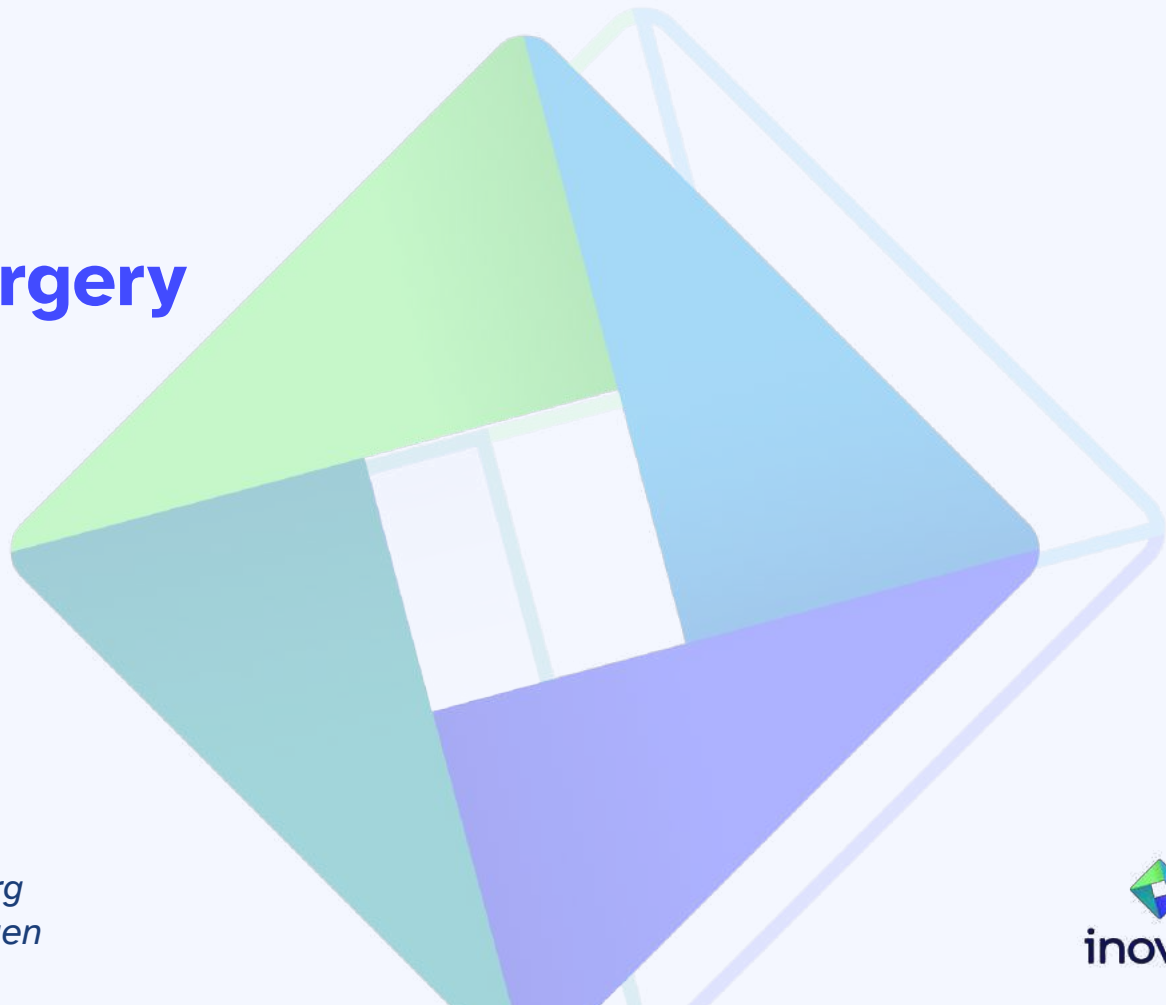


Master of Art Forgery

Developing a Generative AI Game
Meetup Erlangen - 09.11.23

Team inovex

*Karlsruhe · Köln · München · Hamburg
Berlin · Stuttgart · Pforzheim · Erlangen*



Pascal Fecht



Software and Machine Learning Engineer

- at inovex since 2016
- Software Engineering and NLP



Pascal Fecht



@pfecht

Agenda

- Project background and idea
- Demo
- Technical details



“Parrot”: What happened before?

