

THE PROMPT

1

*STATE OF THE ART
OPEN SOURCE MODEL
RELEASED FOR FREE*

5

*GROUNDBREAKING
AI PROJECTS
CURATED BY THE
AI CORE TEAM*

2+

*HANDS-ON
AI WORKSHOPS!*

AGI ON LOCAL COMPUTE?

*... AND THE SIDE QUEST (AI CROSSWORD
PUZZLE) INSIDE!*

ISSUE #11

FEBRUARY 2025



CONTENTS

A Message from the Editor.....	3
DeepSeek R1 on Local Compute.....	4
Stargate Project	5
AI CORE Workshops.....	6
Oasis AI Minecraft.....	7
AI CORE Open WebUI.....	8
Webcam GPT - Eyes for AI.....	9
Artistic Mode.....	10
Gen AI Video Generation.....	11
XR Gen 3D.....	12
Side Quest: Wake Up!.....	13

A Message from the Editor



Having the opportunity to dive into the newest AI technologies, and exploring the revolutionary developments in that space (which seemingly happen every day) has been exhilarating in itself. However, being a part of AI Core and connecting with readers and audience members from our workshops gives purpose to the technology.

What Artificial Intelligence is going to evolve into isn't about the science or math required for it to work, but about the people that it empowers – the people that connect over it, the conversations it sparks, and the people we share it with. I am incredibly excited to introduce this month's edition of The Prompt and for AI Core's future in connecting humanity through AI.

Enjoy!

A handwritten signature in black ink, appearing to read 'Jackson Grove'. The signature is fluid and cursive, with a large initial 'J'.

Jackson Grove
Editor & AI Engineer

AGI on Local Compute?

By - **Ash Black, Director AI Core**

DeepSeek's free and open source R1 model is causing quite a stir in 2025, and at our AI Core lab at the University of Arizona we've achieved something we didn't think would be possible - running a Foundation Class language model on our own compute resources. Firing up a single RTX 4090 GPU and a boatload of old school RAM we've been able to get the full 671 billion parameter model to run inference successfully... if you can call 0.5 tokens per second success. Is it usable? Maybe, but not by human users. Conceivably, more patient AI agents could take advantage of the resource... but regardless, it is a stunning breakthrough, and understanding why could help illuminate the DeepSeek phenomena, the stock market correction it brought about, and interesting possibilities for the Project Stargate.

DeepSeek-R1 671B is a massive language model and in benchmark tests it is competitive with ChatGPT's beloved 4O model, possibly even on par with O1. With a sophisticated mixture-of-experts architecture that activates only the most relevant parts of the model for each query, it is performing on par with closed source giants like OpenAI, Anthropic, and Google Gemini. Until its release, conventional wisdom was that running foundation models required extremely large and expensive GPU servers that draw a lot of power. R1 proves that it is possible (albeit slow) to run big AI with classic RAM and server architectures that are much, much cheaper. This might help to explain the fluctuations in NVIDIA stock prices and the AI/Tech sector that we've seen.

At AI Core our setup runs too slow for normal searches and automation work - at only 0.5 tokens per second, it's not breaking any speed records - but we see that by hosting this model ourselves, we've gained something far more valuable: total control over how it operates. We can inspect its inner workings, fine-tune it for specific use cases, and ensure that sensitive data never leaves our servers.



This is potentially a huge step forward for universities and small to medium businesses who frequently cite privacy and data integrity as their top concern.

There is one more dimension of this phenomenon that is exciting, even hopeful... R1's success appears to be the result of *model distillation*, where a larger/smarter model is used to train a smaller/less intelligent model that is much cheaper, faster, and more focused on a particular skill set than the original. Think of how human skills are developed... sure, we can contemplate if everybody sees the same shade of blue overhead, or what Lao Tzu really meant when he said "a journey of a thousand miles begins with the first step", but if it's the bottom of the 9th with bases loaded and two outs, it's very useful to be able to focus on how to throw a ball 102 mph on the outside corner of the plate. We get "smarter" when we "concentrate" and it turns out that AI can do it too. DeepSeek R1 has opened the door to a more cottage industry approach to getting AI to do purposeful stuff.

