

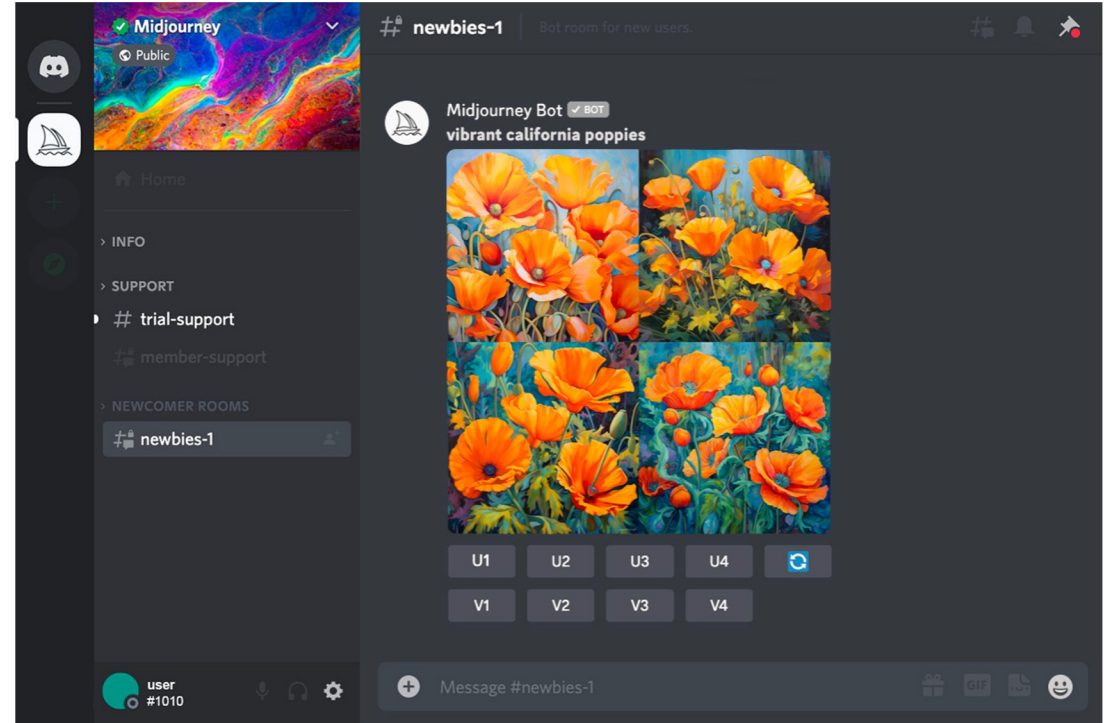
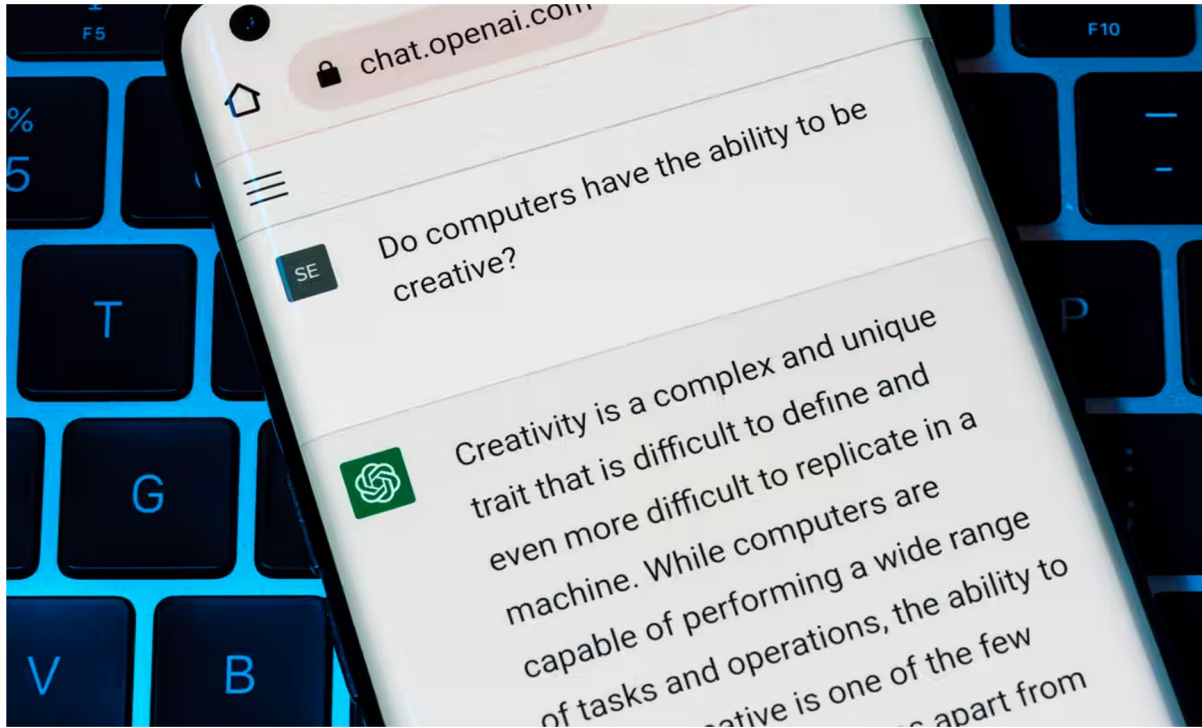


SEED // SEARCH FOR EXTRAORDINARY EXPERIENCES DIVISION
seed.ea.com

Beyond text and images: Generative AI in the game industry

Judith Bütepage
SEED – Electronic Arts

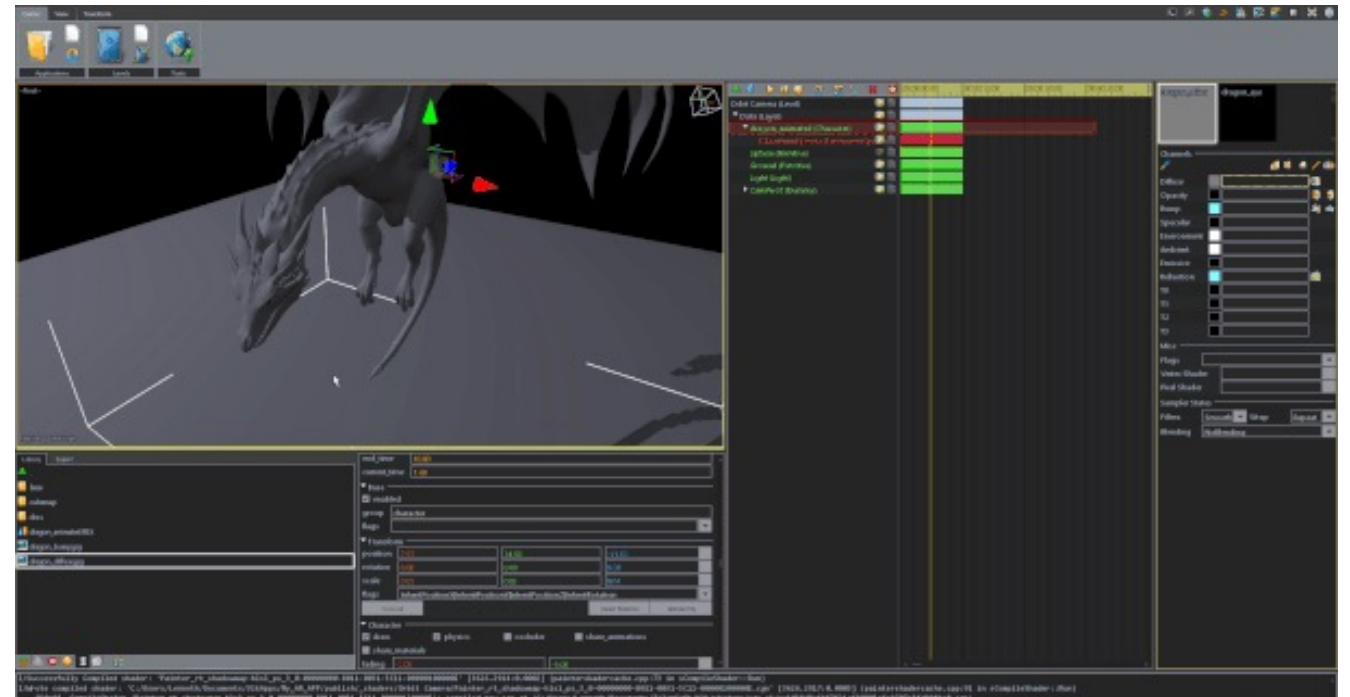
GENERATIVE AI



Link to image

CONTENT IN GAMES

3D Models, Textures, Animations, Character Designs, Environment Designs, Concept Art, Level Layouts, User Interface (UI) Elements, Particle Effects, Sound Effects, Cutscene Animations, Cinematics, Visual Effects, Lighting Designs, Props, Weapons, Vehicles, Rigging, Physics Simulations, Game Scripts, Level of Detail (LOD) Models, Terrain Sculpting, Skyboxes, Skydomes, Water Effects, Day-Night Cycle Systems, Weather Effects, Menu Screens, Loading Screens, Game Trailers, Game Packaging Designs, etc.



Links to image <https://www.cgchannel.com/2018/11/valkyrie-engine-aims-to-let-artists-create-ar-apps-without-coding/>

GENERATIVE AI IS USEFUL IN SITUATIONS WHERE ...

... tasks are repetitive and / or resource intensive.

... the need for development is high.

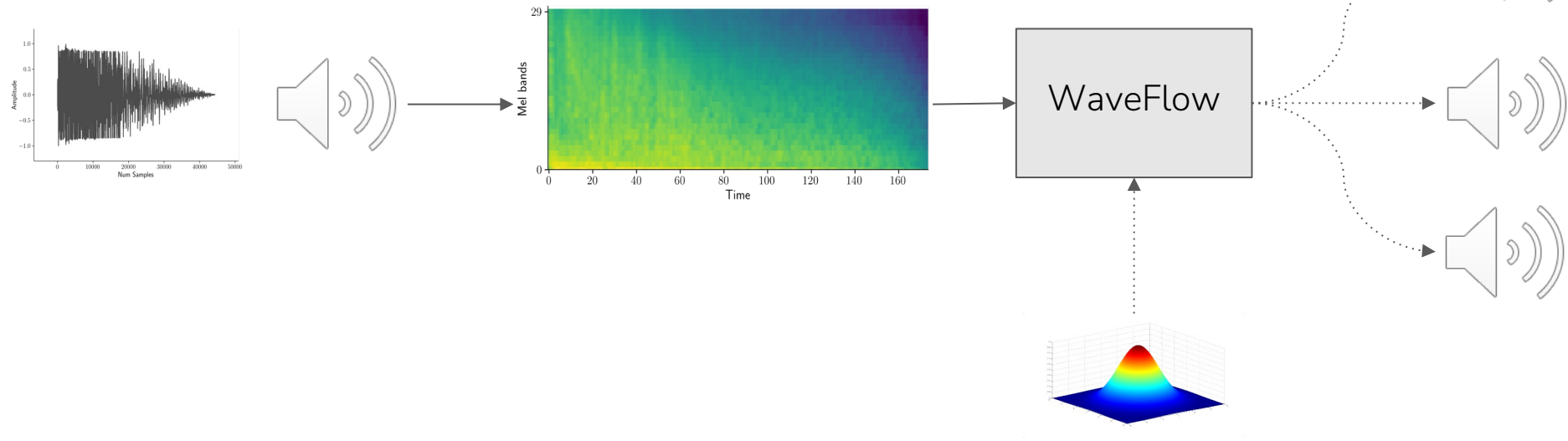
... the ML models can expedite development process