- Student project
- Started in 2020
- Showcase for language models (BERT, ...)



app.inovex.ai



A screenshot of the Parrot application interface. The top navigation bar is dark blue with the Parrot logo and the text "PARROT". To the right of the logo are the links "TEXT COMPLETION" and "PROMPT ENGINEERING". The main content area is white and titled "TEXT COMPLETION" with a sub-link "Advanced Playground". Below the title is a paragraph of text explaining text completion. A "Disclaimer" section follows, warning about sensitive or inappropriate language. The "Complete your sentence" section includes instructions and a text input field with a submit button. The "Generate multiple responses" section explains the probabilistic nature of the models and includes a text input field and a submit button.

March 2023: Product Discovery

What can we do to create an **interesting showcase** for large language models and Generative AI at inovex?

- **Product Discovery** Workshop
- Representatives from Sales, Domain Experts, ...

March 2023: Product Discovery



Vision

What is your motivation for creating the product?
Which positive change should it bring about?

Auf der genAI games Seite von inovex spielen täglich 10.000 begeisterte Menschen AI games und lernen dadurch inovex kennen. Glücklicherweise haben unsere Studierenden mega Spaß daran, AI games zu erstellen und die User Zahl zu vergrößern. Über den "Advanced Weg" generieren wir neue Kontakte.

inovex zur ai company machen

Leute neugierig machen (Kunden, Bewerber, etc.)

Kundenneuererung durch Ansprache verschiedener Zielgruppen

Menschen (i.a. Entscheider oder potentielle Mitarbeiter) lernen inovex über das Spiel kennen, setzen sich danach mit inovex auseinander und beauftragen uns / bewerben sich weil sie begeistert sind.

Spielerisch Kompetenz von inovex im Bereich genAI zeigen

Kompetenz zeigen



Target Group

Which market or market segment does the product address?
Who are the target customers and users?

Studierenden

Spielbegeisterte Menschen

Technikgeisterte Menschen

Advanced User ermöglichen im Kontakt ermöglichen

potentielle Neukundinnen

Bestandskunden

potentielle neue Kolleg:innen

Neugierde

Easy access to a complicated topic (generativeAI)

Kreatives Spielen - "prompt engineering as a game"

Spaß

Wettbewerb / Austausch / Sozialer Kontakt



Needs

Which problem does the product solve?
What benefit does it provide?



Product

What product is it?
What makes it stand out?
Is it feasible to develop the product?

Es ist ein Spiel (z.g. generiere mit Prompt ein Bild, das nahe dem ist, was auf einem vorgegebenen Bild zu sehen ist)

What makes it stand out -> easy to understand and use

Add tutorial video which explains where the AI is doing what

VERY easy to use

AI und UI sind beeindruckend! (Accessibility)

Cooler Easter eggs -> (infos zu GenAI oder inovex, oder cooler)

Simple / Advanced

Komponente (LM, Text-To-Image) über die Zeit austauschbar (mit dem State-Of-The-Art gehen)

App vs Web App?



Business Goals

How is the product going to benefit the company?
What are the business goals?

Reichweite erhöhen (Top of the funnel)

Kompetenzen "beweisen" - v.a. UX/CX und AI

Cross-Team (LoB) Zusammenarbeit

Studierende & MA haben Umgang mit State of the Art Technologien

Neukundengewinnung (im besten Fall) oder neue Projekte mit Bestandskunden

Mitarbeiterakquise

Wissensaufbau

⇒ Let's develop a Generative AI game



April 2023: Kick-Off and Design Thinking Workshop

But... What should the game actually be?

⇒ **Design Thinking** Workshop

- Originally by IBM to develop user-centric products
- Participants: Student team (the actual developers)
- We did a “light-weight” version

