branch**

Coventry University  
Coventry, UK  
boyd@improbobotics.org

**Kory Mathewson**

Improbobotics  
Montréal, QC, Canada  
kory@improbobotics.org

## Abstract

We describe and reflect on the use of image generation in the context of live, interactive performances, performed in front of theatre audiences over 2021-2024. Our performance format centers human improvisational actors reacting to and inspiring real-time image generation, establishing a three-way dialogue between actors, image prompter and curator, and audiences providing suggestions. In this way, our comedic show engages theatre audiences on the ethics of generative AI.

## Introduction

“Show, don’t tell” is a storytelling and science communication principle. Our theatre company, Improbobotics<sup>1</sup>, employs it to demonstrate to lay theatre audiences the formidable advances in generative artificial intelligence, focusing on image generation.

Our work is situated within a larger context of improvisational games with AI tools. Improvisation with physical robots has been pioneered by Bruce et al. (2000), and improvised comedy with physical robots and Large-Language Model (LLM) chatbots has been introduced by Mathewson and Mirowski (2017b), alongside the improvised comedy Turing test (Mathewson and Mirowski, 2017a) and the idea of a human actors performing as *Cyborgs* who take their lines from language models (Mathewson and Mirowski, 2018; Mirowski and Mathewson, 2019; Loesel et al., 2020). Alternative formats that rely on AI include live machine translation in unscripted performances presented in (Mirowski et al., 2020), comedy roasts with AI by company *ComedyBytes*<sup>2</sup> and custom LLMs for improv (Cho and May, 2020). Many joke generation tools have also been developed (Goes et al., 2023; Winters et al., 2018; Winters, 2021), including *Witscript* by Toplyn (2022).

In contrast to that work, in our company’s improvised comedy shows, improvisational actors co-create comedic narratives alongside AI-based image generators. These performances touch equally upon the computer science behind text-to-image generators and upon the socio-technical aspects of AI-generated art.

Our show responds to societal interrogations about AI. Since 2015, image generation algorithms have followed

seemingly exponential improvement from uncanny curiosity to highly detailed reproductions of reality. Generative Adversarial Networks (Goodfellow et al., 2020; Brock et al., 2018; Esser et al., 2021), publicly mostly known for artistic style transfer, gave rise to text-guided diffusion-based image generators (Dhariwal and Nichol, 2021; Rombach et al., 2022), trained on billions of captioned images (Schuhmann et al., 2022) that allow realistic image generation in seconds. We use these text-to-image generators as tools in a variety of co-creative interactions with actors on stage and audiences. Our format demonstrates a creative use of such tools to inspire actors on stage.

## Methods and Performances

The premise of our shows is that actors react to images shown on the screen backstage, and strive at justifying what is shown, irrespective of how absurd or incongruous it may be. Comedy derives from the wit of actors who adapt to any suggestion while keeping a coherent storytelling. Visual improvisation is similar to the improv format known as “Powerpoint Karaoke” or “Improvised TED Talks”, where slide decks are prepared in advance of the show (unbeknownst to the presenter) or during the performance using an automated slide generator (Winters and Mathewson, 2019).

## Progressive Image Generation Process as Part of the Improvised Performance

In the first version of our show, developed in the context of tele-immersive online performances (Branch et al., 2021) performed online (see Figure 1), the procedurally-generated images were rendered as dynamic “art” pieces projected onto a virtual painter’s easel, letting the improvisers and audiences see the image as it is being generated, or revealing it only at the end of the scene as a punchline.

We built a tool based on BigSleep (Murdock and Wang, 2021), employing an image generator combined with a Contrastive Language-Image Pretraining critic (Radford et al., 2021). Image generation was running locally on a desktop with a consumer-grade graphics card (NVIDIA RTX 3090), with results streamed to a media server interfacing with TouchDesigner<sup>3</sup> for projections. As the technology developed, we employed increasingly more capable image gen-

<sup>1</sup><https://improbobotics.org>

<sup>2</sup><https://www.comedybytes.io/>

<sup>3</sup><https://derivative.ca>



Figure 1: Illustration of the improv game “AI Portrait”, with a virtual easel superimposed with the generated image. BigGAN + CLIP was used image generation. The photo on the left shows a stage performance where the image of the painting is projected as appearing on an easel inside an artist’s studio, and shown on the screen above the actors. The two photos on the right illustrate an online performance with the image generated shown on a virtual easel that can be shown to or hidden from the public. Credits: Erika Hughes.

erators, starting with Big GAN (Brock et al., 2018), then VQGAN (Esser et al., 2021) and finally Diffusion Networks (Dhariwal and Nichol, 2021). These generators were unconditional, meaning they would not generate images in response to text; rather they were initialized randomly and had to be coupled with CLIP “critic” module that would match image and text embeddings via gradient descent, then back-propagate gradients from the text embeddings down to image embeddings and image generator, to gradually maximize the similarity between the generated image and target text caption.