EXAMPLE: GENERATING SOUNDS

NEURAL SYNTHESIS OF SOUND EFFECTS

SYNTHESIZED

ORIGINAL

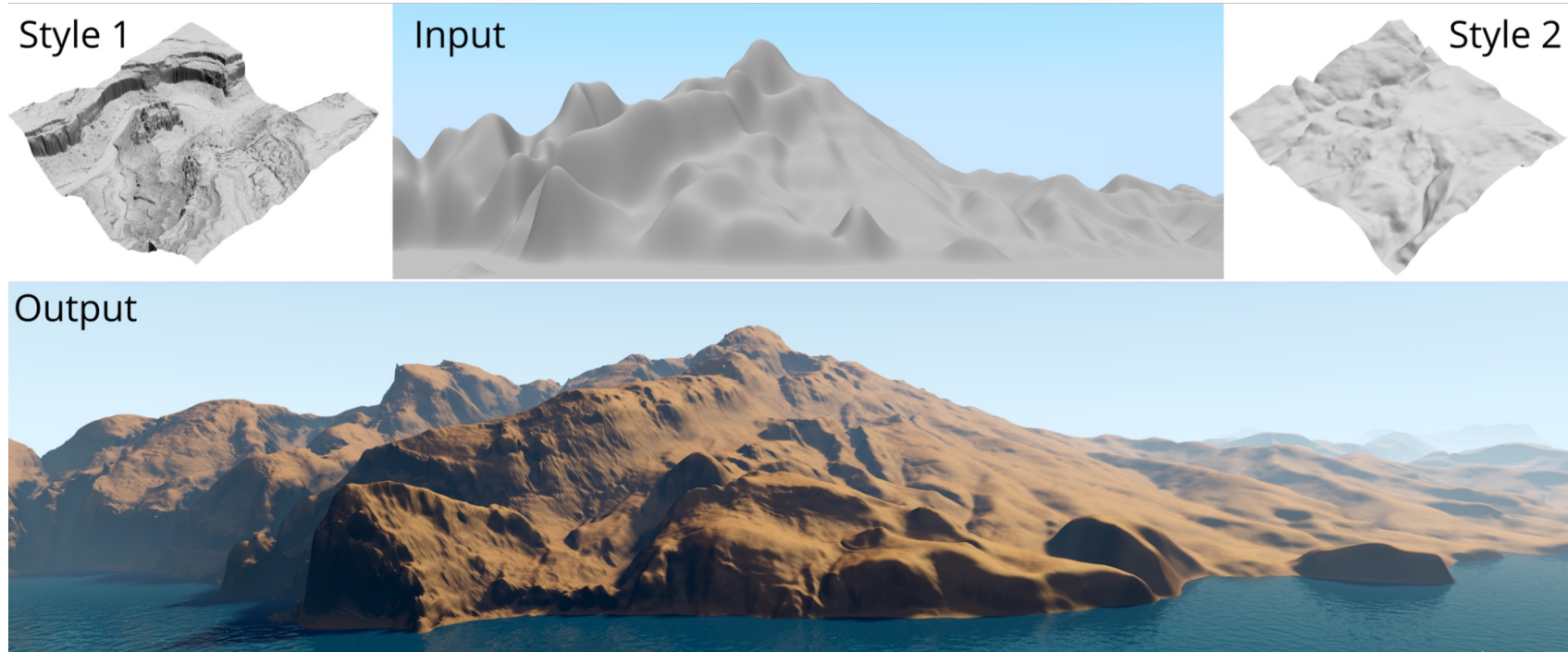


Andreu, Sergi, and Monica Villanueva Aylagas. "Neural Synthesis of Sound Effects Using Flow-Based Deep Generative Models." *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*. Vol. 18. No. 1. 2022.

Ping, Wei, et al. "WaveFlow: A compact flow-based model for raw audio." *International Conference on Machine Learning*. PMLR, 2020.

EXAMPLE: CREATING WORLDS

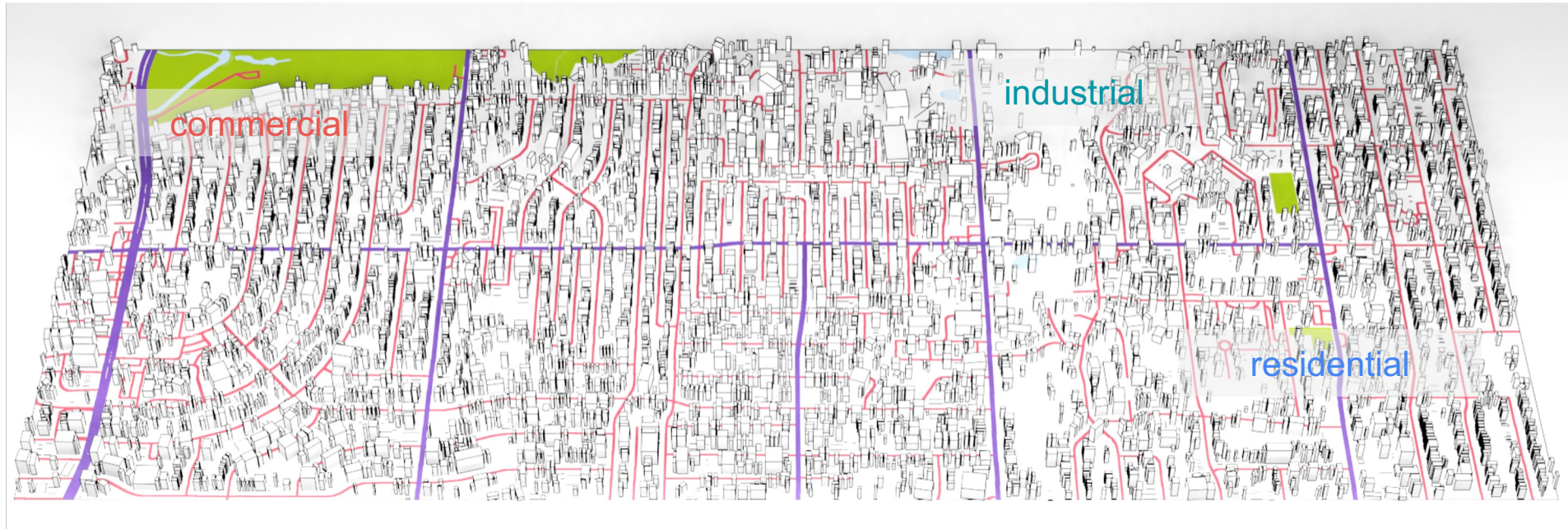
Generative Terrain Amplification



Building Footprint Generation



Building Footprint Generation



EXAMPLE: ANIMATION

EXAMPLE: ANIMATION



Expert Activation



Music-to-Dance Synthesis

Our method can further generalize to different genres of unheard music clips and adapt the choreographies to the beat of the music.

Phase State



Rise Against - Dancing For Rain

SIMILARITIES AND DIFFERENCES

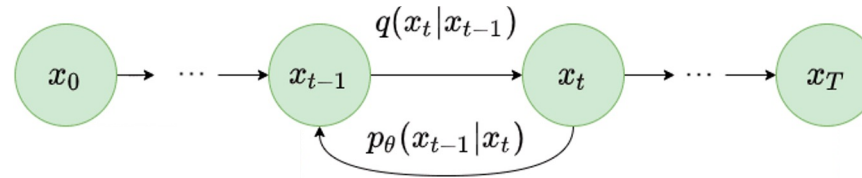
MAINSTREAM ML

VS

ML FOR GAME
DEVELOPMENT

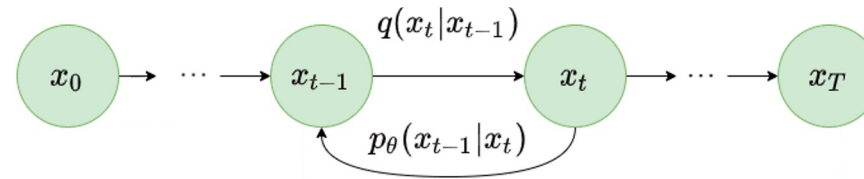
SIMILARITIES AND DIFFERENCES

New generative models

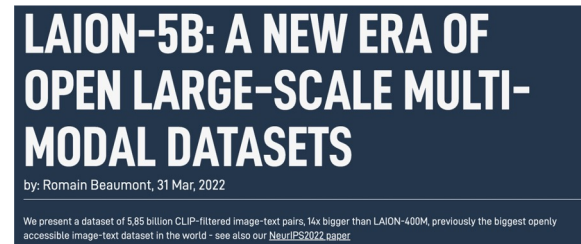


SIMILARITIES AND DIFFERENCES

New generative models

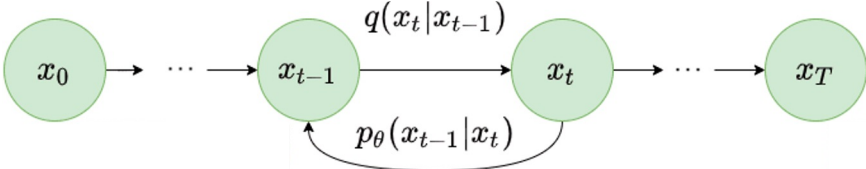


Lots of data



SIMILARITIES AND DIFFERENCES

New generative models

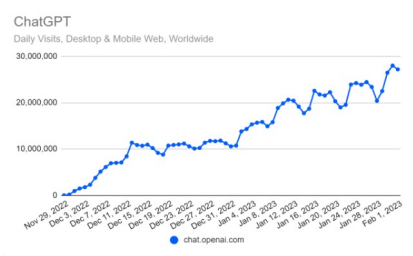


Lots of data

LAION-5B: A NEW ERA OF OPEN LARGE-SCALE MULTI-MODAL DATASETS
by: Romain Beaumont, 31 Mar, 2022

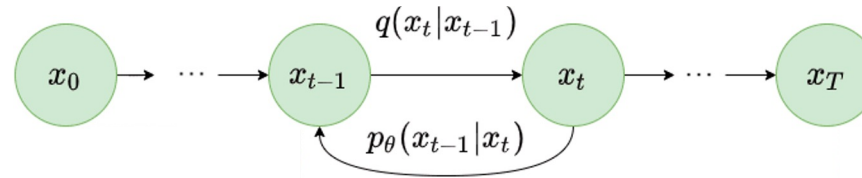
We present a dataset of 5.85 billion CLIP-filtered image-text pairs, 14x bigger than LAION-400M, previously the biggest openly accessible image-text dataset in the world - see also our [NeurIPS2022 paper](#)

Models publicly available

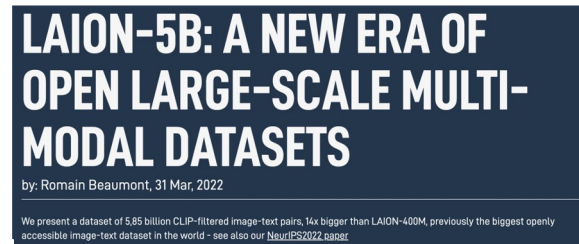


SIMILARITIES AND DIFFERENCES

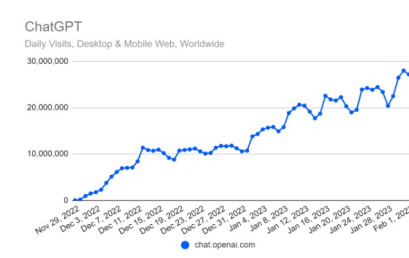
New generative models



Lots of data

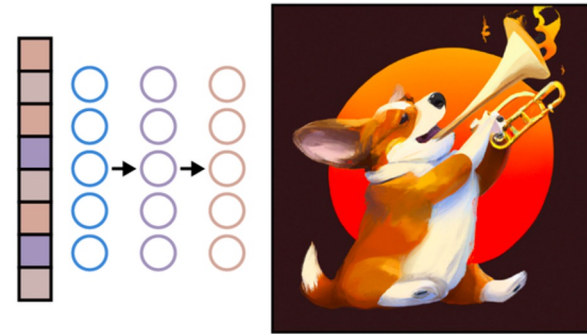


Models publicly available



Intuitive ways of interacting with models
and interpreting their output

THE POWER OF LANGUAGE

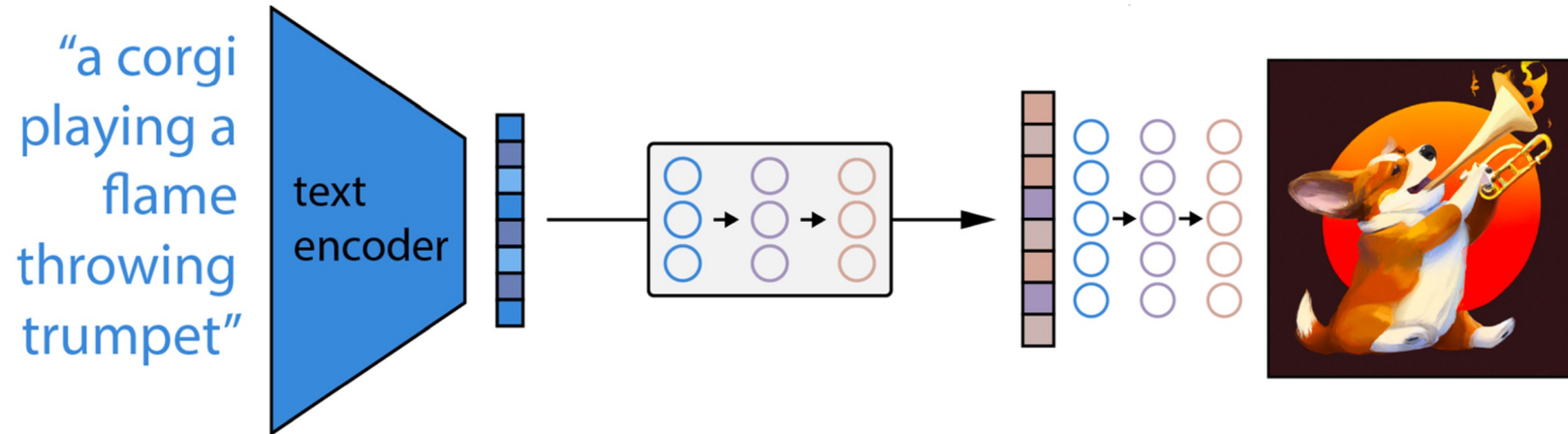


Most generative models generate content
by conditioning the output on a sample in a latent space.