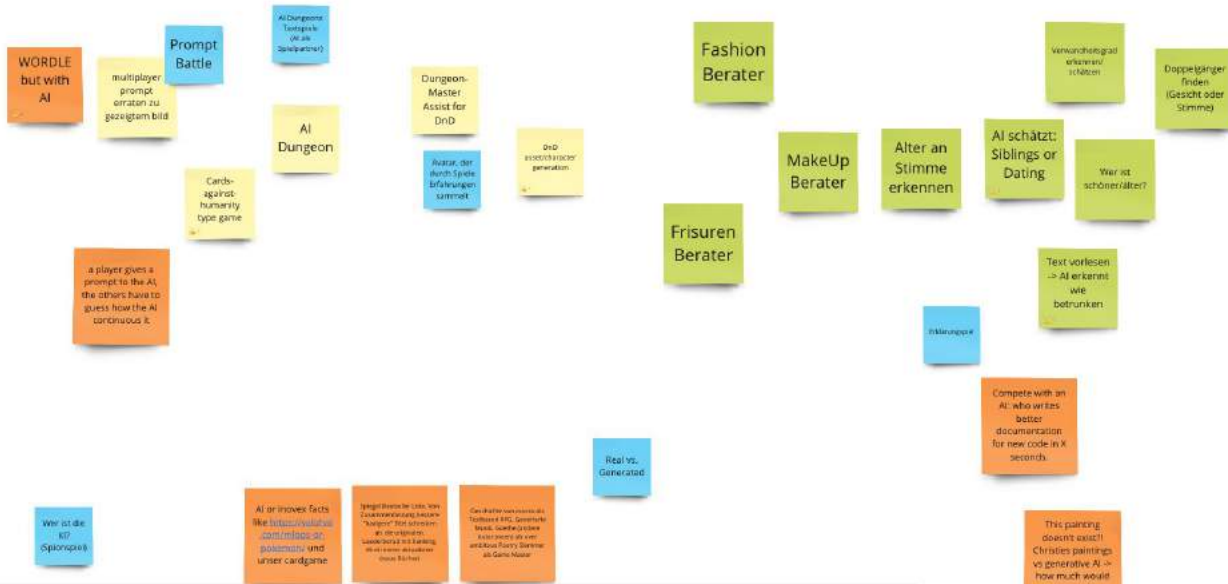
April 2023: Kick-Off and Design Thinking Workshop