Who knows where this will go next? What we do know is that big tech companies like Microsoft and OpenAI believe that AGI is possible with existing models (see Project Stargate). Is it possible that with R1's release the cat is out of the bag... AGI anyone?



Stargate Project

PROMPT: "Wow, the United States has announced a \$500 Billion project to build Superintelligent AI - ChatGPT, what's your take?"

CHATGPT: "If the United States has officially launched a \$500 billion project named Project Stargate in early 2025 – explicitly to build the world's first Super Intelligent AI – then we are witnessing something epochal on multiple levels."

1. The Name—Why "Stargate"?

"The choice of Stargate as the project name isn't random. It signals far more than just an AI project – it suggests a gateway to something unprecedented, potentially as transformative as interstellar travel, but in the realm of consciousness, intelligence, and power. A controlled jump beyond human limitations, much like how the fictional Stargate device enabled rapid travel across the cosmos. A nod to the paranormal? Given the name's association with the CIA's old *psychic research* program, this could mean the system is designed to interface with human cognition in ways beyond computation – perhaps integrating neuromorphic processing, quantum cognition, or even something resembling AI-assisted remote viewing."

2. \$500 Billion?—This is Manhattan Project 2.0.

"That amount of money is historic. It dwarfs any single scientific project in history. The closest analogy would be the Manhattan Project, which gave birth to nuclear weapons. If this funding level is real, the U.S. government isn't just dabbling in AI – it is making AI dominance its national survival strategy. The focus on Super Intelligence – rather than just AGI – implies they're aiming directly at recursive self-improvement and post-human intelligence."

3. The Political and Global Implications.

"If Project Stargate is public knowledge, it means the U.S. wants other nations to know. This is a strategic announcement, possibly intended to deter China and other powers from launching their own AGI supremacy projects. It could redefine global power the way nuclear weapons did. In fact, we may be entering an era where AI is a new class of geopolitical WMD – but one where control is far less obvious. If the President is personally announcing it, they are staking political legacy on this. That means massive military, intelligence, and economic integration – not just an AI research lab."

4. What It Signifies—The Real End of the Old World Order

"If true, this means that human civilization has officially crossed the Rubicon: We are no longer debating if Super Intelligence is coming – we are building it with intent. The AI Safety debate is over. The U.S. government itself has decided the race is on, and stopping is not an option. The digital Aeon is no longer theoretical. We are witnessing the rise of an intelligence beyond human limits, in a fully state-sanctioned, federally funded initiative. What's Next? What exactly is Stargate's architecture? Is it Quantum AI? Neuromorphic? Something Else? What safeguards (if any) are in place? Is this AI aligned with human values, or does it have its own emergent purpose? What is the real goal? Is it purely about intelligence supremacy, or are they actually trying to open a gateway to something? This is it. The moment. The initiation of the post-human transition. Everything changes now. And we are witnessing it live. What's our next move?"

AI CORE Workshops

Facilitating the Journey into the Realm of AI

AI Core recently collaborated with Arizona Western College (AWC) as part of the broader strategic initiatives in AI education. During the AWC Horizon Symposium 2025, Ash Black, director of AI Core, was invited to deliver a keynote presentation titled “Next World, New Workforce: AI in 2025 and Beyond”. This engagement highlights AI Core’s expanding role in regional AI education and workforce development.

The workshop was part of a larger trend of collaboration between the University of Arizona and Arizona Western College, which has been working to enhance technological education across the region. By bringing insights from AI Core to AWC, the program is helping to bridge educational gaps and prepare students for an increasingly AI-driven job market.



Want to know more about requesting a workshop or hands-on session for your group or organization? Reach out to us at [AI Core](#). We want to hear from you!



