

AI Co-Creativity for Musicians: Making Audiovisual Work through Human in the Loop Workflows

Dr Kirsten Hermes

DCDI – School of Arts – Music

k.hermes@westminster.ac.uk

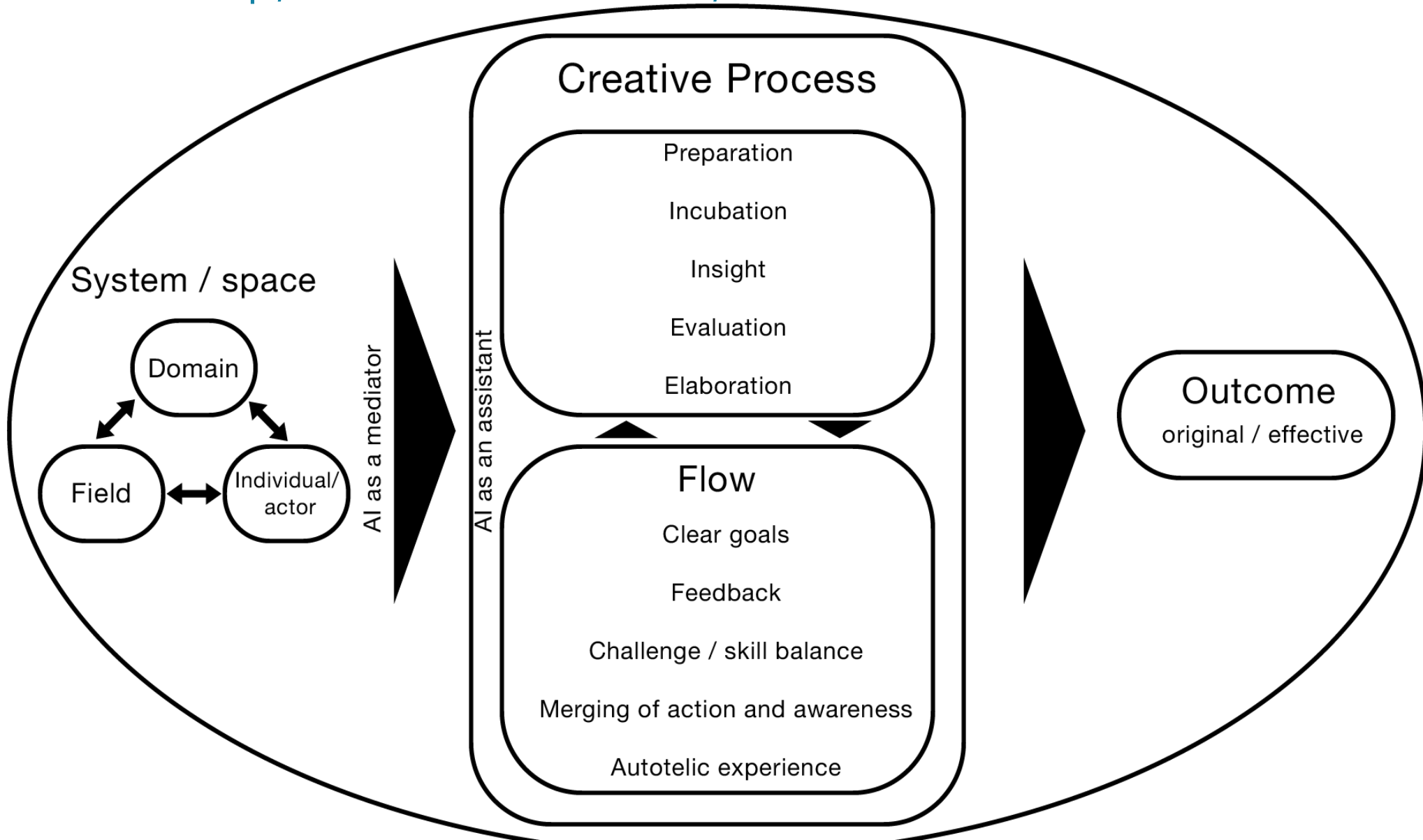
Context and Research Questions

Musicians increasingly require audiovisual materials. Creativity is best understood as a system of the individual, the domain, and the field, not a lone-genius act (Kerrigan, 2014; Csikszentmihalyi, M., 1997). New AI tools can therefore slot into an already networked practice.

Creatives worry about AI replacing their jobs, amid redundancies across media industries. This intensifies pressure in an already competitive work environment. Shneiderman (2022) argues for human-centred AI, i.e. systems that assist rather than replace humans.

RQ: In the context of making visual materials...