As a consequence of the limitations of the image generation algorithm at that time, a single image generation typically lasted for the whole duration of the improv scene (which was about 5 minutes). We therefore needed to build an improv game format around the constraints of the algorithm, and opted for formats inspired by a painter’s studio.

We used the generated images to texture a 3D model of a painter’s easel within the virtual environment. The easel could be moved on and off the screen as well as flipped around so the audience can see either the back of the easel or the generated painting. This allowed us to develop several different games around “painting”. The basic form of the game involved a painter in a studio with an acquaintance wanting a portrait; the painter and the subject would engage in a dialogue that would reveal the inner truth of their mutual relationship, as the AI generated image would progressively be revealed<sup>4</sup> to both the painter and to audience (the sub-

<sup>4</sup>To stage the slow image generation process, the image would

ject would not look at or see the painting). As the image would be revealed, usually looking nothing like what is expected, the improvisers would need to justify it and react to it. The painter’s dialogue would also be informed by the image gradually appearing on the “easel”, and comedy resulted from information asymmetry (Fig. 1).

Many variations of this game can be played including having the painter be a street caricature artist that the couple randomly encounters, to a ‘new age’ art therapist who paints a person’s trauma while listening to their problems (Fig. 2).

As we introduced the format in early 2021, improvisers considered the output of image generators as “abstract art”. As their realism advanced, improvisers would consider them as caricatures and highlight their uncanny elements as source material for improvised comedy. Interestingly, such perceptions followed the embrace of the *glitch aesthetic* prevalent in early image generation<sup>5</sup>.

### Fast Image Generation and Surprise

With the advent of faster image generators based on Latent Diffusion Rombach et al. (2022), we updated both the software interface of our show (replacing image generators by Latent Diffusion then Stable Diffusion 1.5 models, running

be slowly revealed over time and rendered using “brush strokes”, as if it had been “painted”.

<sup>5</sup>Examples of AI-assisted artworks when we started developing the show can be seen on the AI Art Gallery of the NeurIPS Workshop on Machine Learning for Creativity and Design 2020 <https://www.aiartonline.com/>



Figure 2: Illustration of the improv game “art therapy”, showing the image generation process with “brush strokes”. The generated image is projected both on the screen and on a physical easel. VQGAN + CLIP was used for image generation. Credits: Stuart Hollis, Lidia Crisafulli.

either locally or via the Stable Diffusion API<sup>6</sup>) and the game format. We took advantage of much faster generation (of the order of 5-10 seconds), allowing the prompter to continuously listen to and react to the scene, generating new material “riffing off” the scene.

As we were able to generate multiple image candidates, we introduced a new role consisting in curating the images via a visual interface running on a tablet. Figure 4 shows the UI for prompting and curating the images; this UI allows to display a specific image on screen with appropriate comedic timing, as well as to avoid selecting and showing images that could be inappropriate for the show and audience.

With that new tool, we explored various improv game formats, and settled on “Movie Pitch”, where two directors come to meet a film producer to pitch them their film project. As the producer continuously asks them questions about the film, visuals appear on the screen. Each time a new image appears, the directors attempt to justify it as being an element from the film, such as the movie’s poster, events in the film or a celebrity actor cameo.

We also investigate generating scene backgrounds and illustrations that respond to the scene’s dialogue or creatively prompt the actors’ creativity. Note that these images were rendered to support live performance. As the quality of generated images improved, they became more predictable, and our challenge consisted in avoiding merely illustrating the improv scene. Instead, the image prompter took on the role of “yes, and-ing” the suggestions from the improvisers, to introduce new and surprising elements in generated images.

## Performances

Our visual improvisation game format demonstrates some of the most exciting new areas for live performance with co-creative AI, setting the stage for entirely new types of performance experience.