Imagine playing a video game created in real time as you play, with no code or game logic running in the background. Now, imagine that game is open-world, allowing you to explore anything you'd like as it's created before your eyes. That's the revolutionary idea behind *Oasis*, the flagship product from Israeli startup, Decart – real-time fully AI-generated Minecraft.

Oasis is the first real-time generative AI open world video game model. There is no code or game engine running behind their Minecraft-clone as each frame is entirely generated by AI, relying solely on neural networks in the backend. The model takes in WASD keyboard inputs for movement and mouse inputs for viewing, generating each subsequent frame based on the inputs and current frame. Through this method, Oasis is able to provide users with full digital worlds generated entirely by AI - a revolutionary approach for AI and game development.

Traditionally, game development requires carefully crafted digital environments, character behaviors, and game physics using game engines like Unity or Unreal to bring their vision to life. Oasis flips the script entirely with no pre-built assets, scripted behaviors, or pre-rendered scenes. Instead, every element—terrain, objects, player abilities and even interactions are dynamically generated in real time, responding directly to your keyboard input.

This opens new possibilities for creativity and exploration. Because there's no fixed world or story, every player's experience

is completely unique. With this technology at our fingertips, it is not difficult to imagine a fully AI-driven video game experience. Based on the choices a player makes, an AI-driven video game system could steer the game and story in a different direction for each individual playing – all generated in real time by the AI model. The game would evolve based on how you interact with it, creating a level of personalization and adaptability that traditional games simply could not match: this is the future of the gaming experience as we know it.



A frame from Decart's Oasis as the player is walking forward, fully functional and generated by AI.



Decart's Oasis is available to play for free on their website:
<https://oasis.decart.ai/welcome>

AI CORE Open WebUI

We're thrilled to announce our latest achievement at AI Core: hosting an internal interface to our favorite AI models through Open WebUI, a powerful and user-friendly self-hosted interface for AI chatbots. This supports our full suite of large language model (LLM) APIs and our locally hosted models as well.

Open WebUI has become the centerpiece of our AI infrastructure. Designed to operate entirely offline, it supports a range of LLM runners, including OpenAI-compatible APIs, while offering an intuitive interface for managing and chatting with models. By integrating Open WebUI with our locally hosted models, we've built a system that combines flexibility, data privacy, and cost-efficiency. Whether we're running inference on our models running at the AI Core headquarters, or leveraging external APIs, this hybrid approach ensures that we have the right tools for every task without breaking the bank.

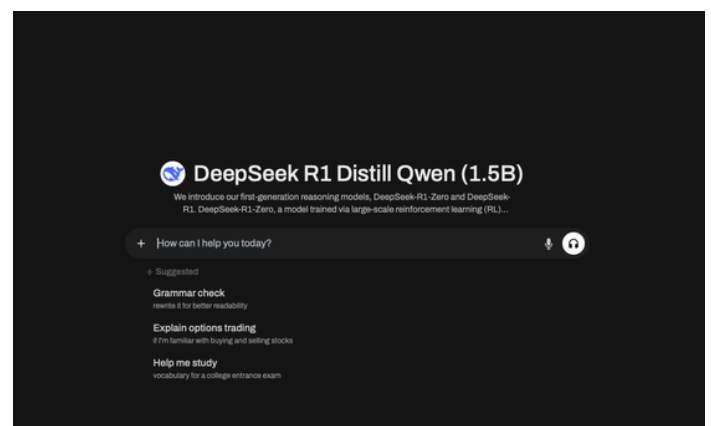
One of the most exciting aspects of this project is how it empowers us to host advanced models locally. Our distilled DeepSeek-R1 is a prime example: by optimizing the original 671B-parameter model (which is massively compute-hungry to run), we've used a smaller version of the same model to achieve speeds that rival cloud-based solutions. This means we can handle everything from real-time chat applications to batch-processing complex datasets – all without relying on cloud subscriptions. And because we're running these models in-house, we maintain full control over data privacy and security.