THE POWER OF LANGUAGE

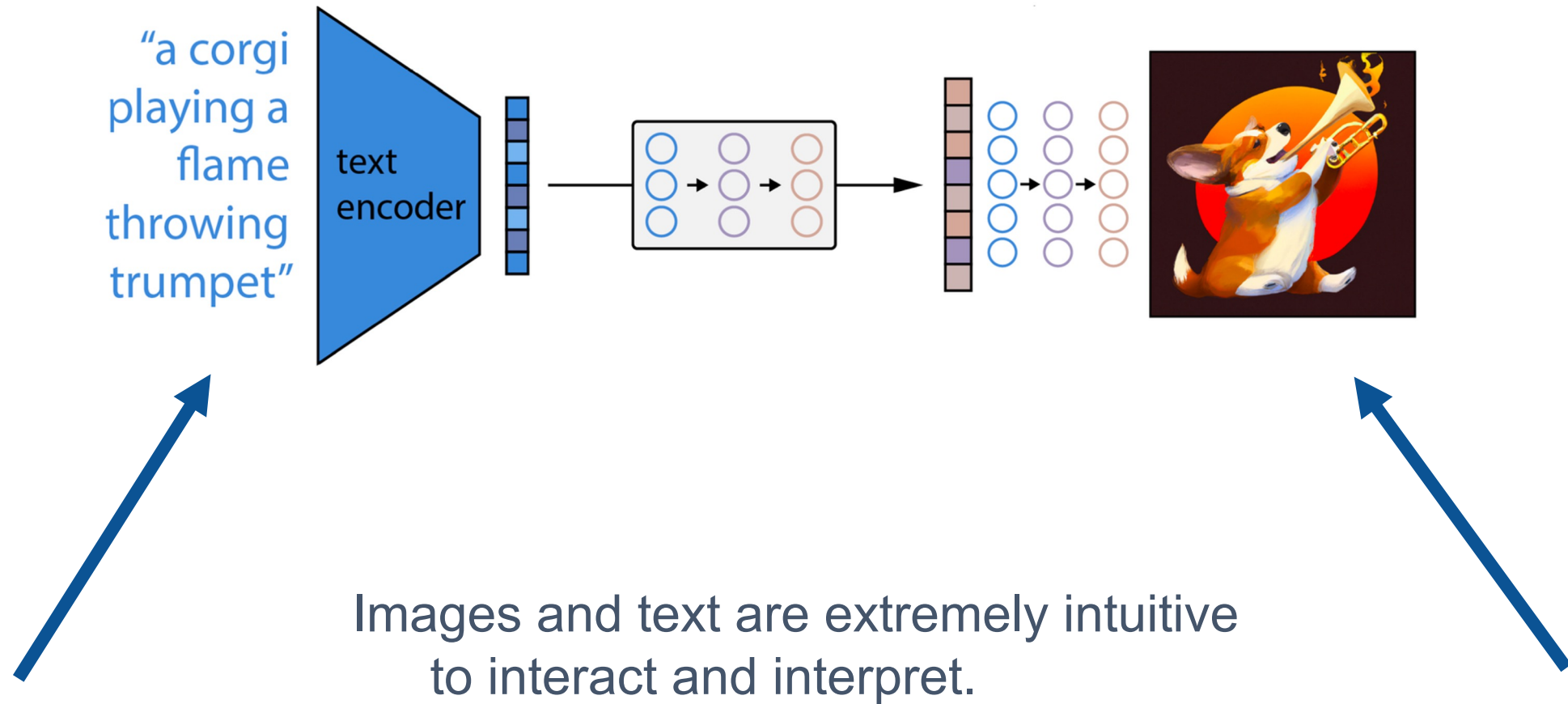
Disclaimer:

This is roughly how DALL-E works.
Diffusion models work a bit differently.
The intuition still holds.



Now you can drive the sampling with
text prompts.

THE POWER OF LANGUAGE

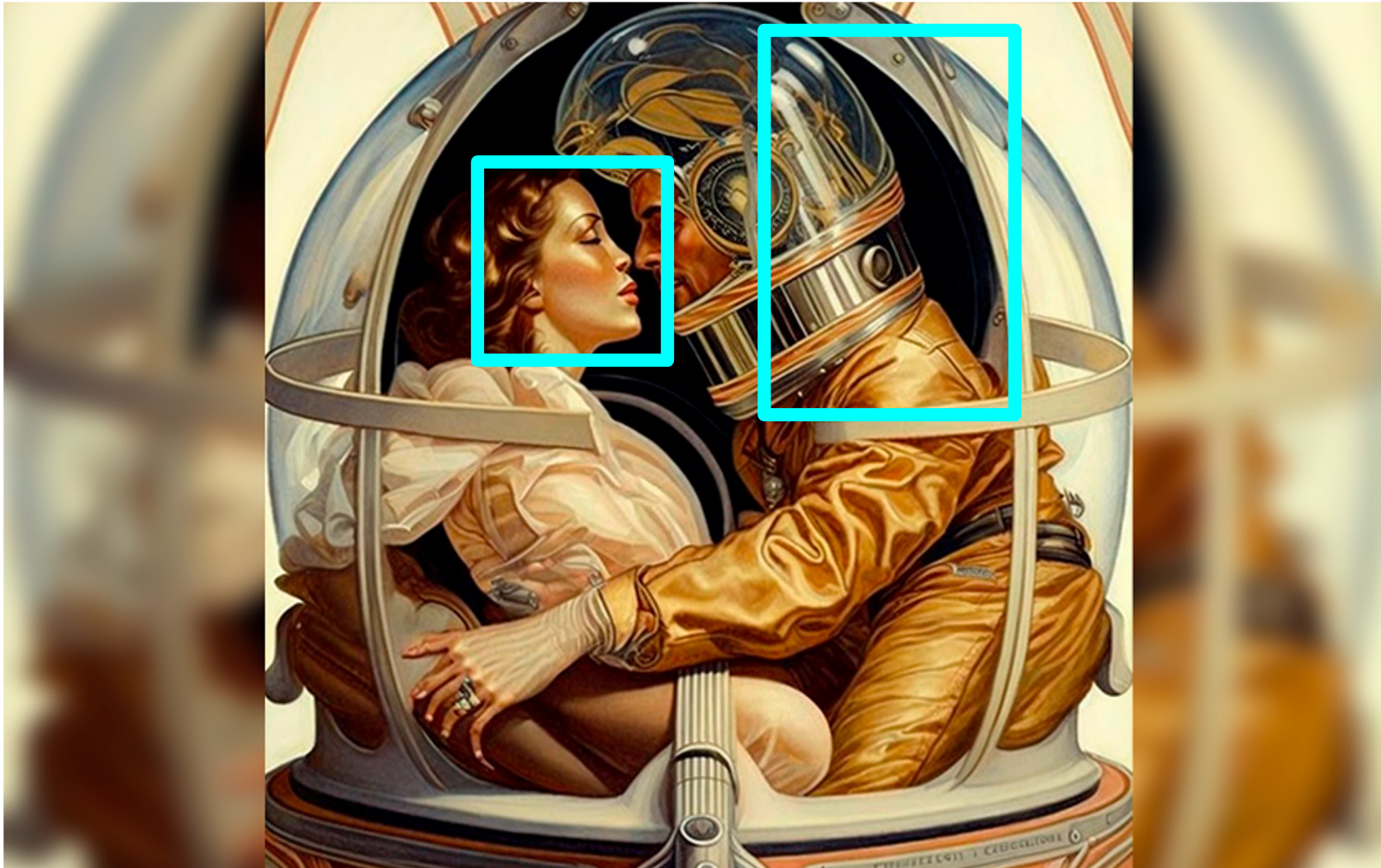


THE POWER OF IMAGES



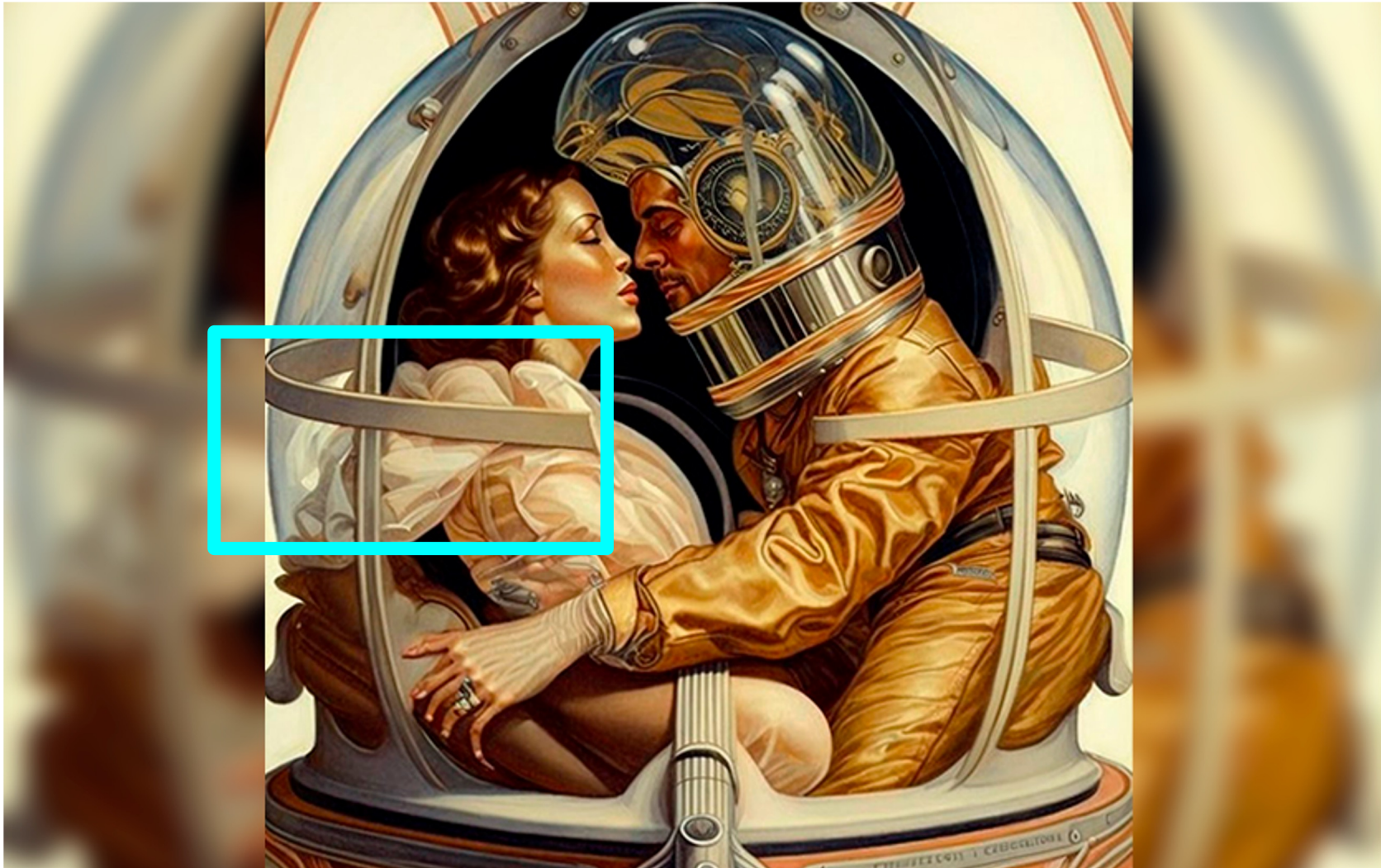
Link to image <https://gagadget.com/en/how-it-works/210102-the-unattainable-pinnacle-of-art-why-does-midjourney-artificial-intelligence-draw-6-fingers-on-hands-and-how/>

