



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

Incorporating a Machine Learning Research Project into Game Audio Production: The **ExFlowSions** Case Study

Mónica Villanueva and Jorge García
SEED – Electronic Arts

TEAM INTRODUCTIONS

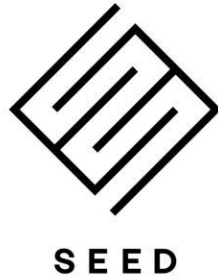


Mónica Villanueva
Research Engineer
SEED



Jorge García
Software Engineer
Tools and Technology - SEED

TEAM INTRODUCTIONS



Search for Extraordinary Experiences Division (SEED) - Electronic Arts

seed.ea.com

We are a cross-disciplinary team within Electronic Arts.
Our mission is to explore, build and help define the
future of interactive entertainment.

- Cross-disciplinary Art, Engineering, and Applied Research team
- Diverse team from Games, Tech, and Academia, circa 2015
- Applied research on risky topics for the benefit of all at EA
- AI, ML, Content Creation, Rendering, Physics, Animation
- Measurable impact and technology artifacts delivered in games
- Sharing through publications, presentations, and open source



AGENDA

1. Introduction to ExFlowSions
2. Motivations and requirements
3. Research codebase and Tool API
4. Tool design
5. Implementation - Gradio
6. Testing approaches
7. Demos
8. Feedback gathering
9. Future work
10. Conclusions and takeaways
11. Q&A



Image generated with Bing-Dall-E3

Prompt: "Explosion with loudspeakers coming out from it and a digital waveform at the bottom"

ExFlowSions INTRO

- Model based on Normalizing Flows
 - Trained on a small dataset (10m)
- Goal
 - Generates variations out of an example
 - Performs style transfer to explosion
- Published at [AIIDE 2022](#)¹

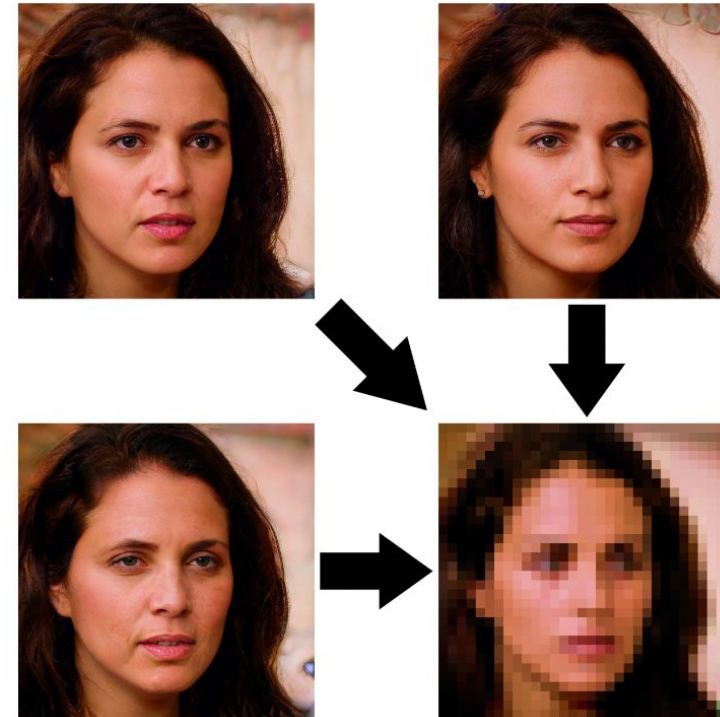
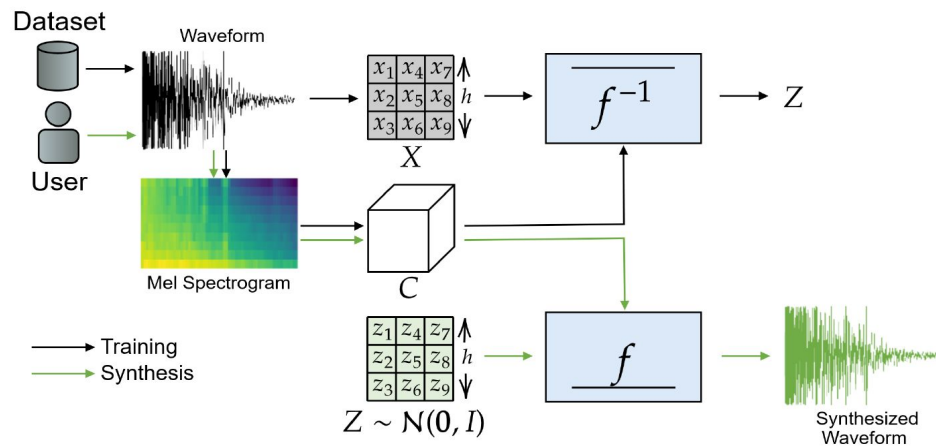
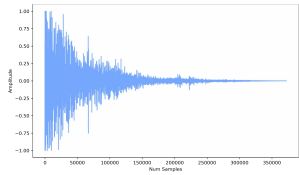


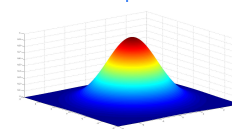
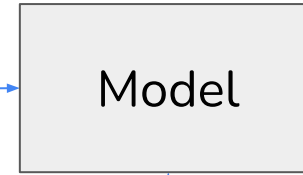
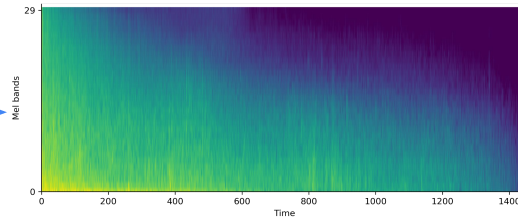
Figure from Menon, Sachit, et al. "Pulse: Self-supervised photo upsampling via latent space exploration of generative models." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020.

ExFlowSions Quick Demo

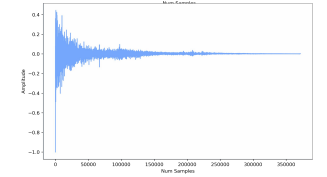
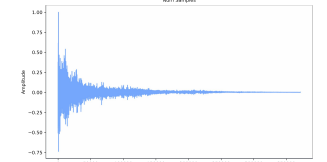
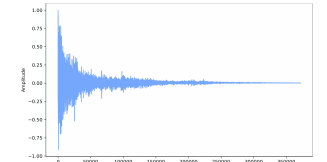
ORIGINAL