Ideate



3

ALL OF INDIVIDUAL TOP 10'S

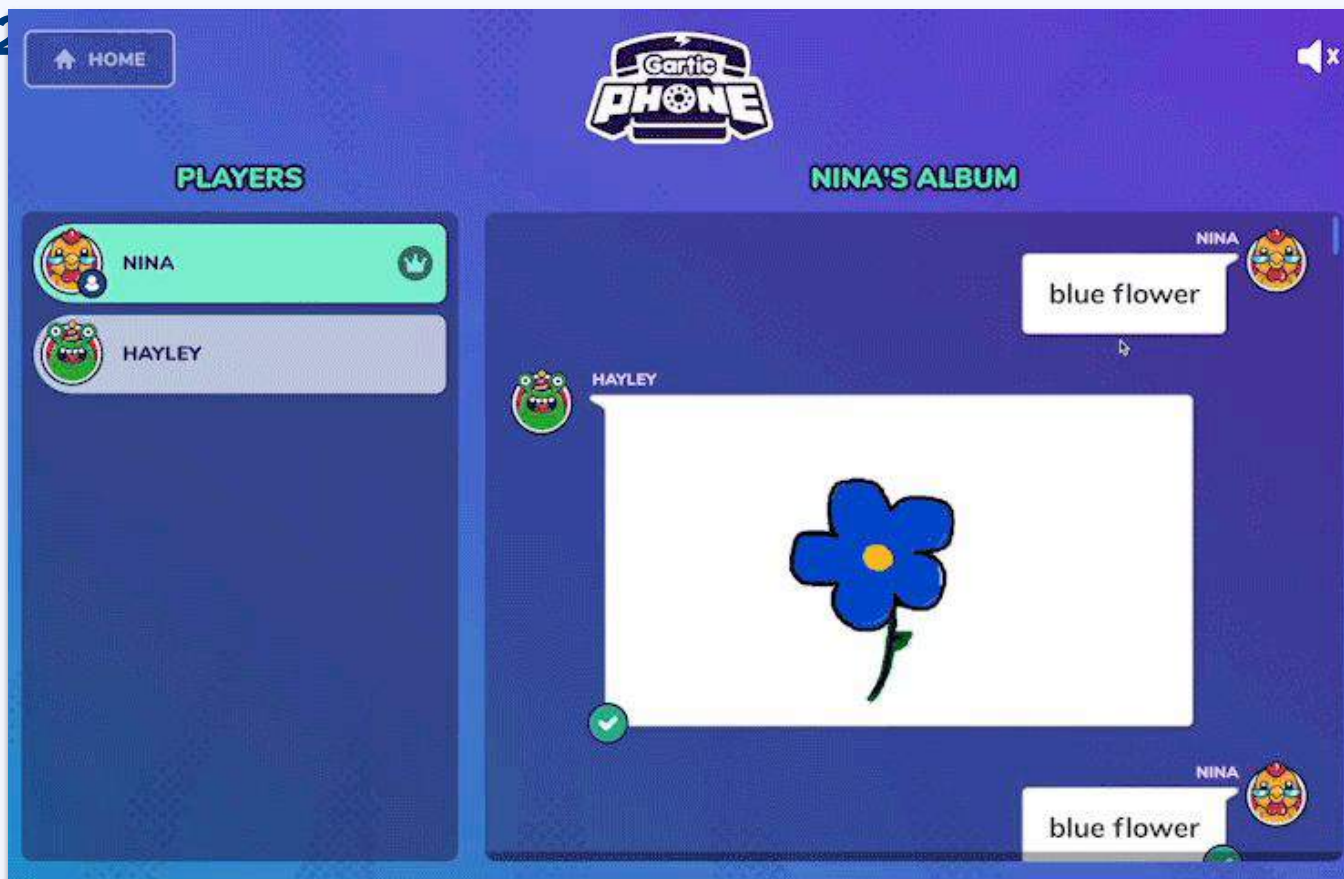


April 2023: Kick-Off and Design Thinking Workshop

How is your concept called? GarlexPhonivation (WIP)		What problem does your concept solve? <p>Showing off the AI competence of iPhones in an effective way by making interest in a wider audience.</p>
Who are key stakeholders? <p>AI interested (potential) customers potential employees sales and marketing team</p>	What is the big idea? <p>Gartic Phone but images and prompts are AI generated.</p>	
Illustrate how it works! <p>Single player Leaderboard</p> <p>4 player Multi Player Lobbying</p> <p>Online Multi 1000 player</p>		
Why might it fail? <p>Multiplayer is challenging Resource heavy</p>	What should we prototype and test? <p>Try to replay in real life Voting</p>	How might we measure success? <p>User numbers User feedback</p>