THE POWER OF IMAGES

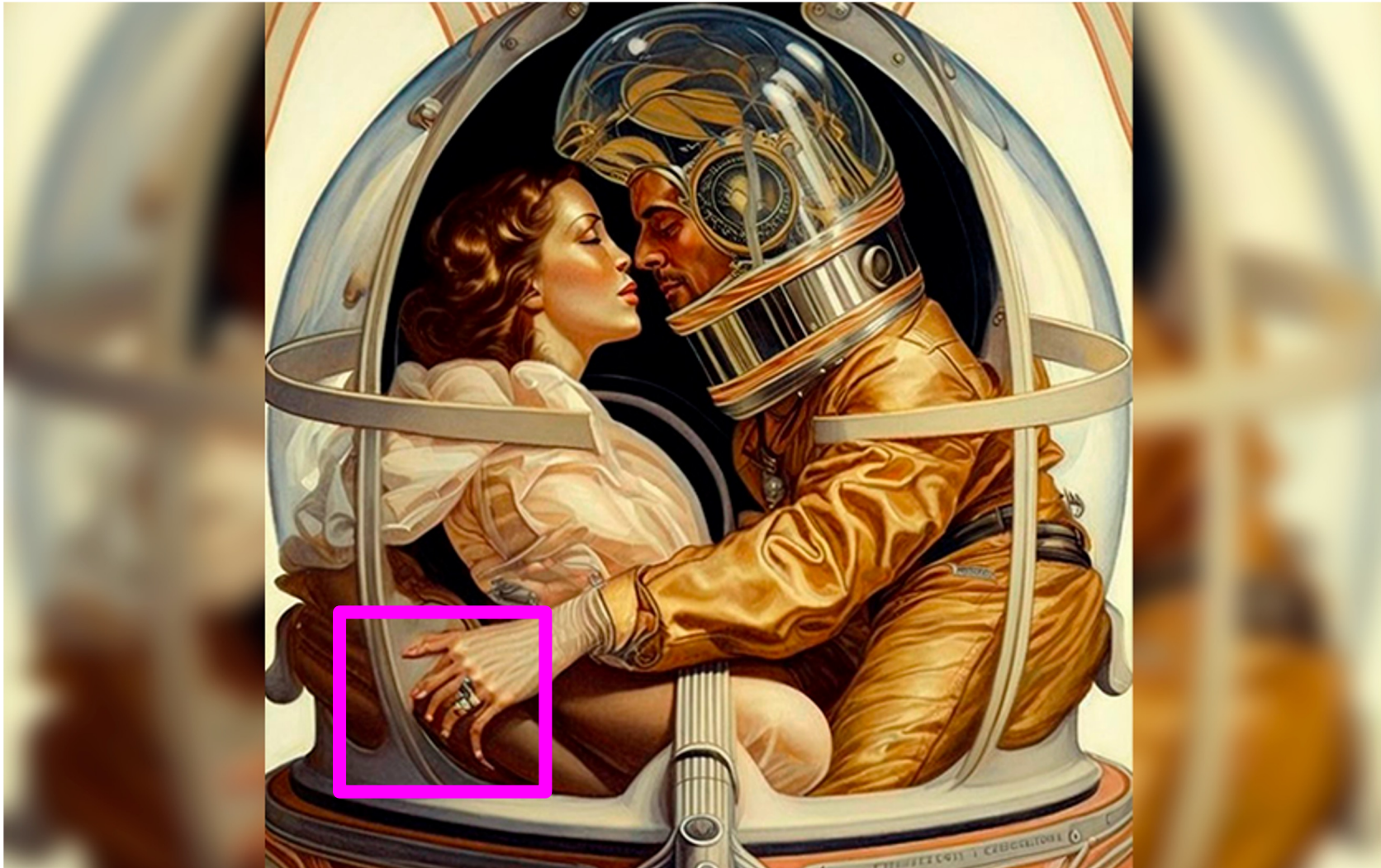


Link to image <https://gagadget.com/en/how-it-works/210102-the-unattainable-pinnacle-of-art-why-does-midjourney-artificial-intelligence-draw-6-fingers-on-hands-and-how/>

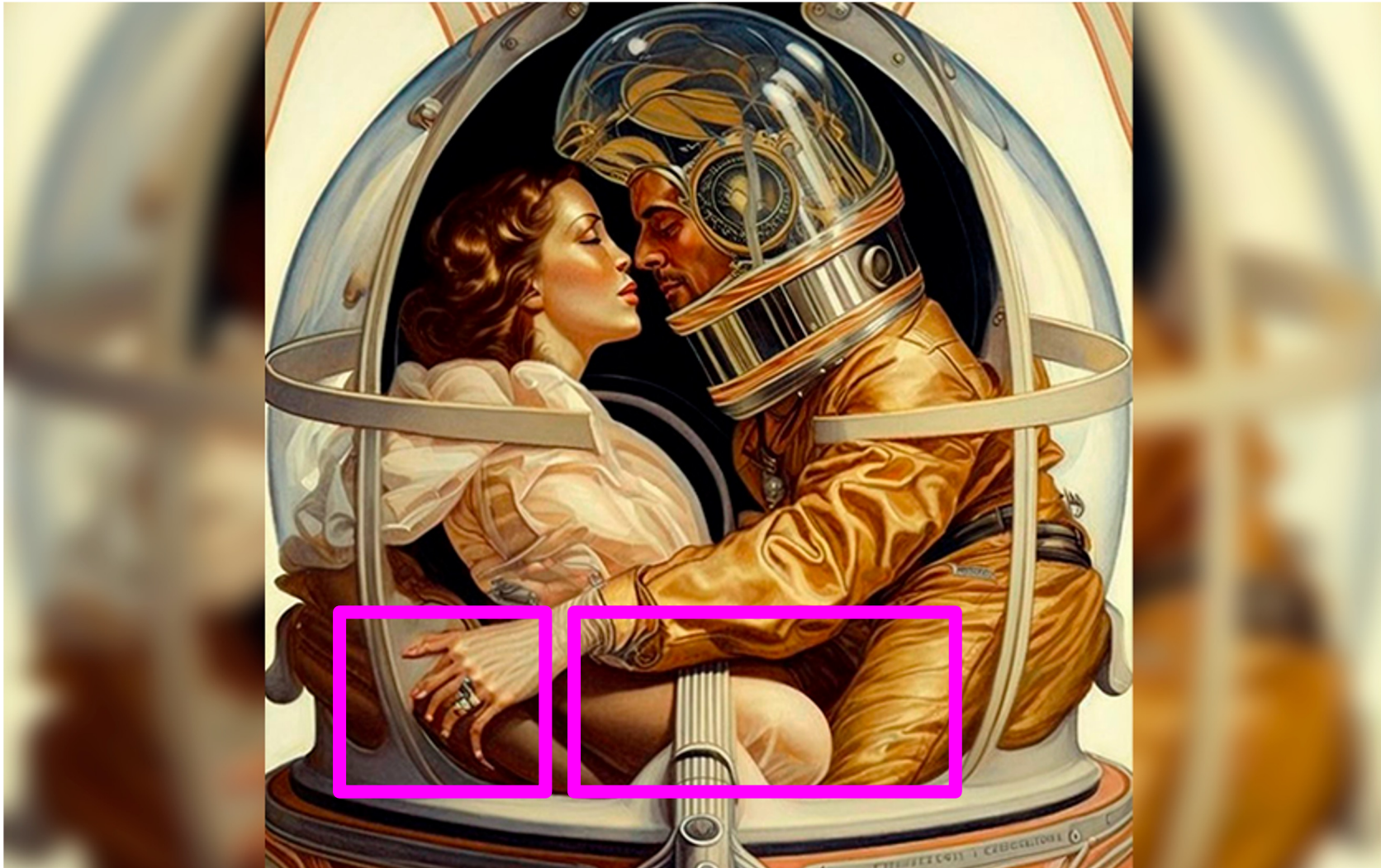
THE POWER OF IMAGES



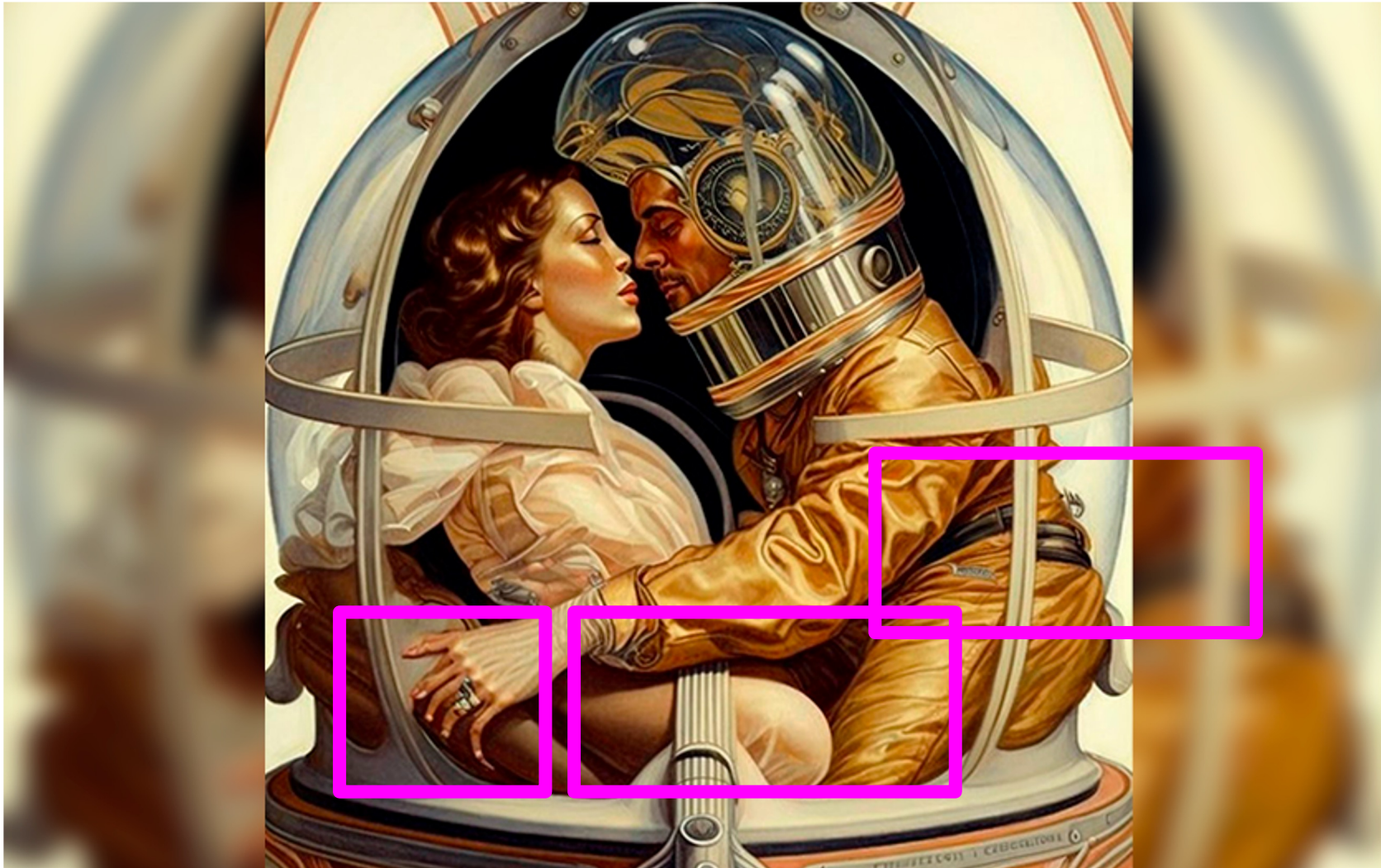
THE POWER OF IMAGES



THE POWER OF IMAGES



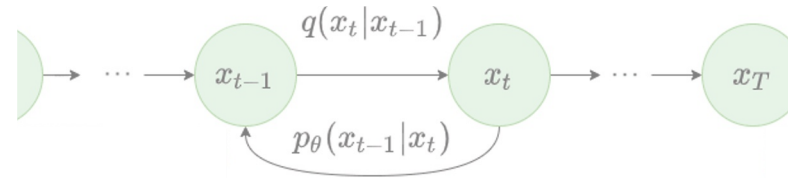
THE POWER OF IMAGES



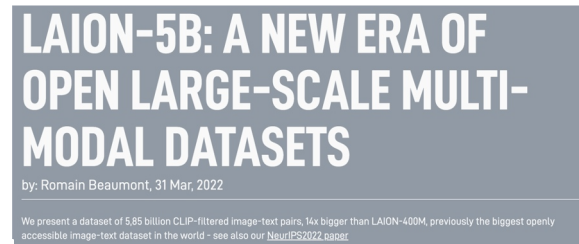
Link to image <https://gagadget.com/en/how-it-works/210102-the-unattainable-pinnacle-of-art-why-does-midjourney-artificial-intelligence-draw-6-fingers-on-hands-and-how/>