<sup>6</sup><https://platform.stability.ai/>

Our format has been presented fifty-five times since 2021 at various theatre and comedy festivals. The “AI Painter” has been presented at online performances (Art AI 2021, Queen City Improv 2021 and Iasi International Festival for Young Audiences 2021) and during staged shows (Camden Fringe 2021 and Leicester Comedy Festival 2022). Rapid image-generation-based shows were presented at Brighton Fringe 2022, Dataiku Everyday AI 2022, Sweden Improfest 2022, British Science Festival 2022 and Birmingham Improv Festival 2022. The “Movie Pitch” format has been presented at Update Required 2023, Brighton Fringe 2023, London AI Festival 2023, Greater Manchester Fringe 2023, twenty-six shows at Edinburgh Festival Fringe 2023, Bristol Improv Theatre in 2023 and Leicester Comedy Festival 2024. The show received positive audience reactions<sup>7</sup>.

## Discussion

### Staging the Image Generation Process

Showing image generation on screen through those different modalities allows us to intuitively communicate two aspects of the algorithm: the iterative generation from an initial noise image (by showing intermediary images produced by the diffusion algorithm or the CLIP optimization), and the stochastic nature of the algorithm (by showing multiple realisations for the same text prompt). This approaches allows to demystify the image generation process for theatre audiences, and in turn helps initiate a dialogue on the current concerns around generative AI.

### Co-creation with Image Generators

Our process for devising and performing alongside generated imagery with a human-in-the-loop ethos, can be situated with recent experimentations of image generation-driven co-creation: the online Discord performance of Bu-

<sup>7</sup><https://theatreandtonic.co.uk/blog/artificial-intelligence-improvisation-review>



Figure 3: Illustration of the improv game “Movie Pitch”. Stable Diffusion 1.5 was used for image generation. Left photo shows the image curator holding a tablet. Right photo shows the image prompter on a laptop. Credits: Lidia Crisafulli.

reau of Multiversal Arbitration<sup>8</sup> by company Aconite (where players engage in prompt battles), live AI performances with co-creation using text and image generators, such as PORTAGING (presented at the NeurIPS Workshop on Machine Learning for Creativity and Design 2022)<sup>9</sup> and Improved TEDx talks by Mathewson and Faïd (reacting to AI-generated presentation slides). This performative aspect extends research on co-creation and iterative design with image generators (Epstein et al., 2022).

An important aspect of improvising with live image generators is the bidirectional flow of information and inspiration, from generated images to live performance, and from live performance to subsequently generated images. If the flow of information was restricted to the former, the actors would only play a *justification* game, demonstrating their adaptability to (incongruous) new suggestions. If that flow of information consisted in only the latter, the image generation system would merely *illustrate* the improvised scene. To succeed in co-creation, the image generation system (or its operator/prompter and curator) needs to input new ideas in response to what the actors say. This can be achieved, for instance, by the reincorporation and combination of multiple previously mentioned ideas into prompts for new images, or by exploring themes adjacent to the currently played scene.

### Ethical Implications

As image generation technology developed, it became widely available to both visual artists and the general public. The fact that images can be generated in the style of specific visual artists gave rise to controversies and ethical concerns about plagiarism and misappropriation of artistic work that cannibalise creative economies (Frosio, 2023).

By employing image generators in the context of a show for diverse theatre festival audiences, we provoked and then engaged members of the general public attending our show



Figure 4: Screenshot of the image generator UI for the prompting and curation. The box in red shows the input box for the image prompt. The images show the results of previous generations. By clicking on one of the image boxes, the image projected on the screen for performers and public is updated.

about their perception of generative AI, illustrating possible uses of AI, inviting their scrutiny during and after the performance, and addressing concerns of the cast members and artists with whom we discussed about the show.

Specifically, we discussed the format and aim of the visual improvisation with our cast members, with members of the public to whom we flyered the show, with audiences in informal discussions before and after the show, with audiences during the performance through a qualitative survey<sup>10</sup>, with journalists from over 10 different press venues who interviewed us, including Daheley (2023); O’Sullivan (2023); Baum (2023); Richardson (2023); Greenberg (2023); Tett (2023), and with participants of a panel on art and AI during Edinburgh Festival Fringe 2023. Common concerns fo-

<sup>8</sup><https://www.bureauofmultiversalarbitration.com/>

<sup>9</sup><https://neurips.cc/virtual/2022/56220>

<sup>10</sup>Results from our survey, approved by the Ethics board of the University of Kent, are the object of a longer publication in (Branch et al., 2024)

cused on copyright and the misappropriation of artists' work when using image generation, and its destructive impact on the creative economies. Additional concerns included the appropriateness of generated images, their multiple representational biases, and the devaluation (via automation) of human creative work.