Explosion sound by: DICE



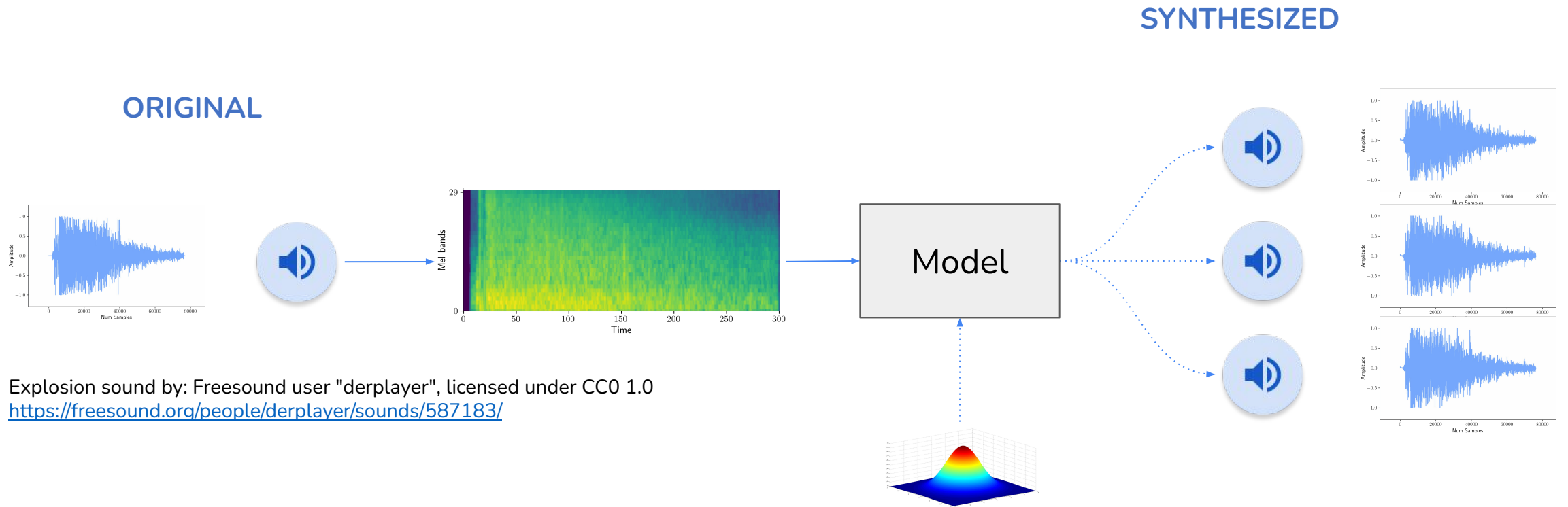
SYNTHESIZED



Results - Target Sample Rate (48kHz)

In Distribution Explosions

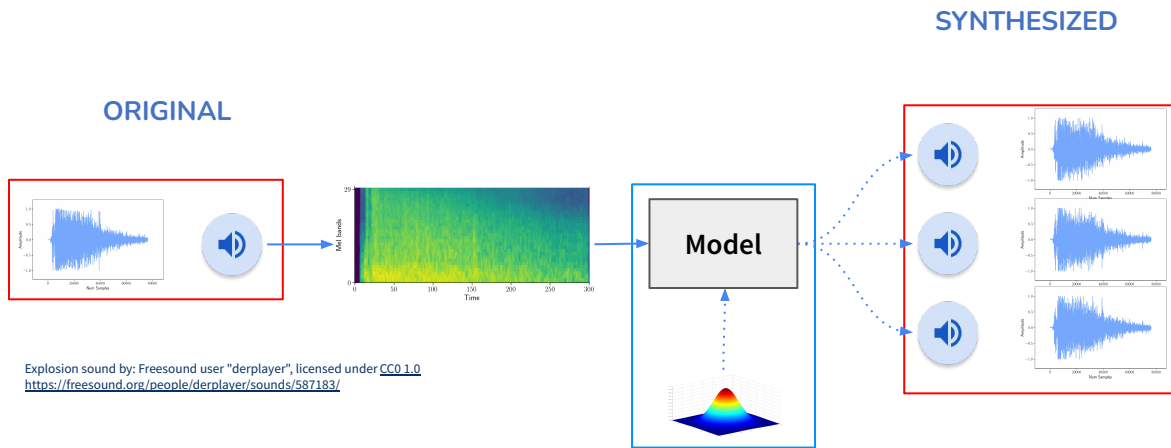
ExFlowSions QUICK DEMO



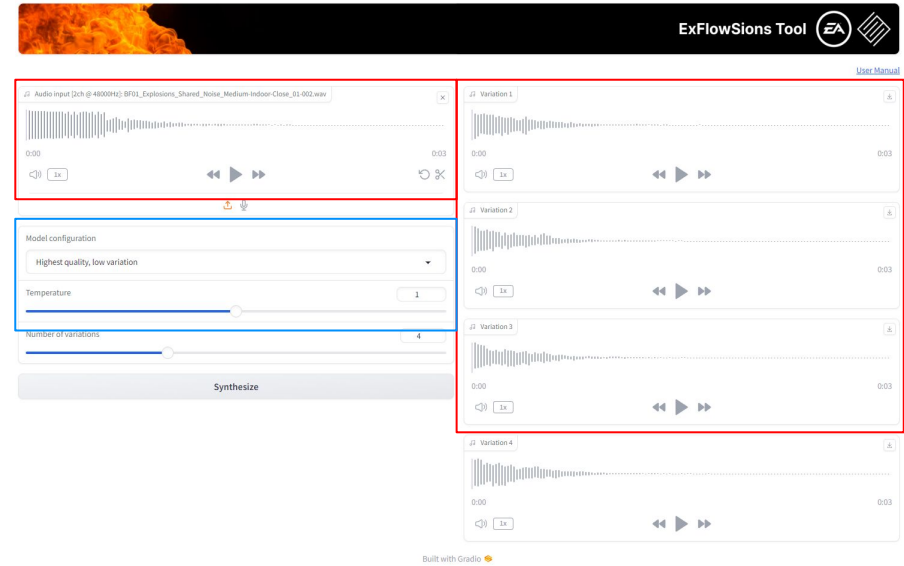
Results - Target Sample Rate (48kHz)

Out of Distribution Explosions

ExFlowSions QUICK DEMO



Explosion sound by: Freesound user "derplaver", licensed under [CC0 1.0](https://creativecommons.org/licenses/by/4.0/)
<https://freesound.org/people/derplaver/sounds/587183/>



TOOL MOTIVATIONS AND REQUIREMENTS GATHERING

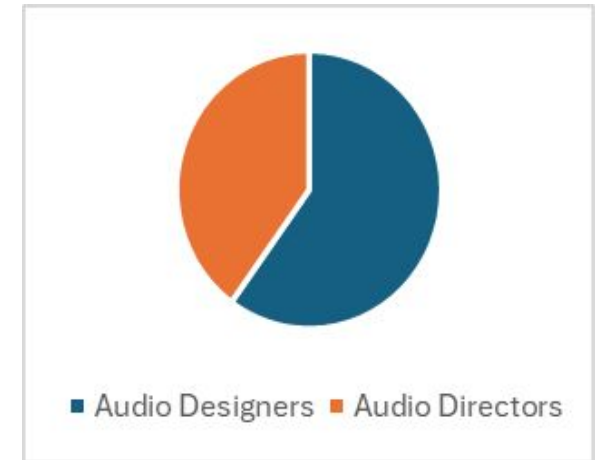
Why create an ExFlowSions tool?