SIMILARITIES AND DIFFERENCES

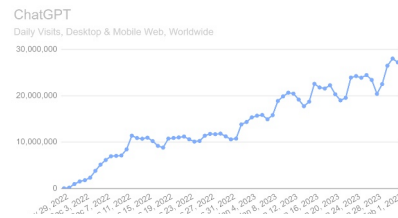
New generative models



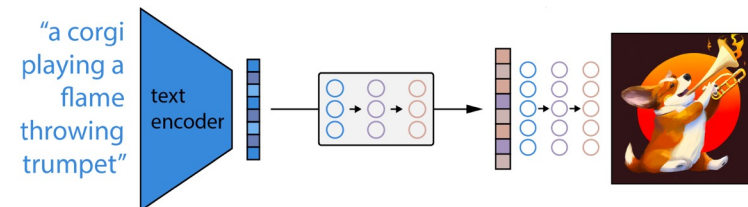
Lots of data



Models publicly available

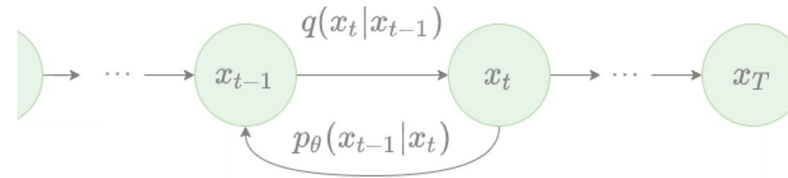


Intuitive ways of interacting with models and interpreting their output

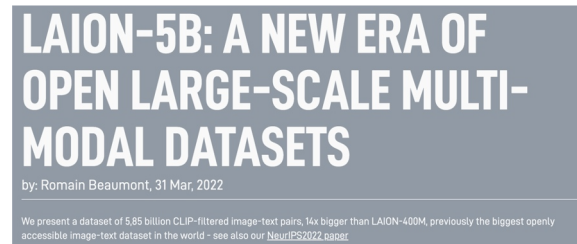


SIMILARITIES AND DIFFERENCES

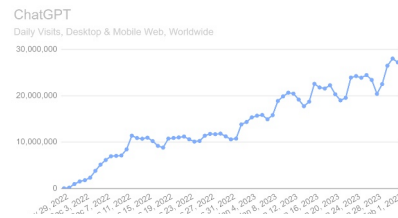
New generative models



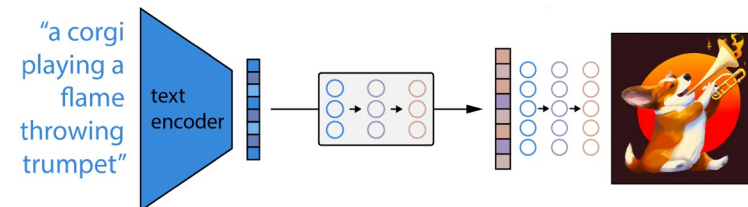
Lots of data



Models publicly available



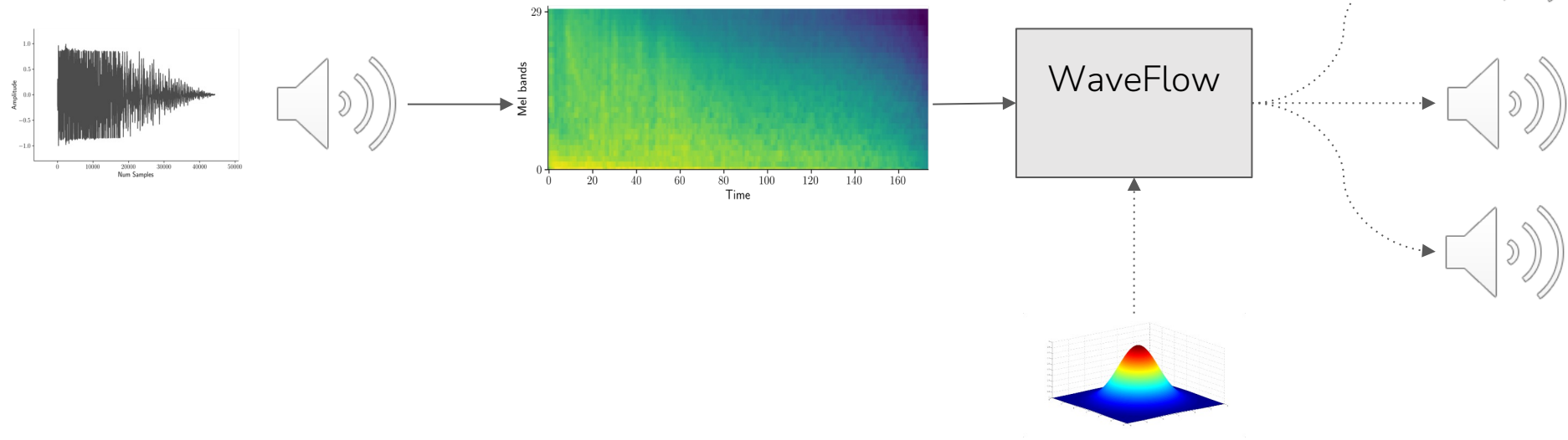
~~Intuitive~~ ways of interacting with models and interpreting their output



NEURAL SYNTHESIS OF SOUND EFFECTS

SYNTHESIZED

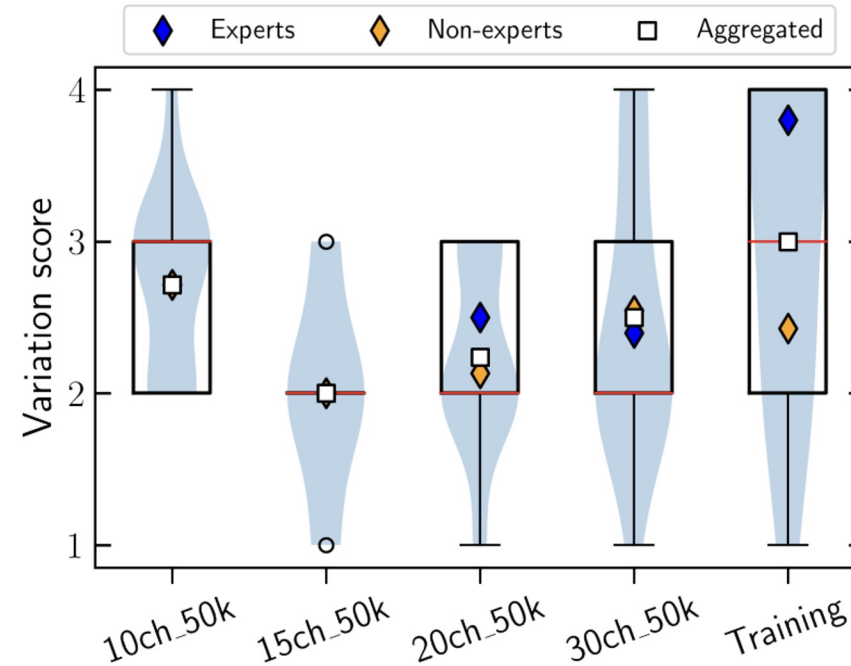
ORIGINAL



Andreu, Sergi, and Monica Villanueva Aylagas. "Neural Synthesis of Sound Effects Using Flow-Based Deep Generative Models." *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*. Vol. 18. No. 1. 2022.

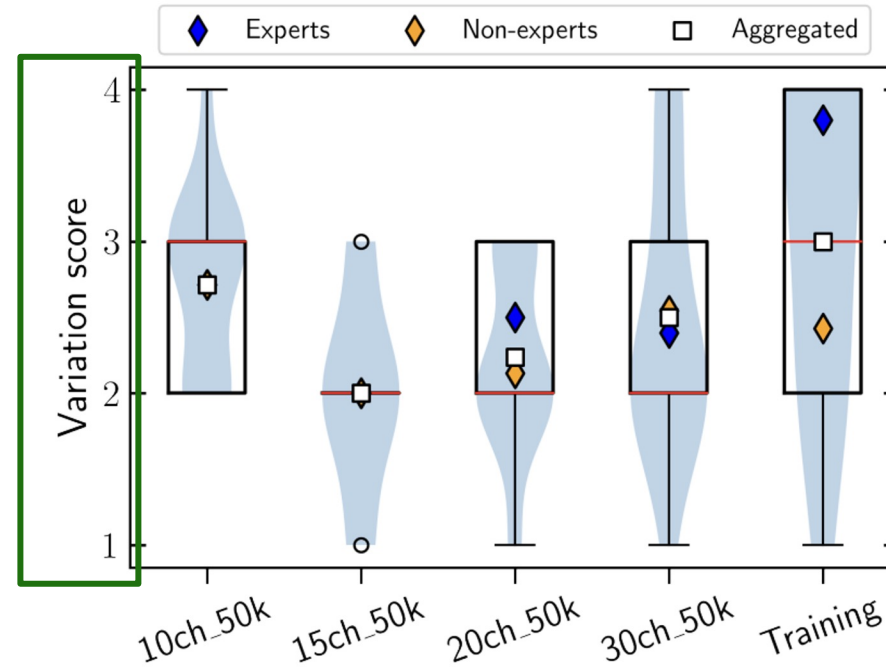
Ping, Wei, et al. "WaveFlow: A compact flow-based model for raw audio." *International Conference on Machine Learning*. PMLR, 2020.

NON-EXPERTS FAIL TO EVALUATE THE QUALITY



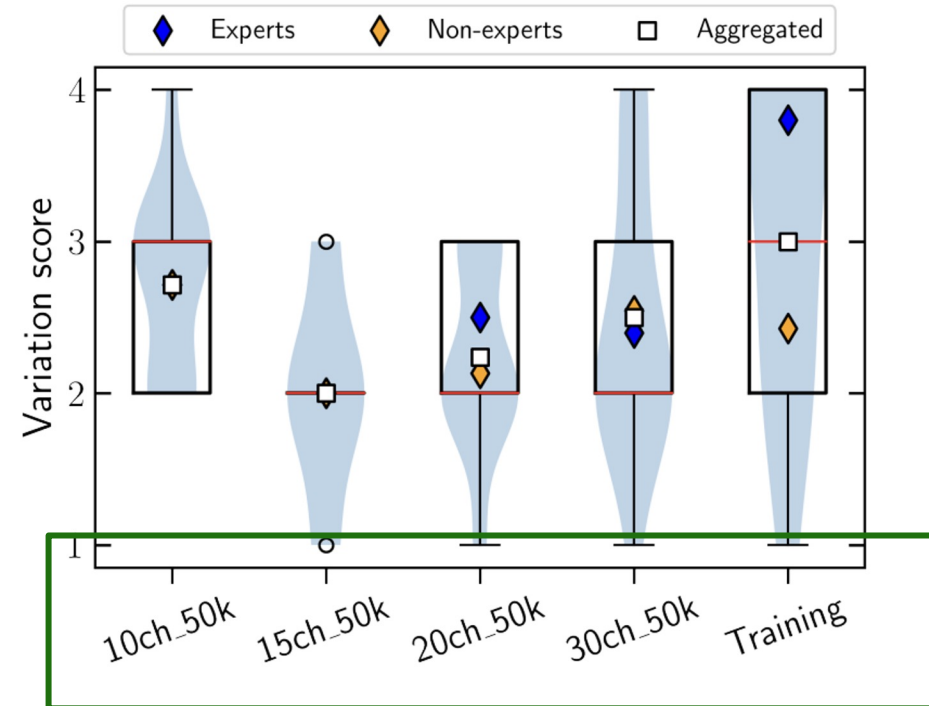
(a) Variation comparison of models with different dimensionality of the mel spectrogram conditioner.

NON-EXPERTS FAIL TO EVALUATE THE QUALITY



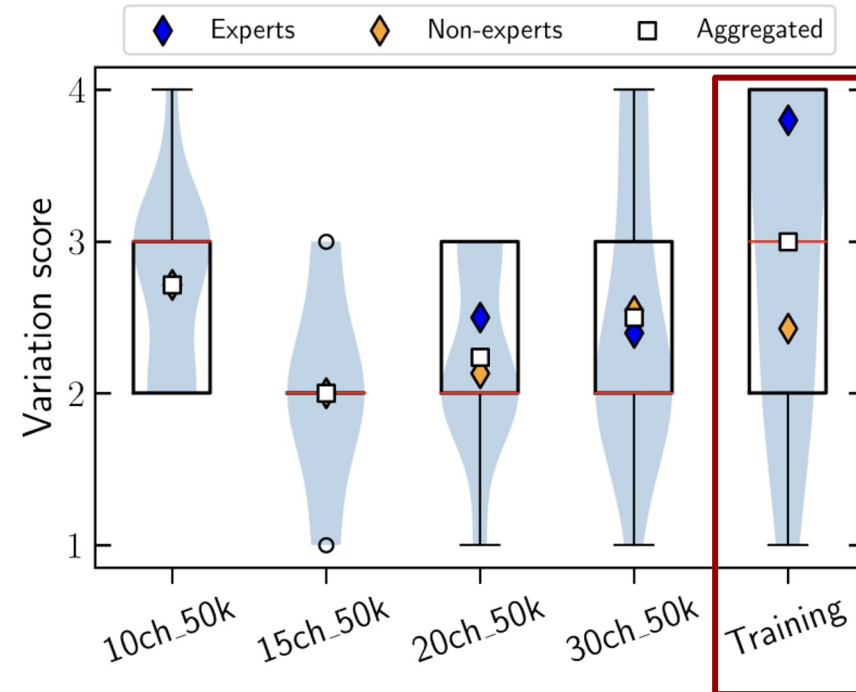
(a) Variation comparison of models with different dimensionality of the mel spectrogram conditioner.