The problem of biases encoded in the image forced us to carefully consider 1) what images would be shown to the audience (thanks to the curation work done by one of the improvisers tasked with selecting the images), and 2) whether the image curator themselves needed a "Not Safe For Work" image detector to minimise the risk of them being exposed to harmful images. As related by performers, over-sexualised images were the most common problematic occurrence, followed by a lack of diversity in representations of people. We partially addressed these issues by proactively avoiding prompts that we suspected could result in sexualised images and also actively aiming at diversifying the prompts and hence the appearance of persons on the generated images.

The problem of attributing generated images to specific artists (who may have unknowingly contributed images to training image generators) is an active area of research and legislation (Zhong et al., 2023). In the context of a live and improvised show, real-time image attribution becomes a highly difficult task. As the performances are typically not recorded and as the generated images are not distributed after the show, the use of generative AI can be put in perspective with best practices for the use of excerpts of printed books or music tracks in some other improvised shows. On the other hand, we can imagine an alternative approach to employing generative AI in a show, by relying on models trained only on licensed work and paid artists.

## Conclusion

In our show format, we presented alternative collaborative and co-creative applications of generative art that invite human performers directly into the generation loop, curating and responding to outputs from image generation systems as part of live interaction. As the audience could witness, generated images were not the final artistic output, and were not distributed post-performance: instead, they were impermanent, like improvised theatre. Generated images went through a layer of transformation by serving as source material to inspire live, improvised performance of human actors.

## Acknowledgements

We thank the audiences of our shows for critical vocal feedback, and all the cast members of Improbotics who contributed their talents and ideas to the improvised game formats with image generation, namely Alex Newson, Ben Lovell, Em Stroud, Fiona Howat, Harry Turnbull, Holly Mallett, Jenny Elfving, Jenny Rowe, Jillian Ellis, Jodie Irvine, Julie Flower, Marouen Mraïhi, Mike Prior, Paul Little, Roel Fox, Sarah Davies, Thomas Jones and Tommy Rydling.

## References

- Baum, G. (2023). Why ai isn't funny: At least not yet. <https://www.hollywoodreporter.com/business/digital/why-ai-isnt-funny-at-least-not-yet-1235503678/>. Hollywood Reporter.
- Branch, B., Efstratiou, C., Mirowski, P., Mathewson, K. W., and Allain, P. (2021). Tele-immersive improv: Effects of immersive visualisations on rehearsing and performing theatre online. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–13.
- Branch, B., Mirowski, P., Mathewson, K., Ppali, S., and Covaci, A. (2024). Designing and evaluating dialogue llms for co-creative improvised theatre. In *Proceedings of the 15th International Conference on Computational Creativity*. Association for Computational Creativity.
- Brock, A., Donahue, J., and Simonyan, K. (2018). Large scale gan training for high fidelity natural image synthesis. *arXiv preprint arXiv:1809.11096*.
- Bruce, A., Knight, J., Listopad, S., Magerko, B., and Nourbakhsh, I. R. (2000). Robot improv: Using drama to create believable agents. In *IEEE International Conference on Robotics and Automation*, volume 4, pages 4002–4008. IEEE.
- Cho, H. and May, J. (2020). Grounding conversations with improvised dialogues. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 2398–2413.
- Daheley, T. (2023). What the ai revolution means for arts. <https://www.bbc.co.uk/programmes/w3ct37sv>. BBC World Service - Cultural Frontline.
- Dhariwal, P. and Nichol, A. (2021). Diffusion models beat gans on image synthesis. *Advances in Neural Information Processing Systems*, 34:8780–8794.
- Epstein, Z., Schroeder, H., and Newman, D. (2022). When happy accidents spark creativity: Bringing collaborative speculation to life with generative ai. In *Proceedings of the 11th International Conference on Computational Creativity*. Association for Computational Creativity.
- Esser, P., Rombach, R., and Ommer, B. (2021). Taming transformers for high-resolution image synthesis. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 12873–12883.
- Frosio, G. (2023). Generative ai in court. *Court (September 1, 2023)*. in Nikos Koutras and Niloufer Selvadurai (eds), *Recreating Creativity, Reinventing Inventiveness-International Perspectives on AI and IP Governance (Routledge, 2023, Forthcoming)*.
- Goes, F., Sawicki, P., Grześ, M., Brown, D., and Volpe, M. (2023). Is gpt-4 good enough to evaluate jokes? In *Intl Conf Computational Creativity*.
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. (2020). Generative adversarial networks. *Communications of the ACM*, 63(11):139–144.