- Have a tool sitting between **research** and **production** to evaluate its value
 - Put the ExFlowSions **model into the hands** of audio designers
 - **Evaluate the audio quality** of the model more easily
- Investigate what are the possibilities of integrating the model into production
- Offer a tool that can be used creatively and explore alternate use cases and requirements

TOOL MOTIVATIONS AND REQUIREMENTS GATHERING

We interviewed five power users

- They all wanted a tool that is **easy to use**
- The tool had to work on **Windows**, their main development environment
- We clarified the **input and output formats** that needed to be supported: WAV 16-bit, 24-bit and 32-bit
- **Output sample rate** is 48kHz (initial requirement for the model)
- Having **multi channel support** (stereo and above) was also important and requested during interviews
- **CPU inference**, on top of GPU inference, was provided for users without compatible graphics cards



RESEARCH CODEBASE AND REFACTORING

Working from a research codebase

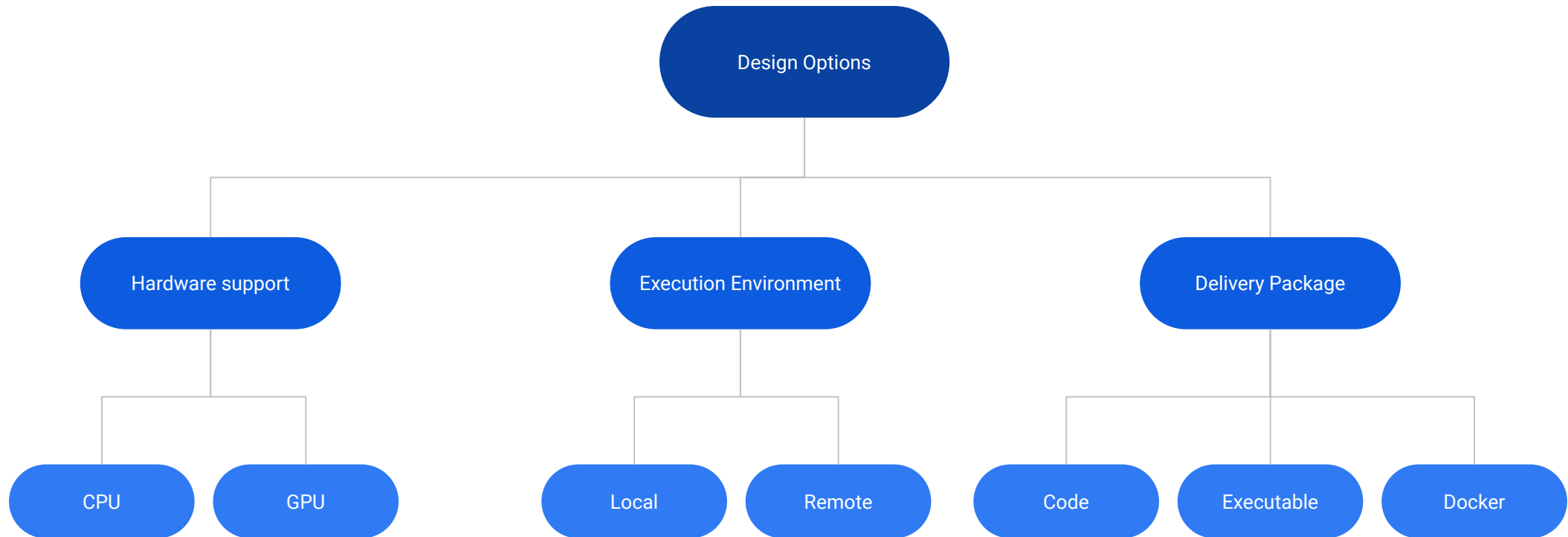
- The original codebase included **various code experiments**, some unused
- Some code needed to be **refactored** for production use (e.g., Python modules adaptation)
- **Naming conventions** were consolidated
- Code **updates and migration** to newer PyTorch versions were also carried out (e.g., FP16 support)