NON-EXPERTS FAIL TO EVALUATE THE QUALITY



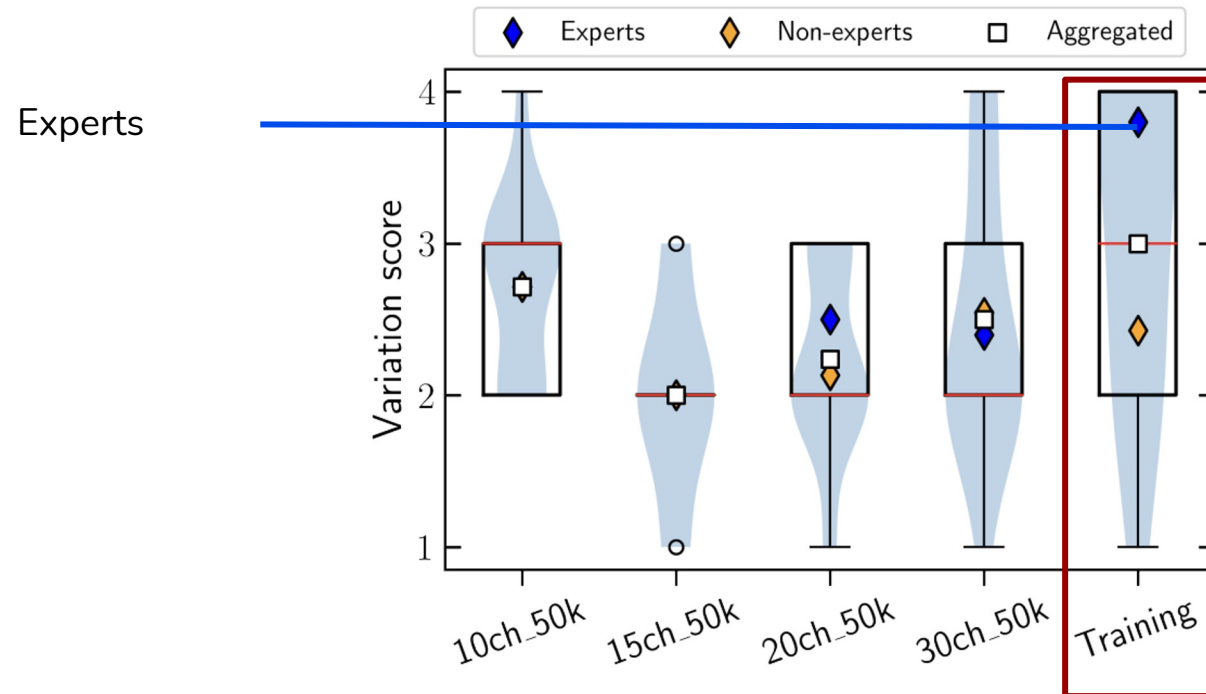
(a) Variation comparison of models with different dimensionality of the mel spectrogram conditioner.

NON-EXPERTS FAIL TO EVALUATE THE QUALITY



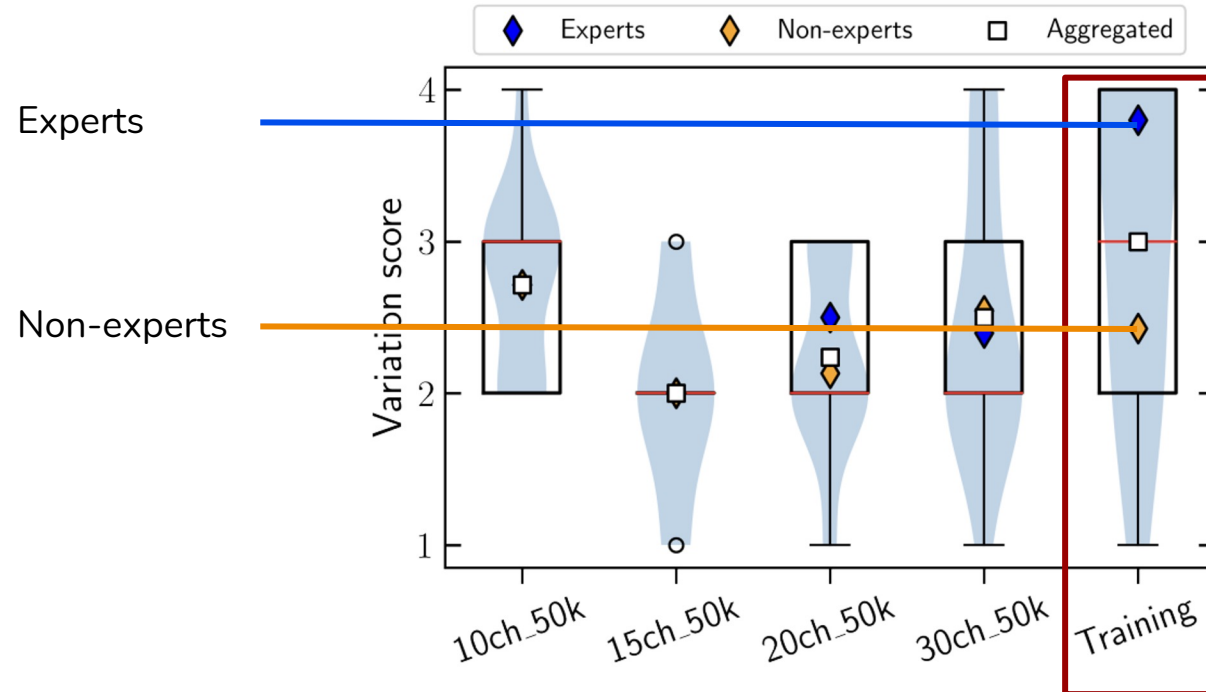
(a) Variation comparison of models with different dimensionality of the mel spectrogram conditioner.

NON-EXPERTS FAIL TO EVALUATE THE QUALITY



(a) Variation comparison of models with different dimensionality of the mel spectrogram conditioner.

NON-EXPERTS FAIL TO EVALUATE THE QUALITY



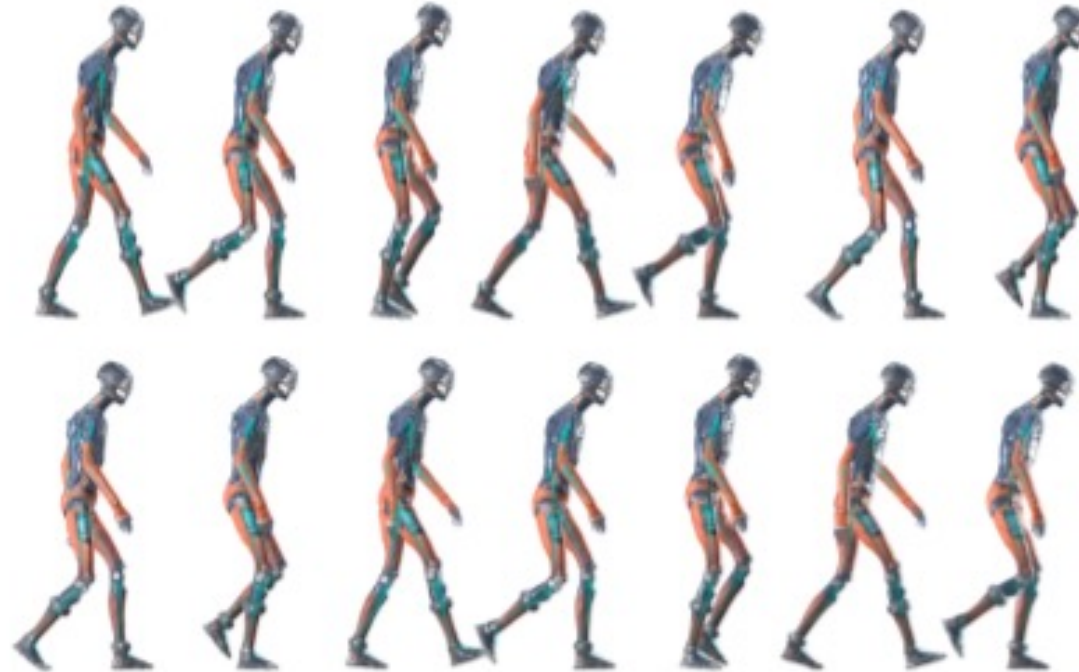
(a) Variation comparison of models with different dimensionality of the mel spectrogram conditioner.

HOW ABOUT THE LOSS?