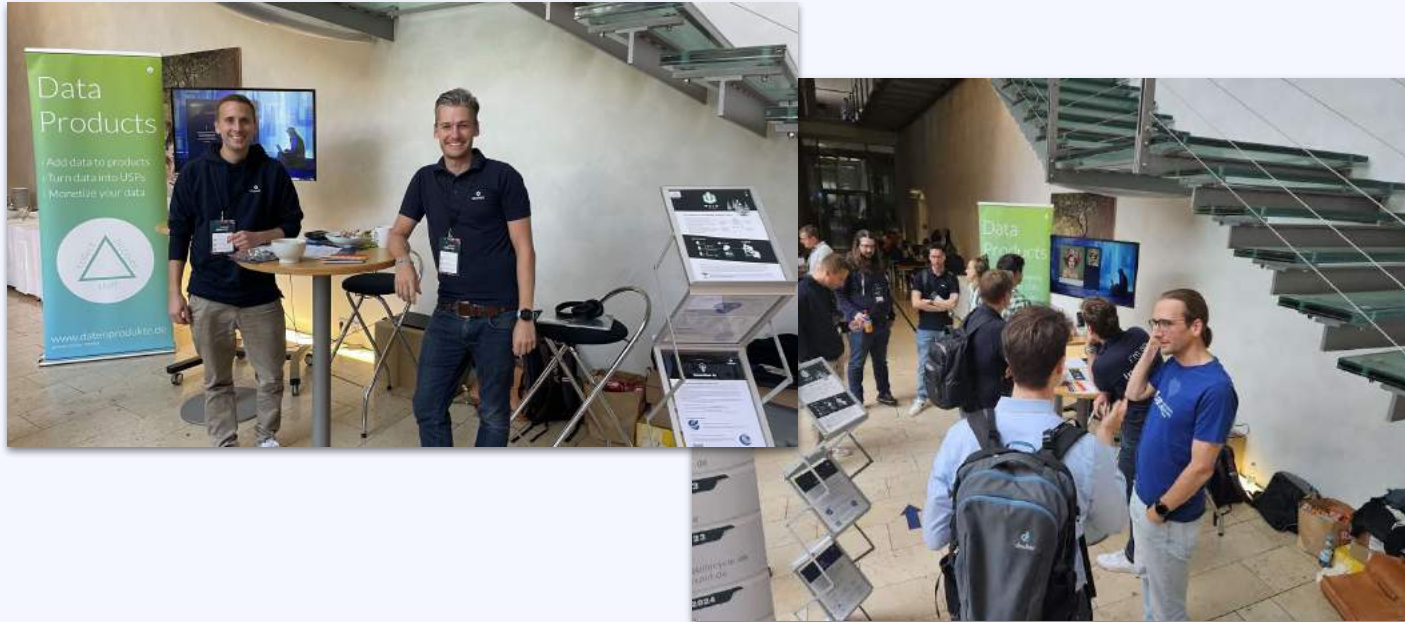
Concept Poster

April 2



October 2023: Release Version 0.1 🚀

12.10: Release and stress test at data2day 2023 conference



DEMO!



**So, what's the game
all about?**

Architecture

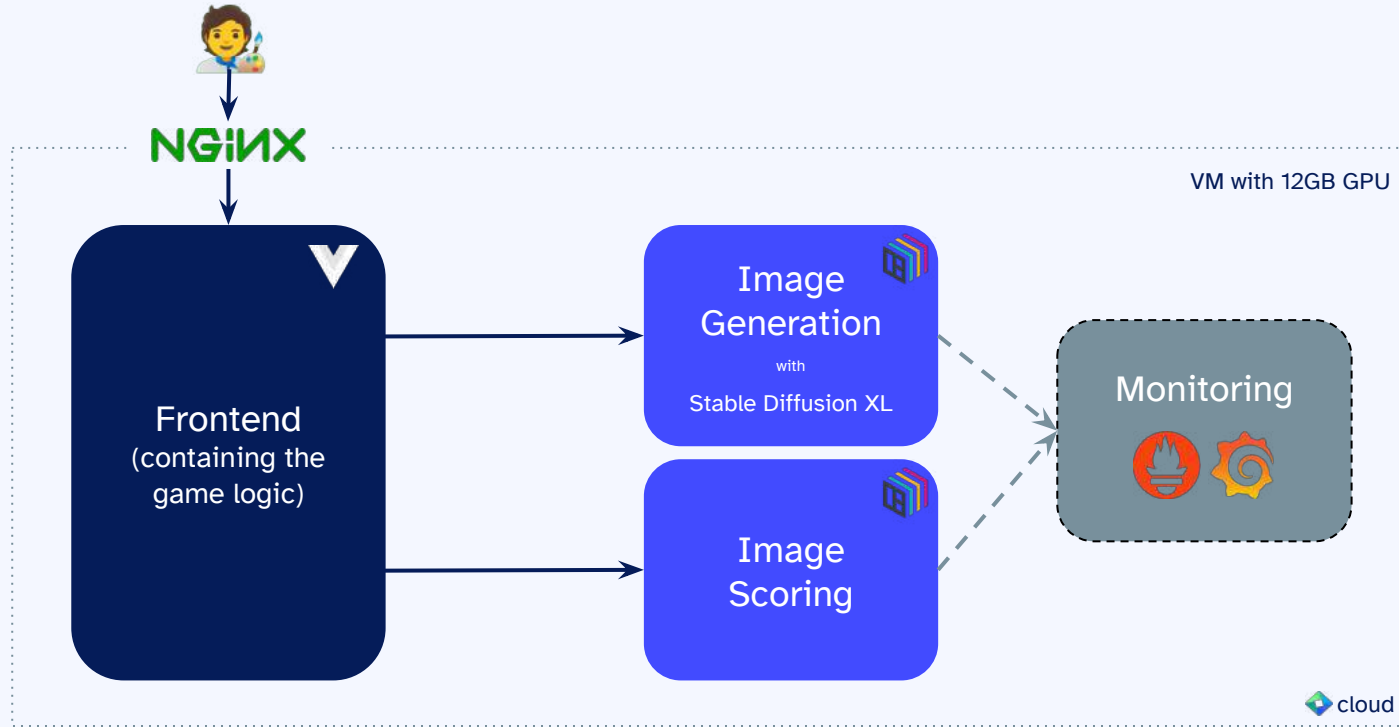


Image Generation Service

[stabilityai/stable-diffusion-xl-base-1.0](https://huggingface.co/stabilityai/stable-diffusion-xl-base-1.0) from huggingface