TOOL API INTRO

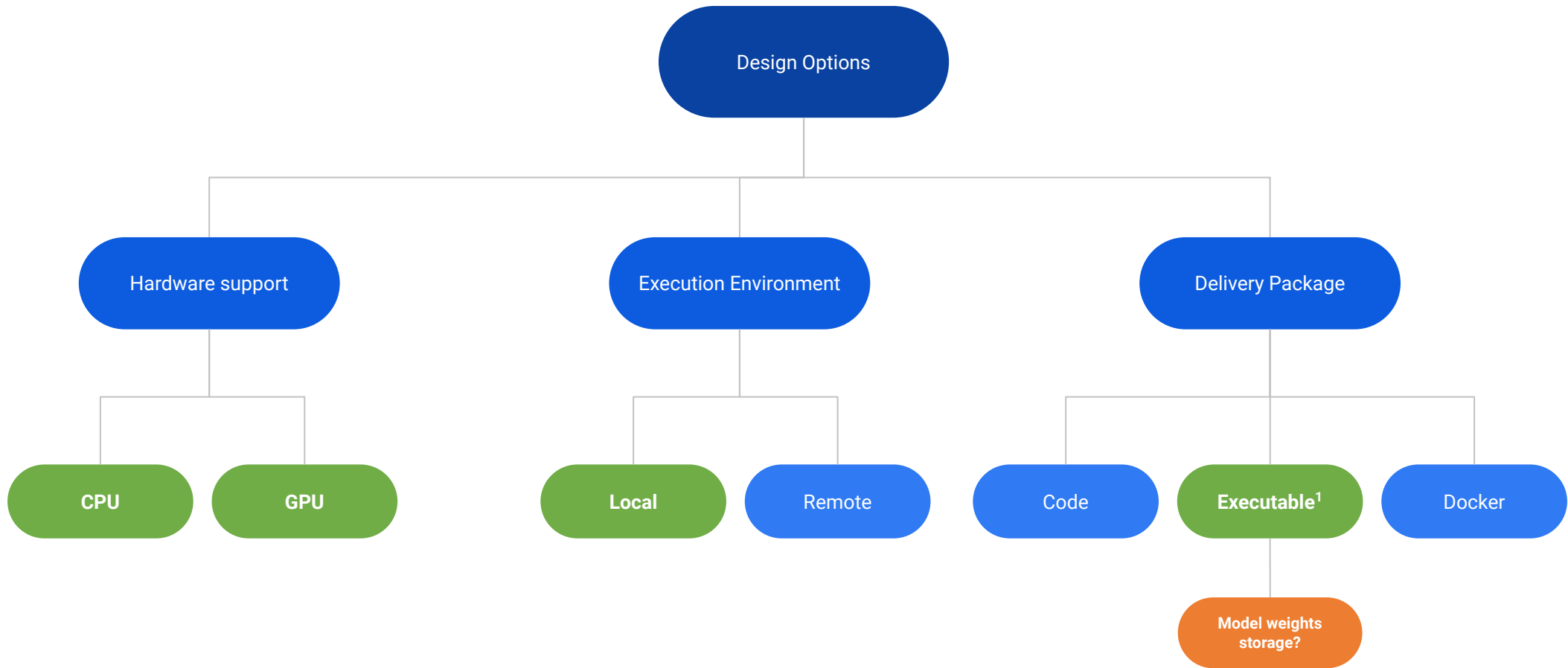
Three functions are used by the tool

- Model initialization
- Data loader creation
- Synthesis function

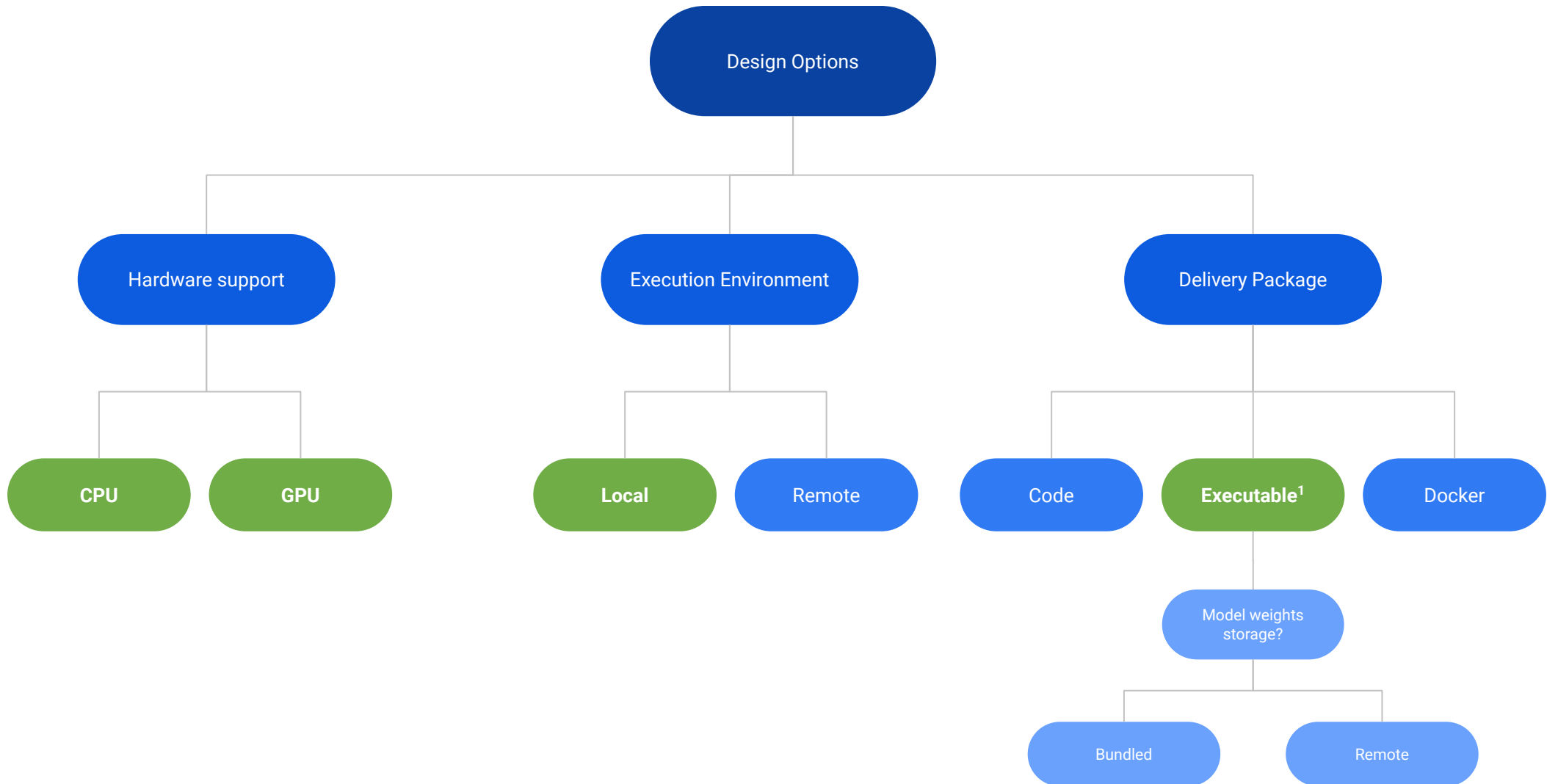
TOOL DESIGN



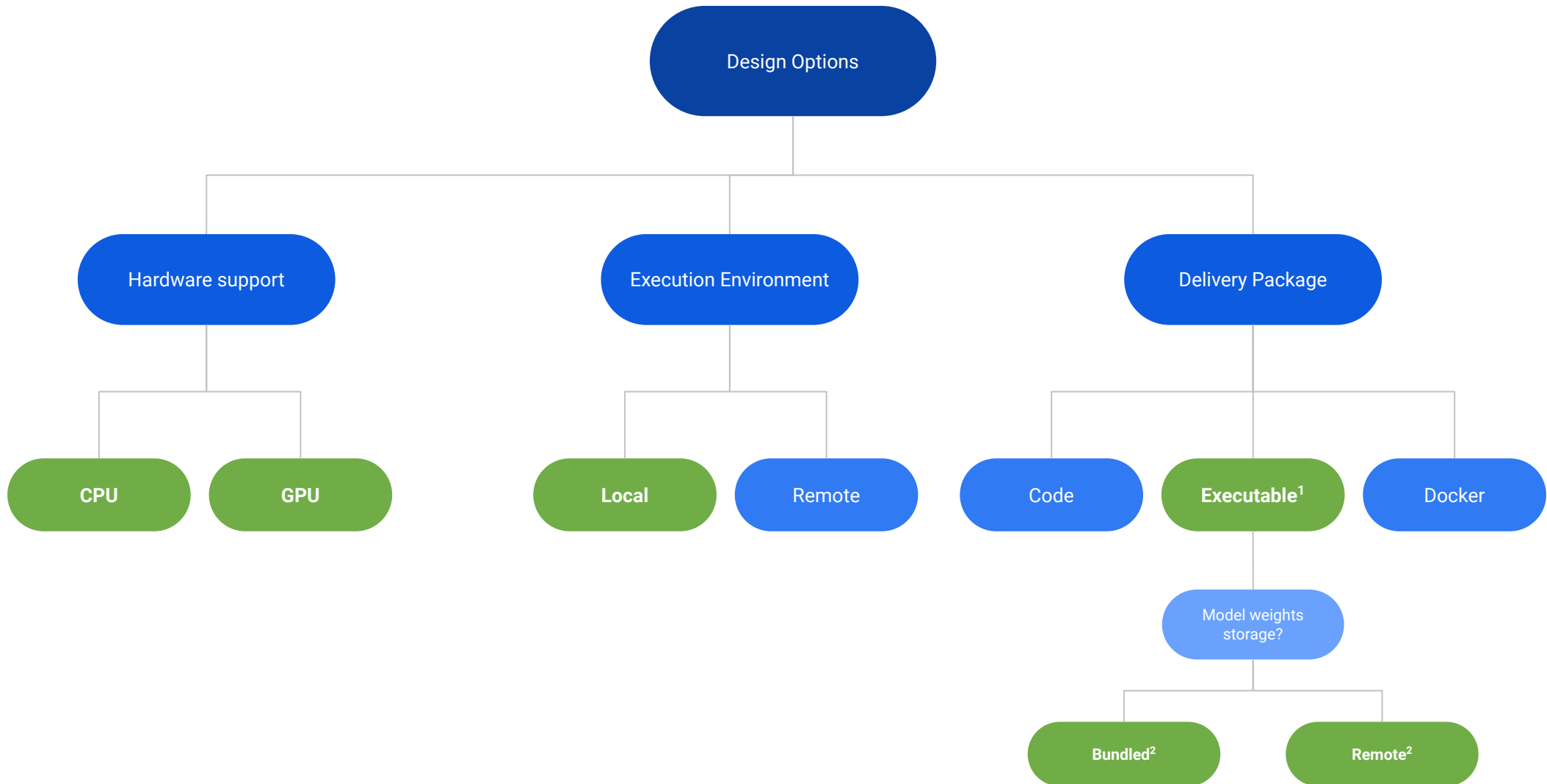
TOOL DESIGN



TOOL DESIGN

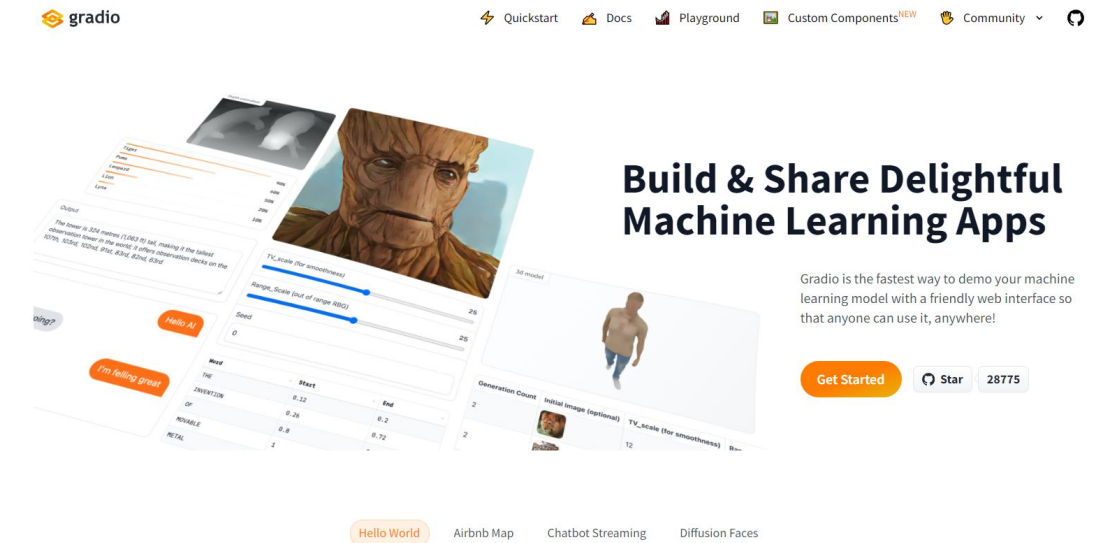


TOOL DESIGN



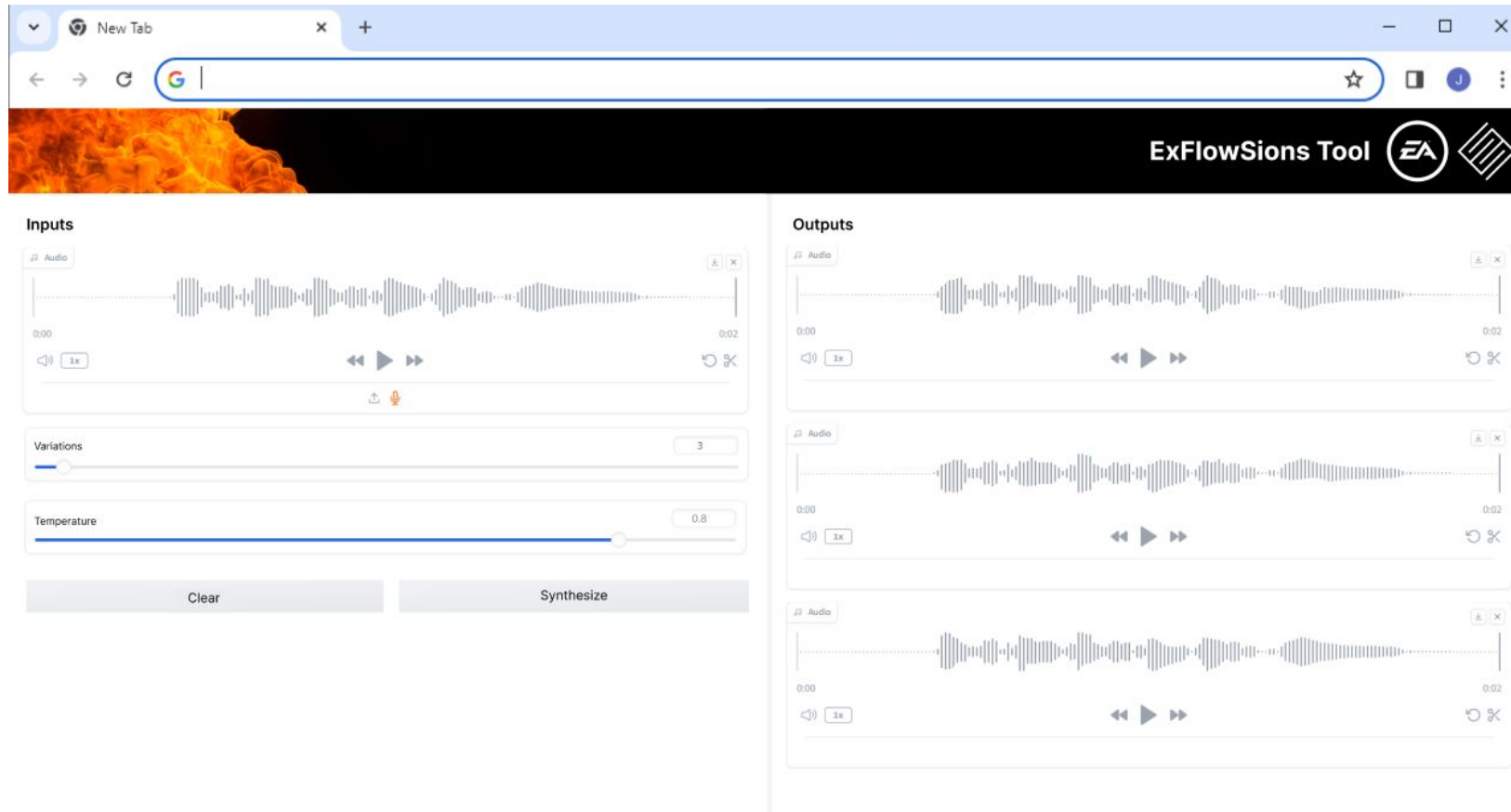
TOOL IMPLEMENTATION - GRADIO

- Alternatives evaluated: Gradio, Streamlit
- Decided on **Gradio**: audio components availability, features and easiness of use
- All UI is **written in Python**, no need to write wrappers for the ML model from other language
- **Rapid prototyping** and iteration time
- **UI customization**
- **Gradio cons and limitations**