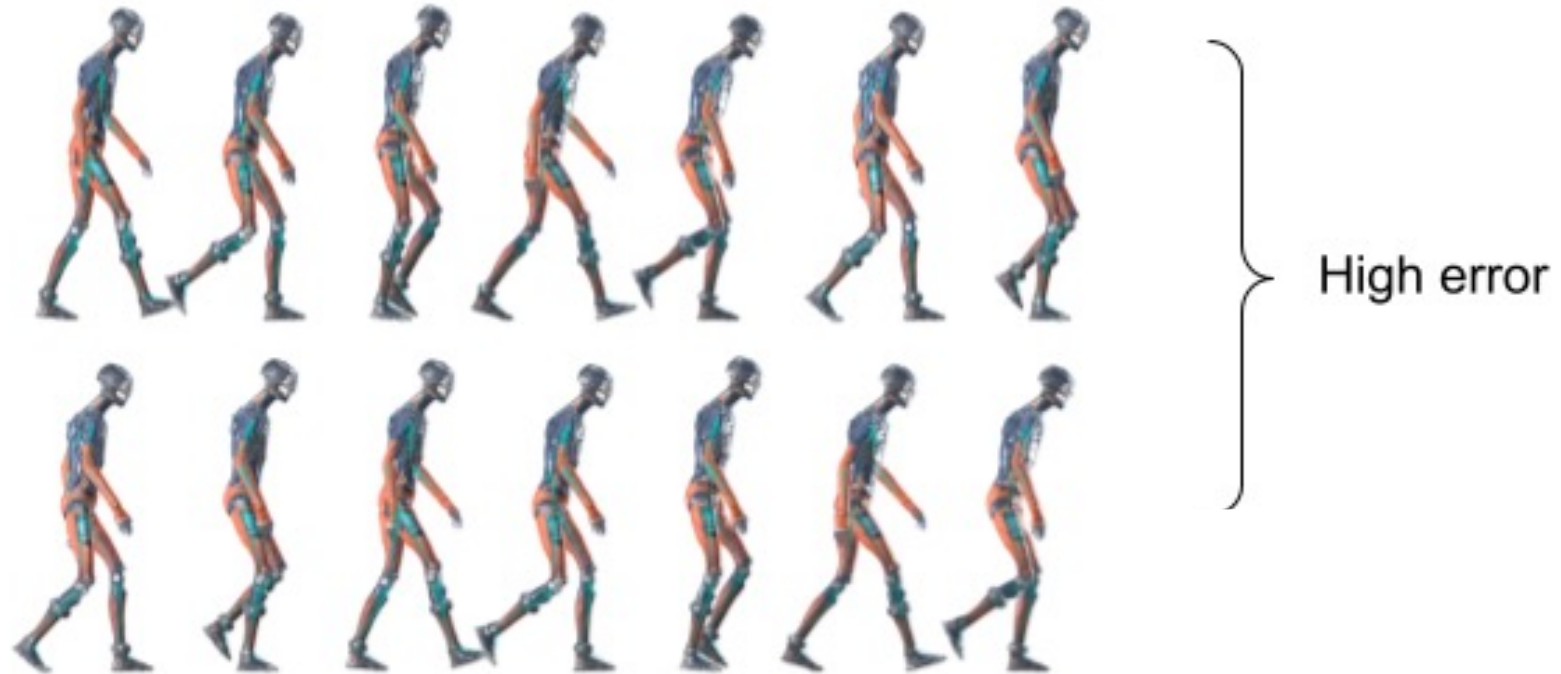
THE LOSS IS OFTEN USELESS



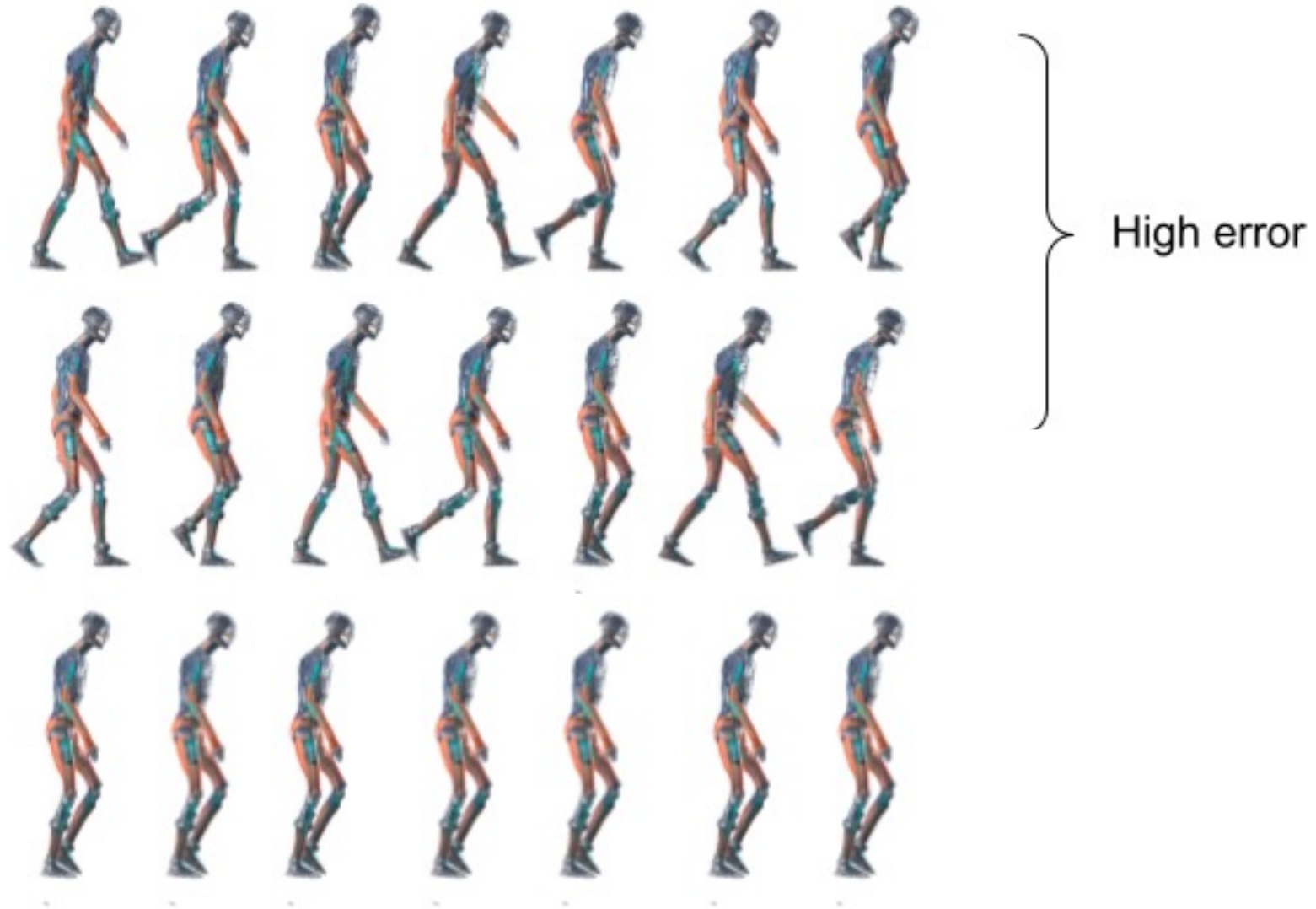
THE LOSS IS OFTEN USELESS



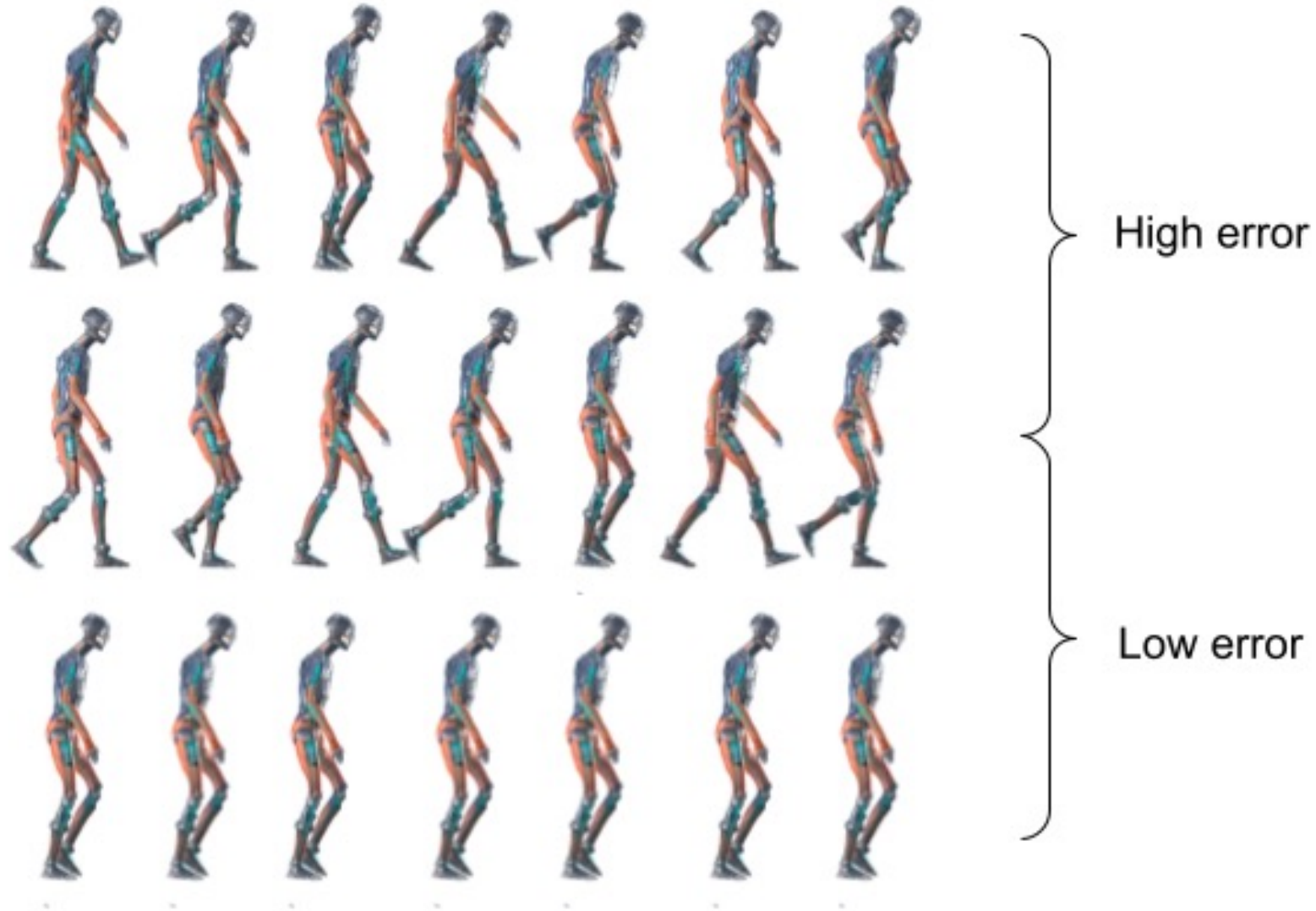
THE LOSS IS OFTEN USELESS



THE LOSS IS OFTEN USELESS



THE LOSS IS OFTEN USELESS



THIS IS A PROBLEM.

THIS IS A PROBLEM

In Game Development

we need to meet certain **quality standards**.

Failing those might mean that the model
is not adopted by the end-user.



THIS IS A PROBLEM

It is **impossible to quickly iterate models** because there exists no good automatic way to assess quality.



SOLUTIONS

HEURISTICS

LIP-SYNC ANIMATIONS

Bilabial consonants (/p b m/) correspond to face poses requiring closed lips.

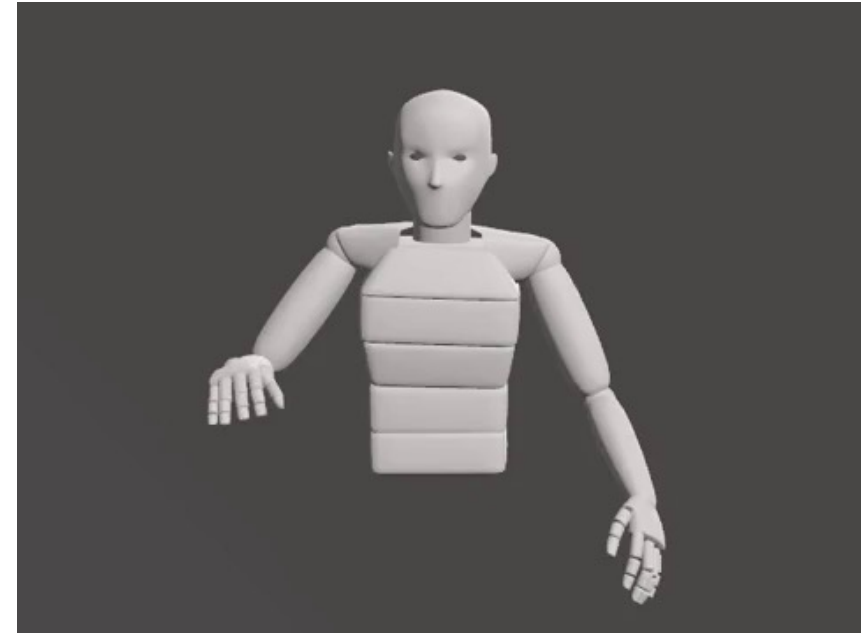
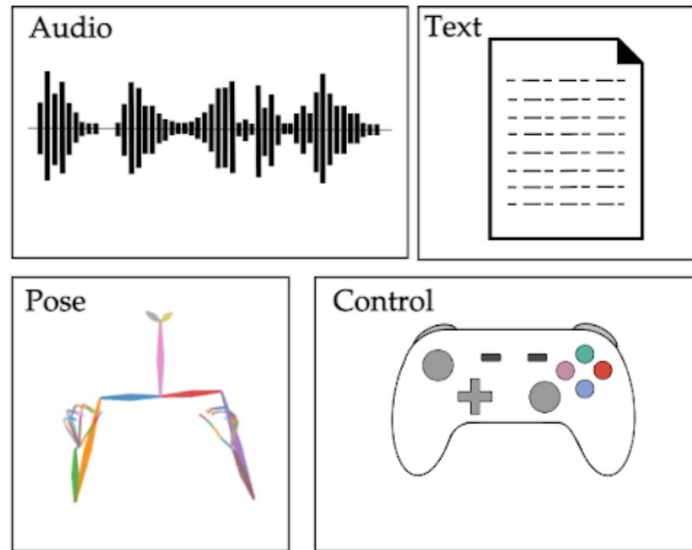
Table 2: *Quantitative metrics for the ablated models and VOCA. We report mean and standard deviation over 5 runs, aggregating also over 12 gold IDs for VOCA.*

Model	Precision %	Recall %	F1 %
<i>Ours</i>	62.01 ± 4.54	55.86 ± 14.43	57.55 ± 7.13
VOCA	86.87 ± 6.93	30.06 ± 13.46	42.79 ± 16.14

HUMAN EVALUATION

GENEA* CHALLENGE

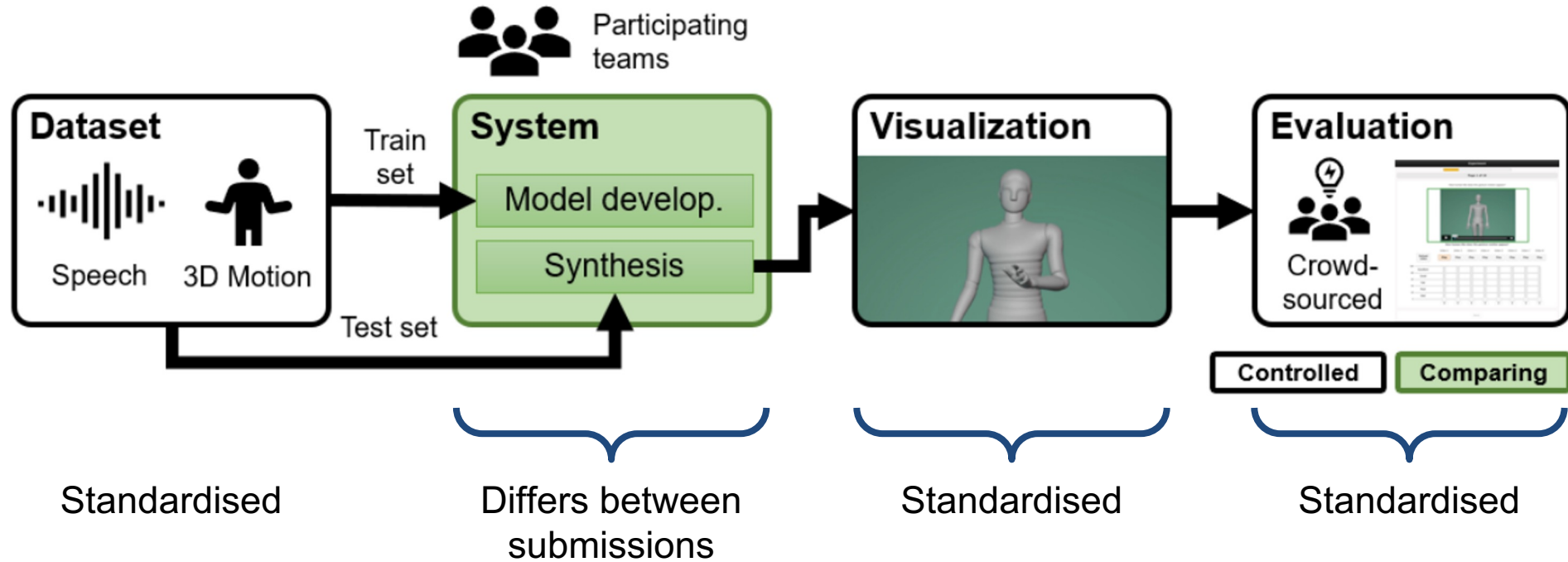
*GENERATION AND EVALUATION OF NON-VERBAL BEHAVIOUR FOR EMBODIED AGENTS



Yoon, Y., Wolfert, P., Kucherenko, T., Viegas, C., Nikolov, T., Tsakov, M., & Henter, G. E. (2022, November). The GENE Challenge 2022: A large evaluation of data-driven co-speech gesture generation. In *Proceedings of the 2022 International Conference on Multimodal Interaction* (pp. 736-747).