- Off-the-Shelf AI model by Stability AI
- Open Source
- Content-Review-Policies
 - Bad word filter (dictionary-based)
 - Stable Diffusion safety filter

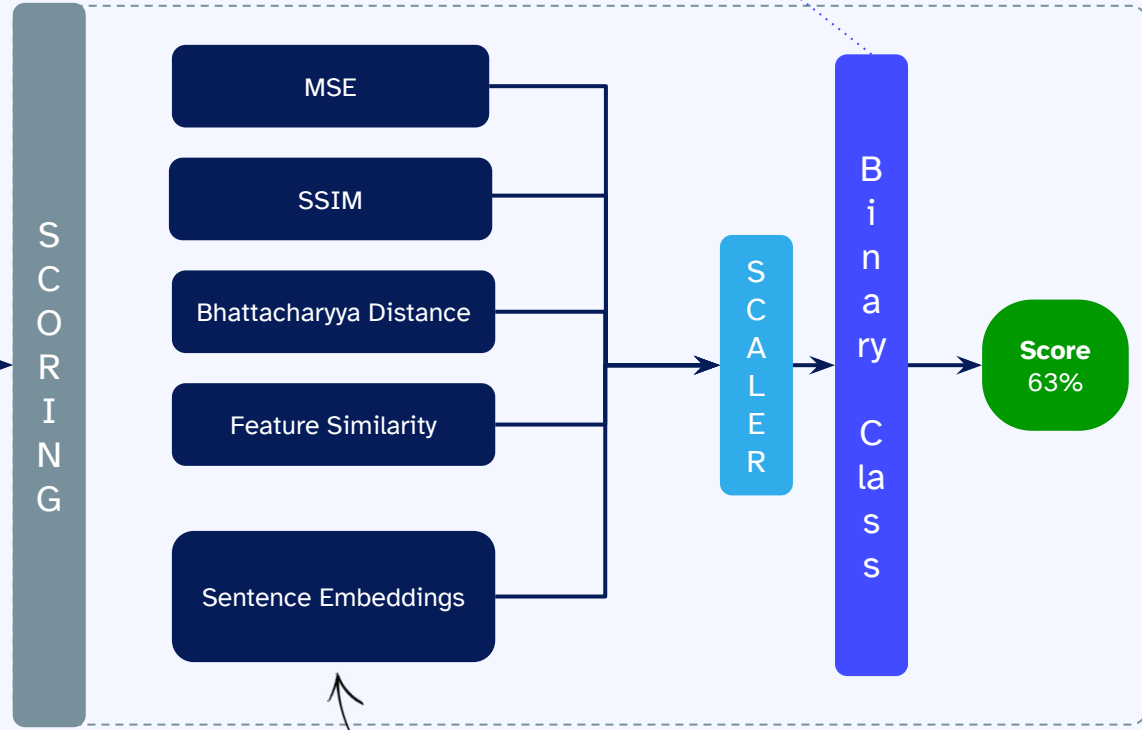
Scoring Service



Target Image
Two pirates playing chess on a boat



User Image
Pirates playing chess



Trained on our own dataset

SCORING

MSE

SSIM

Bhattacharyya Distance

Feature Similarity

Sentence Embeddings

SCALER

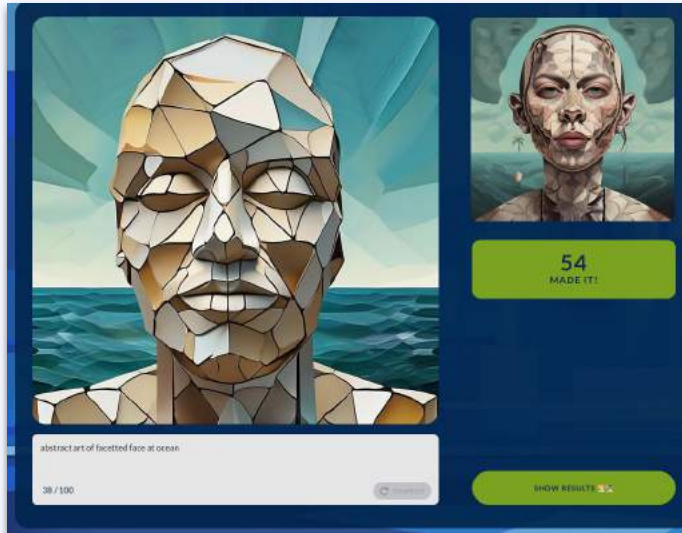
Binary Class

Score
63%

based on user & target prompt

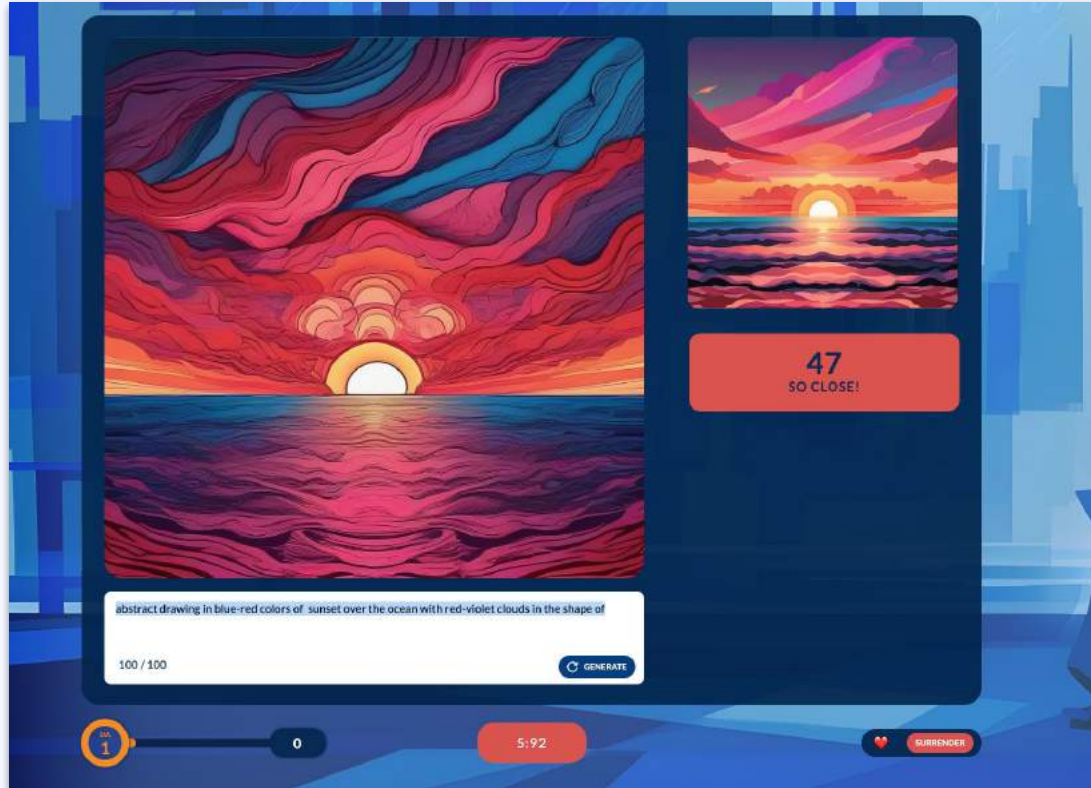


Challenge: Limitations of scoring metrics



- Metrics mainly focus on image features, not their meaning
- **Sentence Embeddings** helped to address this

Challenge: “Score is not fair 😞”



It's a game, we don't have to be fair 😞

Challenge: “Score is not fair 😞”

Experiment: Multi-class instead of binary problem

- However, multi-class classification did perform worse
- More important to let users survive than to match the target score
- **But:** Adding neutral images to survived class achieved better results than before with binary dataset.

class	meaning	score
0	tot. disagree	0
1	disagree	0.25
2	neutral	0.5
3	agree	0.75
4	totally agree	1

Challenge: Scoring overfitted on current images

- Dataset only contains ~30 different target images

Thus, the model:

- does not generalize on other images
- is very sensitive to small changes
- makes it difficult to add new images

⇒ **Possible solution:** Larger and more diverse dataset

What's next?

Check out our social media accounts to get notified



⇒ **Public** Go Live in December 2023

and then

- Develop a more sophisticated scoring (DL approach)
- Gather more user feedback
- Add more features to
 - give users more feedback
 - explicitly **learn Prompt Engineering**
 - add multiplayer functionality (?)

Thank you!



Pascal Fecht
Software Developer

pascal.fecht@inovex.de
01733181068