<http://www.gradio.app>

TOOL IMPLEMENTATION - FIGMA PROTOTYPE

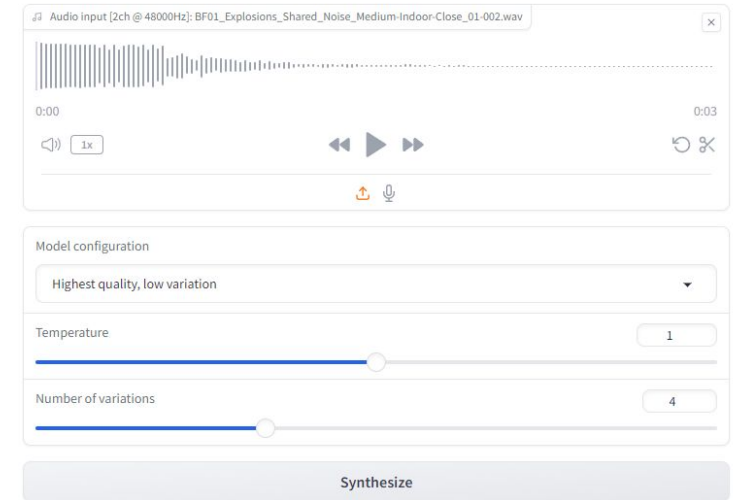


Validate the initial UI approach

TOOL TESTING OVERVIEW

Various approaches

- Manual testing: five test cases covering the hot paths (inference process, parameter changes, etc)
- Unit testing: seven unit tests with the following coverage
 - Four unit tests for the UI parameters (variations, temperature...)
 - Synthesis forcing CPU mode
 - Synthesis with mono input file (mono output)
 - Synthesis with many channels (six) in the input file
- Continuous Integration (CI) pipeline in GitLab uses the Gradio client API for triggering and running the unit tests



TOOL DEMO

References

- Explosion sounds:
 - DICE
- WSHH_Whoosh Futuristic Sci Fi Short Swish Surround_PSE_DRKM_038a.wav:
 - Pro Sound Effects artist “Saro Sahihi”, “Dark Matter” library, 2023 - <http://www.prosoundeffects.com>

FEEDBACK GATHERING (FORM)

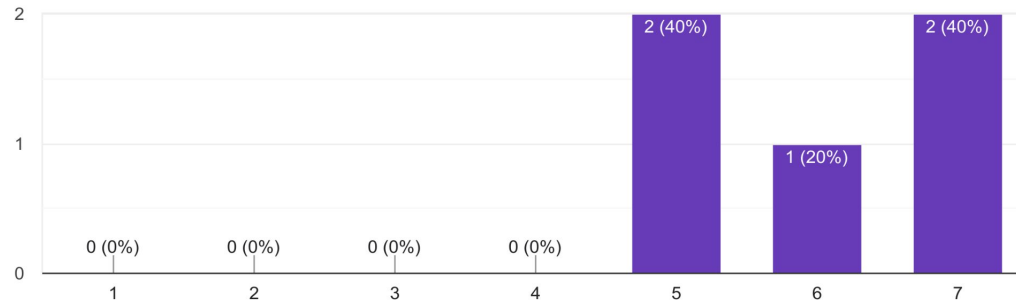
We asked a set of questions...

- Tool UI and **usability** (two questions)
- Model **performance and quality** (one question)
- Tool **impact and value** (two questions)
- **Open** feedback (one question)

FEEDBACK ANALYSIS

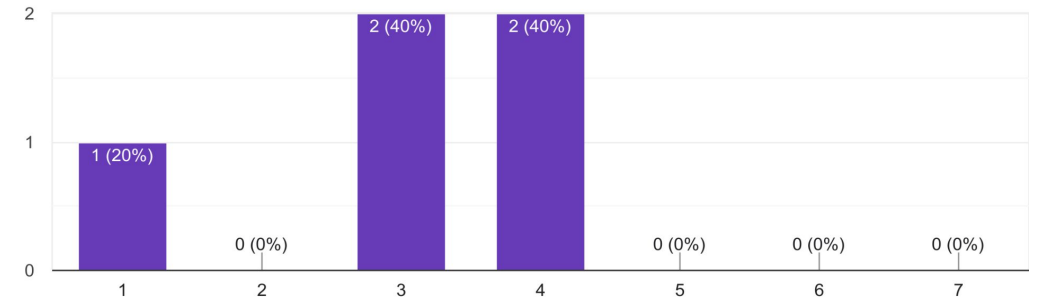
How easy was the tool to use?

5 responses



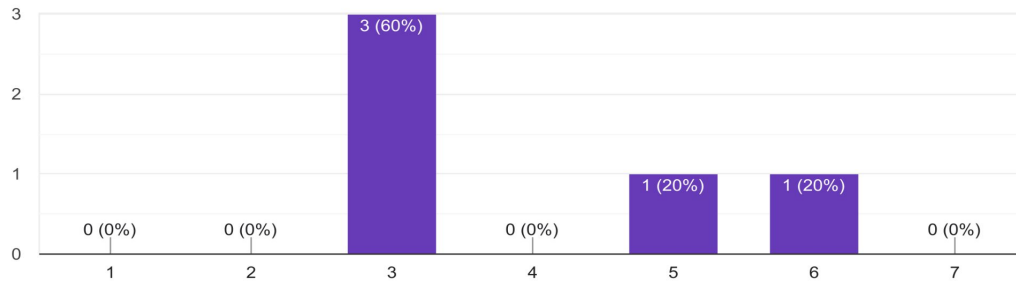
How easy was it to get good sounding results?

5 responses



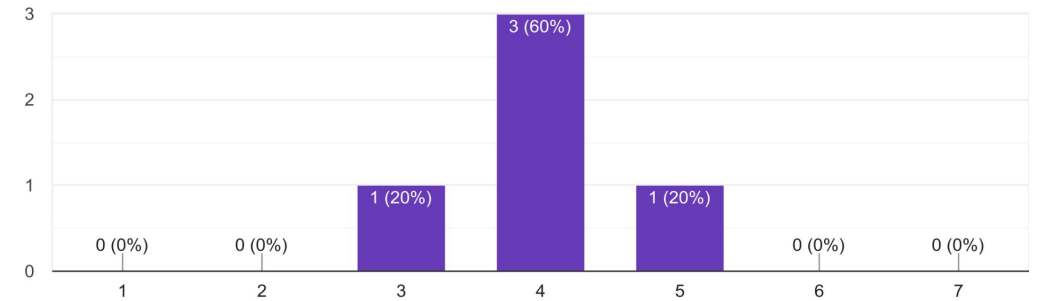
How good was the audio quality of the results that the tool produced? Take into account mono vs multichannel (if applicable)

5 responses



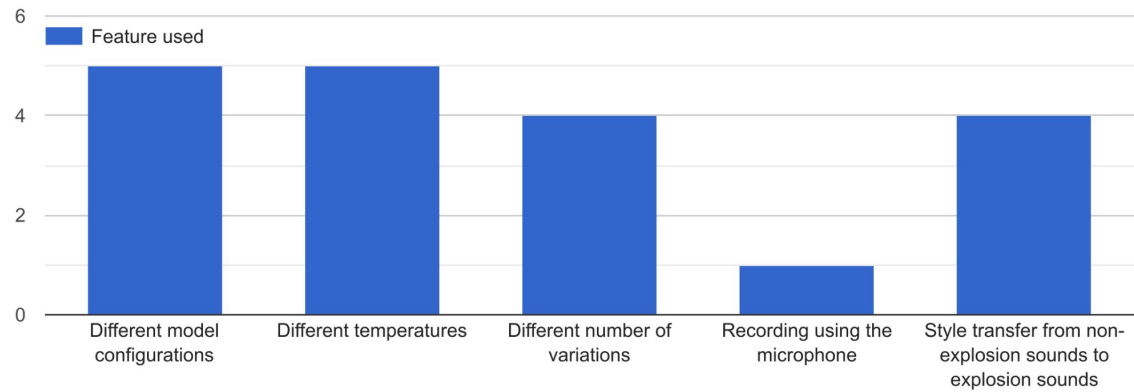
What are the time savings and positive impact this tool could bring to your audio workflows?