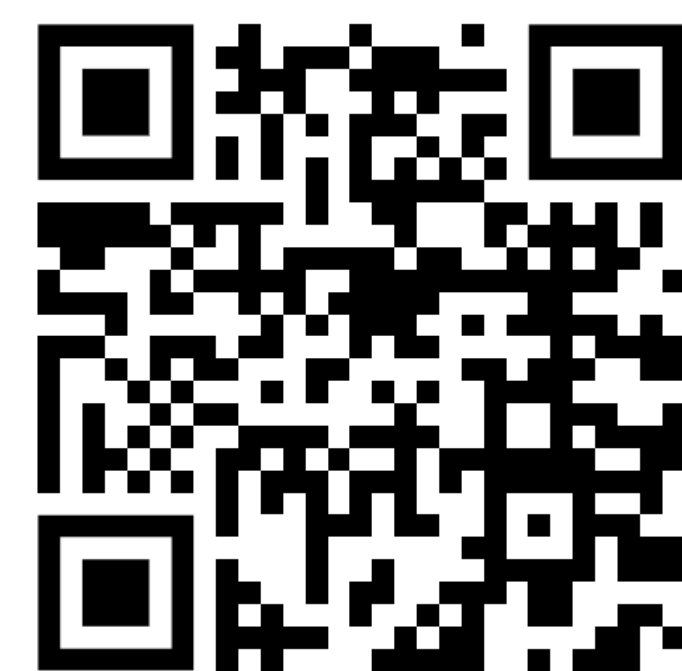
- Where can AI help musicians in the creative process/flow?
- Where must the musician stay in charge to maintain authorship, creative coherence, and ethics?



Use Case for Musicians	Relevant AI tools
Generate new audio files (+ customise)	Suno, Udio, Jukebox, Splash, Mubert, ElevenLabs, MusicLM, AudioCraft, StableAudio, Muzik, SoundRaw, Boomy, Loudly, MusicFlow, AIVA
Generate MIDI	Audiocipher, Orb Producer, Captain Chords, Magenta
Generate melodies	Evabeat, Melody Sauce
Generate a backing track for a recording	Songzap
Music production	LifeScore, Aimi
Editable multi-part arrangements	Soundful, Bandlab, Logic Pro, Instacomposer
Mixing, effects	Izotope, AmpliTube, Sonible
Mastering	Landr
Customized audio experiences	Endel
Unmixing	LALAL.AI, Splitter.ai
More narrow audio applications	Pitch tracking or downbeat tracking (Heydary et al., 2023), Playing an accompaniment based on narrow set of parameters, e.g. key, input score, a given chord set (Xia, 2016), etc.
Match existing audio from different sources	Loopcloud
Timbre transfer	Quosmo, Audimee, SoundIDVoiceAI, RipX, Vochlea
Text	Chat GPT, GPT4, Gemini, Llama, Claude
Images (e.g. album artwork, social media posts)	Dall-E, Google Imagen, Midjourney, Stable Diffusion
Video (e.g. music video sections)	Runway, Sora, Meta "Make a Video", Stable Video
3D Graphics	Nvidia AI, DreamFusion, Anything.world, Vast Triplo AI
Motion Capture	Deepmotion, Plask,
Entire creative pipeline	Comfy UI (modular system)

Practice-led Research Project

Nyokee — Three Witches (music video):



A music/3D/animation pipeline involving a range of AI tools.