GENEA* CHALLENGE

*GENERATION AND EVALUATION OF NON-VERBAL BEHAVIOUR FOR EMBODIED AGENTS



Yoon, Y., Wolfert, P., Kucherenko, T., Viegas, C., Nikolov, T., Tsakov, M., & Henter, G. E. (2022, November). The GENE Challenge 2022: A large evaluation of data-driven co-speech gesture generation. In *Proceedings of the 2022 International Conference on Multimodal Interaction* (pp. 736-747).

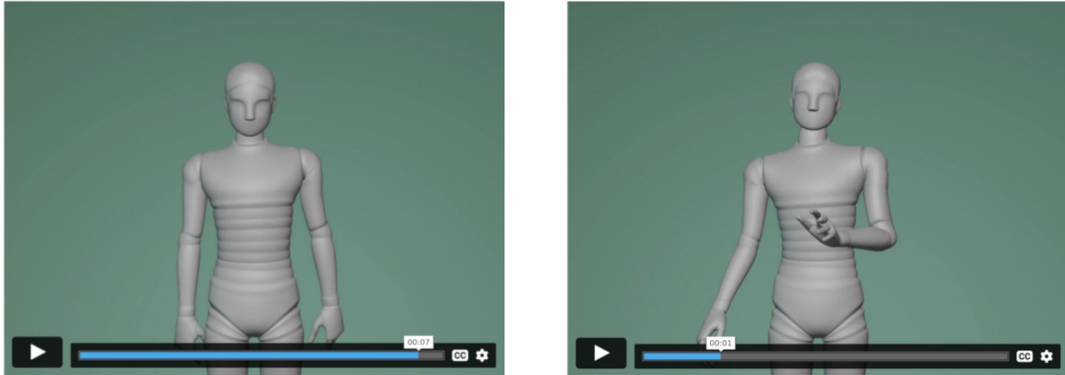
GENEA* CHALLENGE

*GENERATION AND EVALUATION OF NON-VERBAL BEHAVIOUR FOR EMBODIED AGENTS

Experiment

Page 1 of 40

Please indicate which character's motion best matches the speech, both in terms of rhythm and intonation and in terms of meaning.

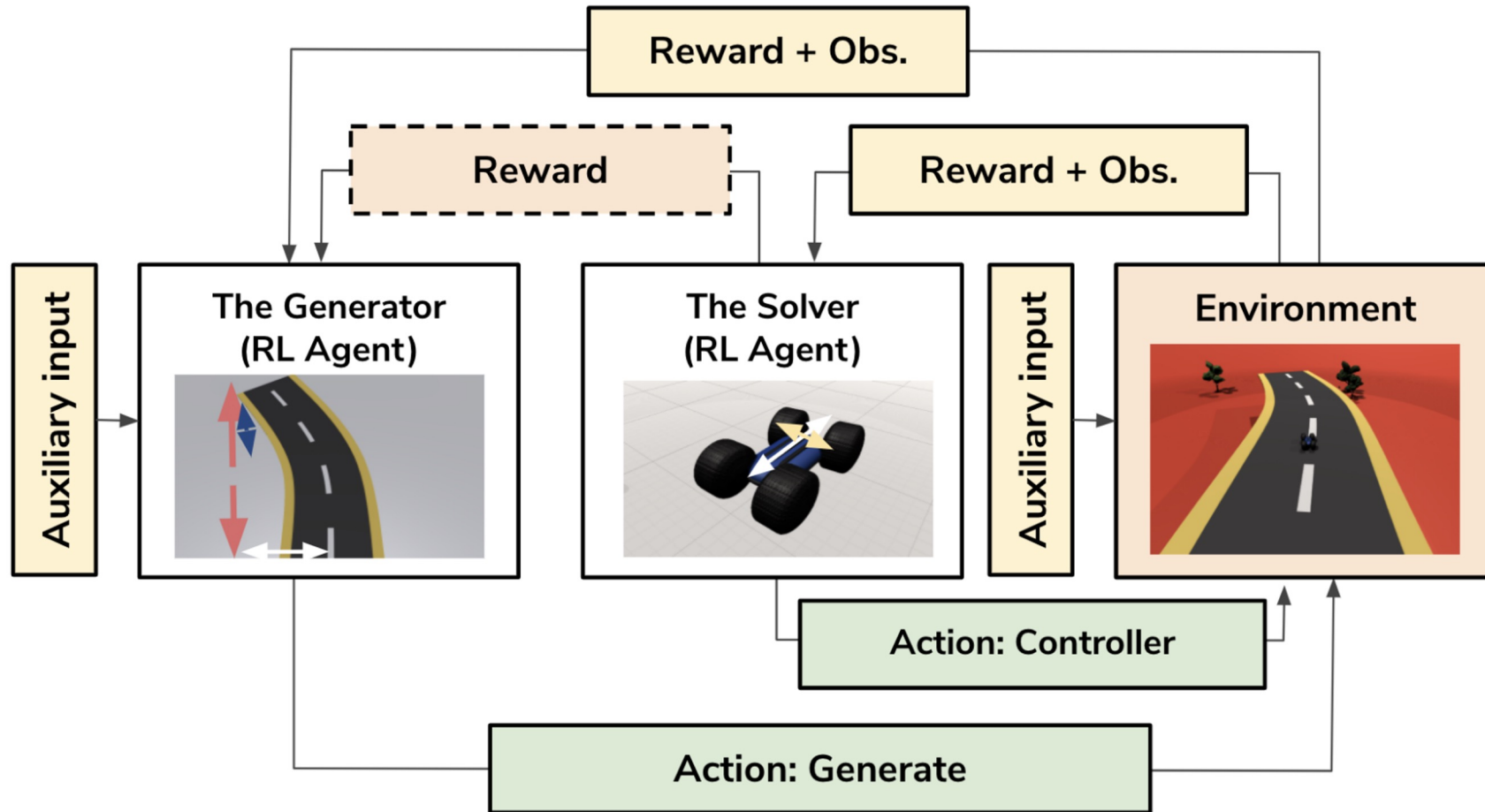


- The character in the video on the *left*
- The character in the video on the *right*
- They are equal

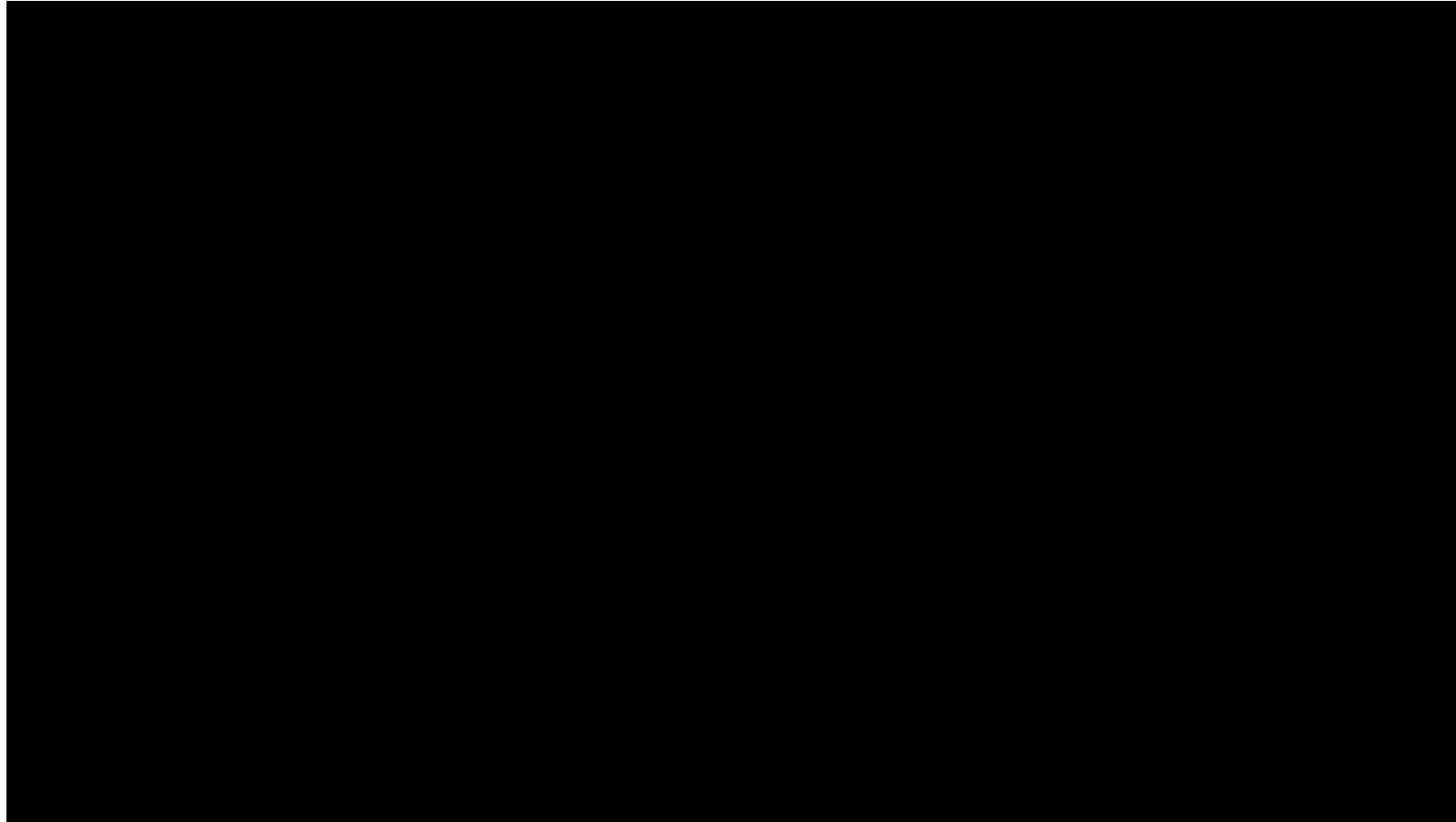
Yoon, Y., Wolfert, P., Kucherenko, T., Viegas, C., Nikolov, T., Tsakov, M., & Henter, G. E. (2022, November). The GENE Challenge 2022: A large evaluation of data-driven co-speech gesture generation. In *Proceedings of the 2022 International Conference on Multimodal Interaction* (pp. 736-747).

REINFORCEMENT LEARNING

ADVERSARIAL REINFORCEMENT AGENTS



ADVERSARIAL REINFORCEMENT AGENTS



Reliable,

automatic,

WHAT IS STILL MISSING?

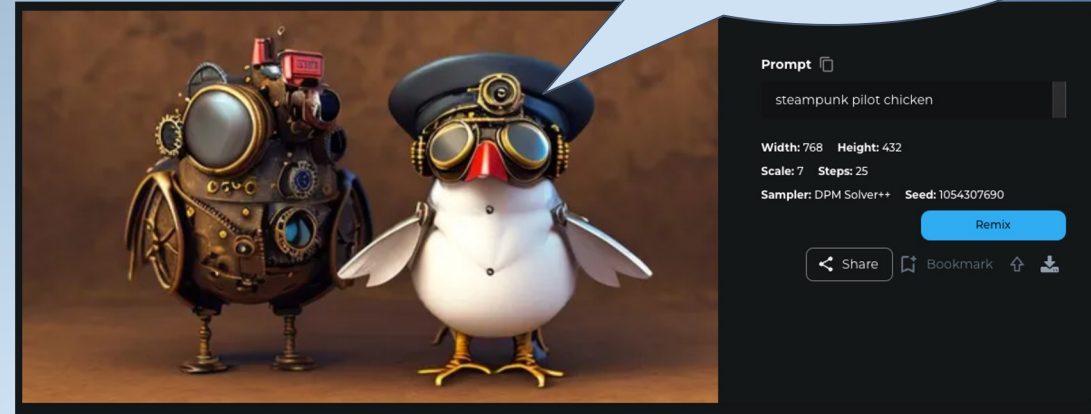
scalable

ways to assess quality.

For questions,
ideas or fika invites
contact me at

Judith Bütepage
jbutepage@ea.com

or find me on LinkedIn.



SEED // SEARCH FOR EXTRAORDINARY
EXPERIENCES DIVISION
seed.ea.com