5 responses



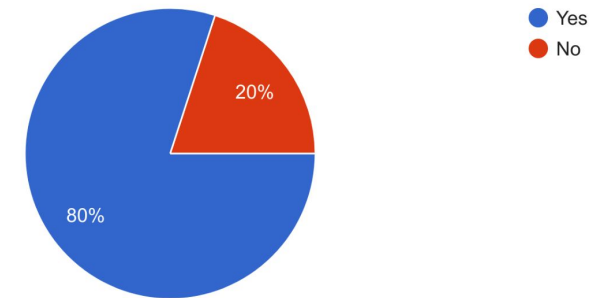
FEEDBACK ANALYSIS

Regarding the different features available in the tool. Have you tried any of the following? (please mark accordingly):



Has this prototype fulfilled your initial expectations and requirements?

5 responses



FEEDBACK ANALYSIS

Positive comments

“It felt **very intuitive** to use the interface, dragging and dropping samples in, using the sliders and previewing the results afterwards.”

“Very easy to use, **neat interface.**”

“... using **default settings sounded good** - although the variation wasn't very pronounced.”

“Generally the **output audio was good**, in the sense that it had interesting artifacts, and it didn't sound ‘low-fi’.”

Areas of improvement

“... the potential is very high but **in its current form I wouldn't use it.**”

“The example we have right now **doesn't provide anything I would use at the moment, or could replicate with existing tool.**”

“For variation creation **I don't think we are there yet.**”

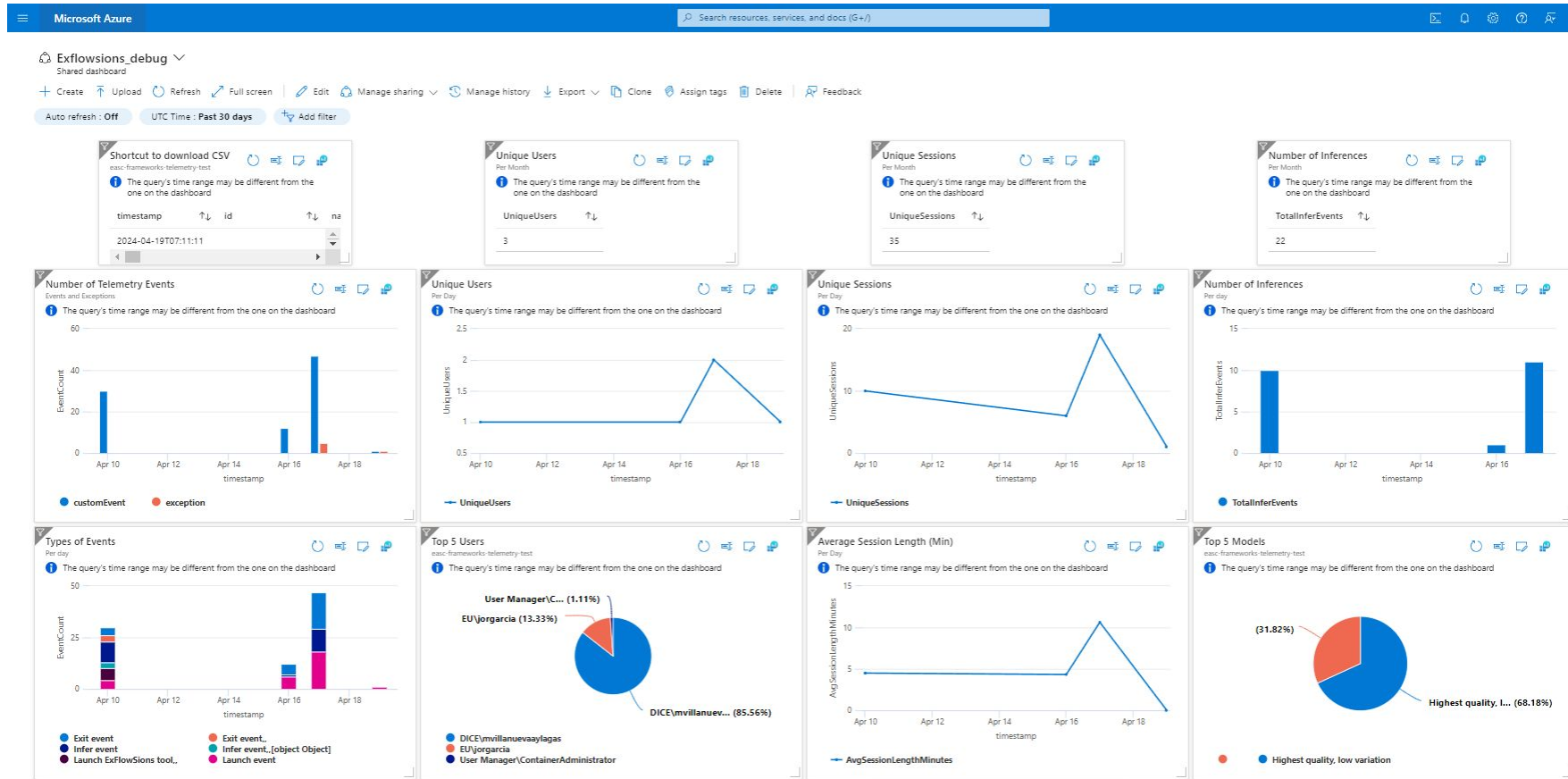
“One time I saw an error related to file length and **processing time was very long** - I ended up cancelling which also took quite a long time- making my pc unusable to some time.”

FUTURE WORK

Tool and model tentative roadmap

- Improve **first inference time**
- Train the model with a **different dataset**
- Improvements:
 - **UI flow performance**
 - User interface **QoL** (e.g., download all generated variations at once)
 - Improve or remove **model post processing**

FUTURE WORK



Telemetry Debug Dashboard

CONCLUSIONS AND TAKEAWAYS

Having a prototype tool bridging **research** and **production**...

- Helps **validating research impact** of ML models
- Facilitates **gathering feedback** from designers
- Provides a **framework for quick experimentation**
- Helps **showcasing research** work
- Promotes **adoption** from game teams
- Provides a **head start on** a future **integration** into production workflows

LEARNINGS & CHALLENGES

Learnings

- UI
- Packaging
- Actionable feedback from users

Challenges

- Gradio bugs and advanced functionality
- Requirements gathering
- Unreliable feedback loop
- [Limited time]