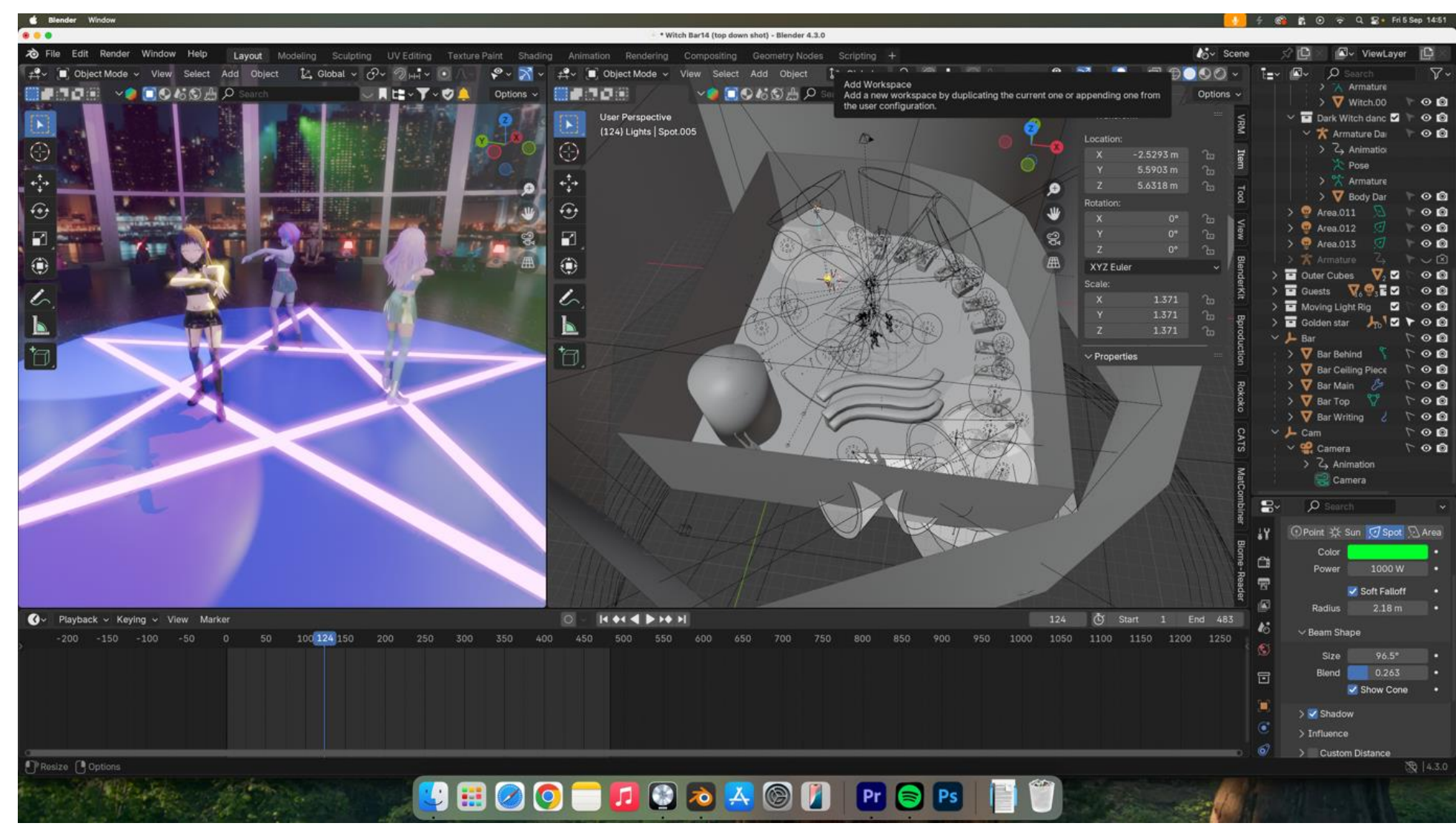
Practice-led workflow analysis + reflexive notes from making the video; mapping creative challenges and affordances to Creative Process and Flow parameters.

Methodology

Musical ideation in Suno, followed by music production work in Logic Pro.



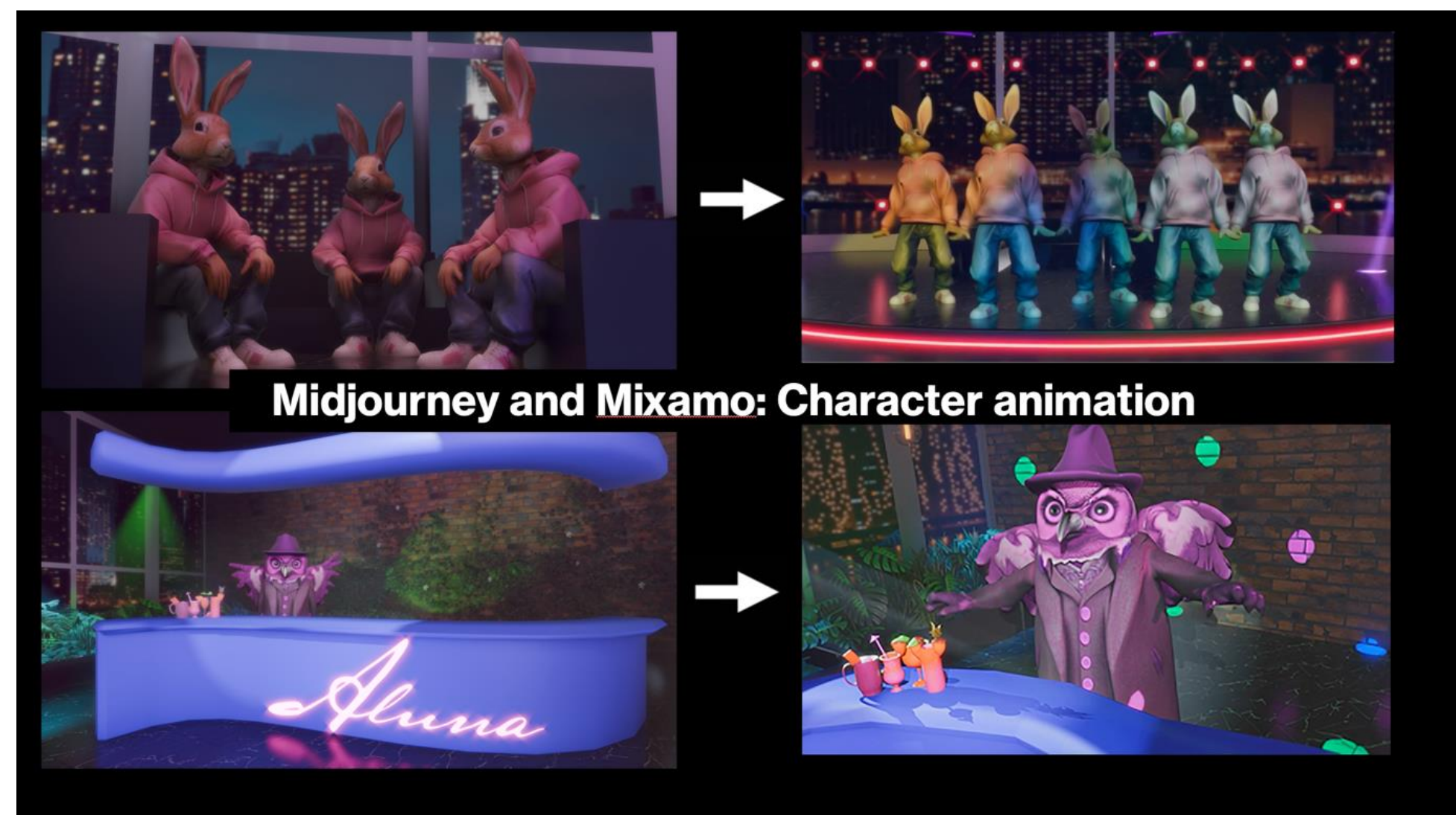
Custom scene in Blender with characters created in VRoid:



AI materials are introduced into the scene:



Hybrid Midjourney / Blender / Mixamo animation workflow:



Final video edit in Premiere Pro.



Findings

Flow Parameters

- Clear goals: set by the musician. AI can suggest options, not intentions.
- Feedback: AI enables instant iterations; community chat/analytics provide social feedback.
- Challenge–skill balance: AI lowers entry barriers; the musician selects difficulty and develops tool competence.
- Merging action and awareness: disparate tools create friction; cohesive pipelines required.
- Autotelic experience: rapid loops support playful exploration; but authorship must sustain intrinsic motivation.

Preparation

- AI helps: rapid ideation (text-to-image/music drafts), mood boards, asset search.
- Human leads: set concept, ethical boundaries, and IP.

Incubation

- AI helps: low-friction variant generation keeps ideas active without heavy investment.
- Human leads: select promising directions; align with artist identity and story world.

Insight

- AI helps: surprising prompts/visuals can catalyse novel combinations.
- Human leads: decide which serendipities serve the narrative and brand.

Evaluation

- AI helps: swift A/B tests of shots, lighting, character design; quick music stems for comparison.
- Human leads: uphold quality thresholds, taste, cultural sensitivity, and audience fit; draw on community feedback.

Elaboration (production)

- AI helps: character/asset generation, motion capture/retargeting, background plates, animation.
- Human leads: final edit, timing, sound mix, and credits.

Conclusion

AI can aid artists with ideation, skill development, immersion, and innovation, allowing them to view their work from new perspectives. Generative AI tools can help artists engage with cultural data in novel ways, offer clear references in collaborative settings, reduce repetitive tasks, and promote creative flow. In this way, musicians can use AI to enhance efficiency in the creative process. On the other hand, AI tools are often specialized and fragmented, requiring new skills and specific workflows. Current platforms rely on written prompts, which may not suit all users, and interfaces can feel rigid, limiting fine-grained control. Since AI cannot replicate human thought, reasoning, or lived experience, artists must interpret and create meaning from AI-generated outcomes. Many creatives desire control and ownership over their work, which AI tools do not always provide. Overall, while AI has the potential to enhance creativity, but further development is needed for a smoother integration into the creative process, depending on the artist's goals and the project's context.

References

- Csikszentmihalyi, M. (2013). Flow: The Psychology of Happiness. Germany: Penguin Random House.
- Csikszentmihalyi, M. (1997). Creativity: Flow and the psychology of discovery and invention. NY: HarperPerennial.
- Hermes, K. 2025 'Augmented creativity: using AI collaborators to make a music video', Dancecult Journal. (in press)
- Hermes, K., 2025. Augmented creativity: does AI help or hinder musicians' creative processes?. In: Hepworth-Sawyer, R., Hodgson, J., Paterson, J., Toulson, R. (Eds.) "Innovation in Music". Routledge, New York (accepted for publication).
- Kerrigan, S. (2014). Accommodating creative documentary practice within a revised systems model of creativity. Journal of Media Practice. 14, 111-127.
- Shneiderman, B. (2022), Human-Centered AI. Oxford: Oxford University Press.