- Greenberg, E. (2023). Yes-anding ai: Artificial intelligence stars at the edinburgh fringe. <https://www.digit.fyi/yes-anding-ai-artificial-intelligence-stars-at-the-edinburgh-fringe/>. DIGIT News.
- Loesel, G., Mirowski, P., and Mathewson, K. W. (2020). Do digital agents do dada? In *Proceedings of the 11th International Conference on Computational Creativity*. Association for Computational Creativity.
- Mathewson, K. and Mirowski, P. (2018). Improbatics: Exploring the imitation game using machine intelligence in improvised theatre. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, volume 14, pages 59–66.
- Mathewson, K. W. and Mirowski, P. (2017a). Improvised comedy as a turing test. *arXiv preprint arXiv:1711.08819*.
- Mathewson, K. W. and Mirowski, P. (2017b). Improvised theatre alongside artificial intelligences. In *Thirteenth Artificial Intelligence and Interactive Digital Entertainment Conference*.
- Mirowski, P., Mathewson, K., Branch, B., Winters, T., Verhoeven, B., and Elfving, J. (2020). Rosetta code: Improv in any language. In *Proceedings of the 11th International Conference on Computational Creativity*, pages 115–122. Association for Computational Creativity.
- Mirowski, P. and Mathewson, K. W. (2019). Human improvised theatre augmented with artificial intelligence. In *Proceedings of the 2019 on Creativity and Cognition*, pages 527–530.
- Murdock, R. and Wang, P. (2021). Big sleep. <https://github.com/lucidrains/big-sleep>.
- O’Sullivan, M. (2023). Artificial intelligence can create, but lacks creativity, say critics. <https://www.voanews.com/a/artificial-intelligence-can-create-but-lacks-creativity-say-critics/7068177.html>. Voice of America.
- Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., et al. (2021). Learning transferable visual models from natural language supervision. In *International Conference on Machine Learning*, pages 8748–8763. PMLR.
- Richardson, J. (2023). Ai is taking over fringe comedy; can robots be funnier than humans? <https://www.scotsman.com/arts-and-culture/edinburgh-festivals/ai-is-taking-over-fringe-comedy-can-robots-be-funnier-than-humans-4238383>. The Scotsman.
- Rombach, R., Blattmann, A., Lorenz, D., Esser, P., and Ommer, B. (2022). High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 10684–10695.
- Schuhmann, C., Beaumont, R., Vencu, R., Gordon, C., Wightman, R., Cherti, M., Coombes, T., Katta, A., Mullis, C., Wortsman, M., et al. (2022). Laion-5b: An open large-scale dataset for training next generation image-text models. *Advances in Neural Information Processing Systems*, 35:25278–25294.
- Tett, G. (2023). Can ai crack comedy? <https://www.ft.com/content/818f2cab-57ff-42c3-917b-4a83f1d87802>. Financial Times.
- Toplyn, J. (2022). Witscript 2: A system for generating improvised jokes without wordplay. In *Intl Conf Computational Creativity*.
- Winters, T. (2021). Computers learning humor is no joke. *Harvard Data Science Review*, 3(2).
- Winters, T. and Mathewson, K. W. (2019). Automatically generating engaging presentation slide decks. In *International Conference on Computational Intelligence in Music, Sound, Art and Design (Part of EvoStar)*, pages 127–141. Springer.
- Winters, T., Nys, V., and De Schreye, D. (2018). Automatic joke generation: Learning humor from examples. In *Distributed, Ambient and Pervasive Interactions: Technologies and Contexts: 6th International Conference, DAPI 2018, Held as Part of HCI International 2018, Las Vegas, NV, USA, July 15–20, 2018, Proceedings, Part II 6*, pages 360–377. Springer.
- Zhong, H., Chang, J., Yang, Z., Wu, T., Mahawaga Arachchige, P. C., Pathmabandu, C., and Xue, M. (2023). Copyright protection and accountability of generative ai: Attack, watermarking and attribution. In *Companion Proceedings of the ACM Web Conference 2023*, pages 94–98.