ACKNOWLEDGEMENTS

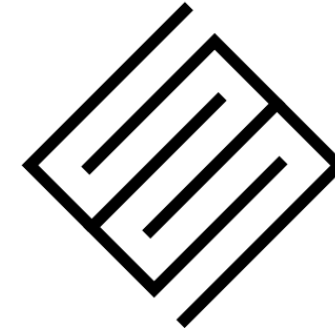
Development

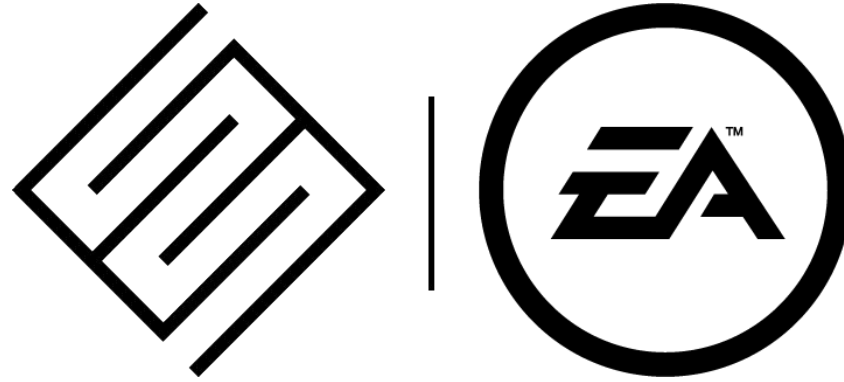
- Sergi Andreu
- Timur Solovev

Users

- Devan Kraushar
- Richard Adrian
- Fernando San Nicolás
- Jon Brunson
- Luc Blanchard
- Rasmus Thorup
- Jean Xu
- Tom Wright

All the SEED team for support and discussions





Questions?

Mónica Villanueva - mvillanuevaaylagas@ea.com

Jorge García - jorgarcia@ea.com

seed.ea.com

STOCKHOLM – MONTRÉAL – LOS ANGELES – SAN FRANCISCO

LONDON – VANCOUVER – REMOTE



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

Presentation Template for Technical Talks

Author Name

SEED – Electronic Arts



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

Presentation Template for Technical Talks with Very Long Titles on Many Lines and Many Authors

First Author
SEED – Electronic Arts

Second Author
Super Long Author's
Name from Long
Institution's Name

Third Author
Third Affiliation

Fourth Author
Forth Affiliation

Fifth Author
Fifth Affiliation

NOTES

- Title slide and final slide:
 - Title slide should be customized
 - Final slide should be updated
- Color of the title on each slide:
- Specs:

TEXT WITH FIGURE

- This sentence will stop close to figure on the right, it should be aligned justify with the border of the figure.

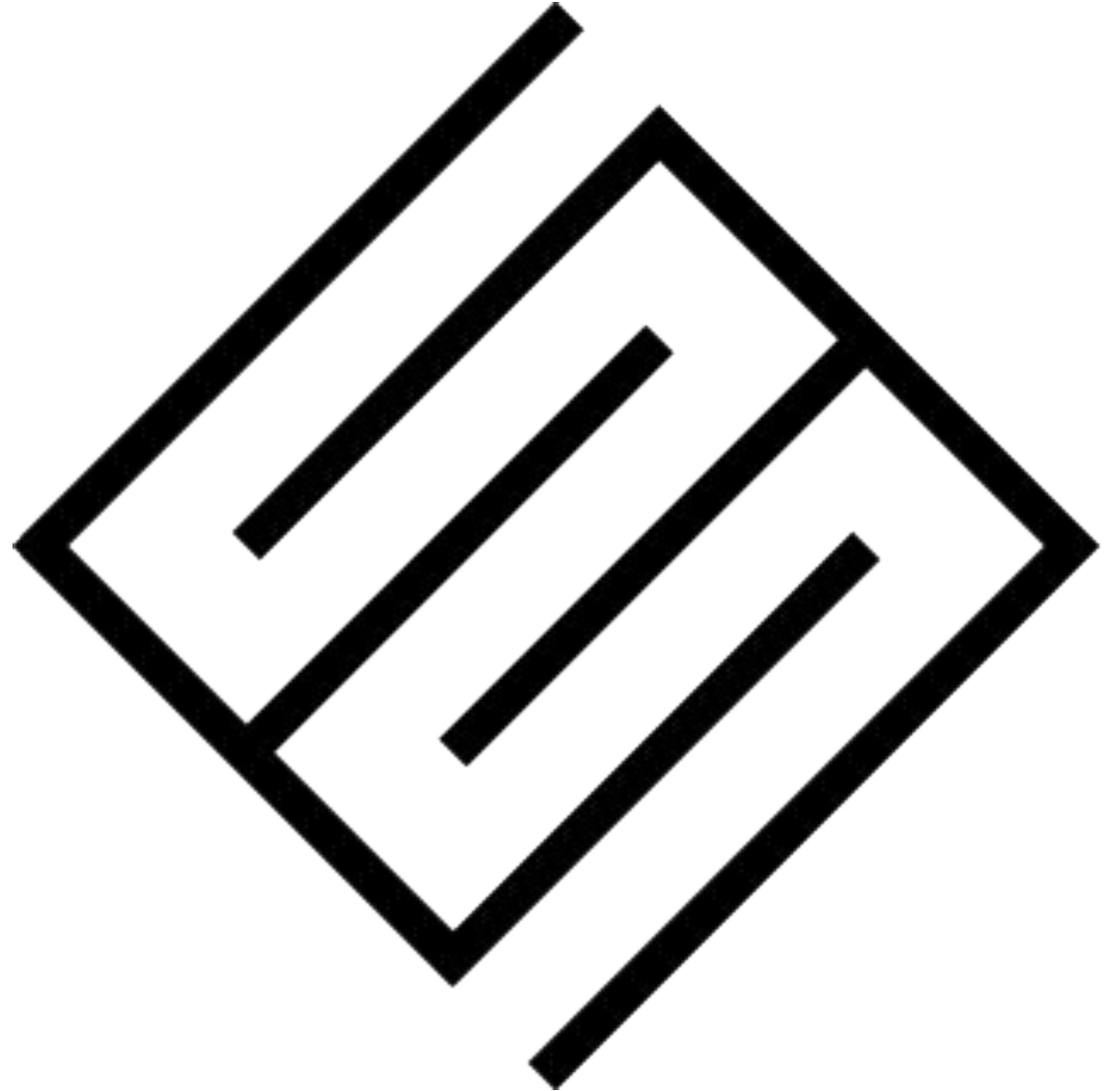
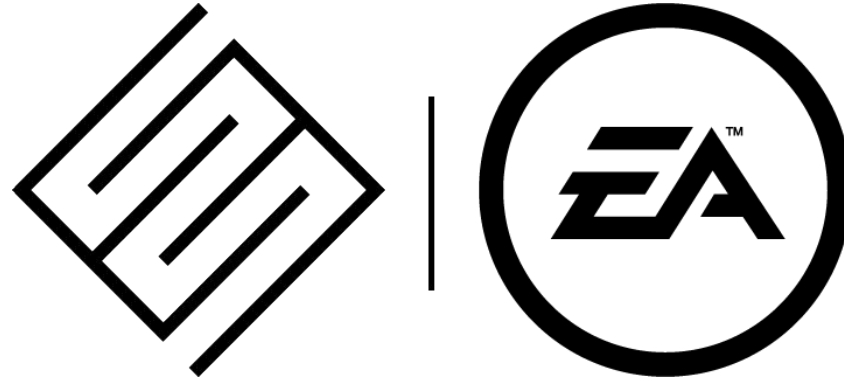


FIGURE WITH CAPTION AT THE BOTTOM



- This is the caption of the figure



seed.ea.com

STOCKHOLM – MONTRÉAL – LOS ANGELES – SAN FRANCISCO
LONDON – VANCOUVER – REMOTE