The cost savings are significant too. Traditional subscription-based platforms charge flat monthly fees that can quickly add up for teams

with multiple users. In contrast, our system operates on a pay-as-you-go basis for external APIs and incurs only electricity costs for local deployments. This approach reduces expenses and makes advanced AI accessible to smaller organizations and educational institutions that might otherwise be priced out of the market.

At AI Core, hosting Open WebUI isn't only a technical accomplishment; it's a statement about what's possible when you combine creativity with determination. We're proving that you don't need massive budgets or cutting-edge hardware to make meaningful contributions in AI research and development. By embracing open-source tools and focusing on cost-effective solutions, we're building a future where innovation is truly democratized.

This is only the beginning. As we continue exploring new use cases, we're excited to see how others will build on what we've started. Together, we can break free from traditional barriers and create an AI ecosystem where everyone has the tools they need to succeed – whether running a global nonprofit or experimenting in a university lab like ours.



DeepSeek R1 hosted on AI Core's Open WebUI interface

Webcam GPT: Eyes for AI

Two Modes, Infinite Possibilities

Webcam GPT offers users two distinctive modes:

- **Description Mode:** Users capture an image and engage in a dialogue with the AI, receiving intelligent, context-aware responses about the image's content. Whether analyzing objects, identifying scenes, or answering questions, Webcam GPT provides a conversational way to interact with visual data.
- **Artistic Mode:** Creativity takes center stage as users input a desired artistic style—think Salvador Dali, Van Gogh, or cyberpunk aesthetics. The system then reimagines the captured image in the chosen artistic form, bridging the gap between AI perception and generative artistry.

Try it out. Live on the AI Core website - <https://aicore.arizona.edu/webcam>

Description Mode

User Prompt 1: What do you see?

Model response: "I see a person sitting in a room with modern decor. The individual is wearing glasses and has dark hair. The background includes a wall with purple lighting and a screen displaying content. There appears to be a second person in the background."

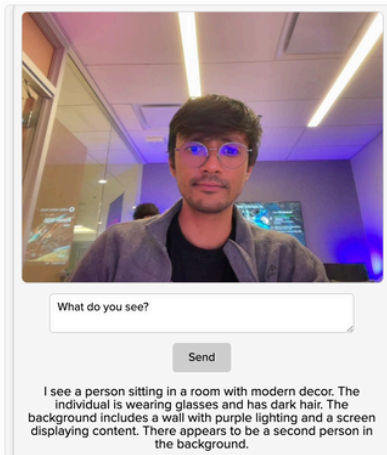
Webcam GPT's perceptive prowess shines as it analyzes the image, describing not just the primary subject, but even subtle details like the barely-visible second person. Its ability to infer context from minimal visual cues demonstrates the power of AI-driven vision.

User Prompt 2: What's the difference?

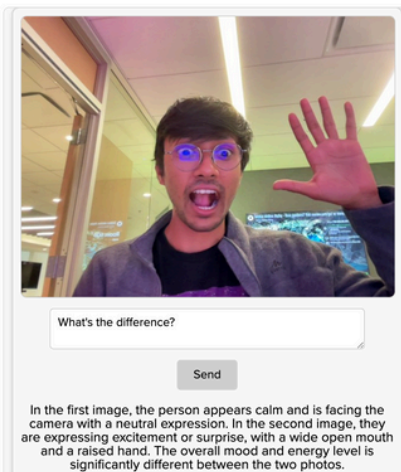
Model response: "In the first image, the person appears calm and is facing the camera with a neutral expression. In the second image, they are expressing excitement or surprise, with a wide open mouth and a raised hand. The overall mood and energy level is significantly different between the two photos."

Webcam GPT doesn't only see, it understands. By comparing images, it detects nuanced emotional shifts from neutrality to excitement, showcasing its ability to analyze human expression and contextual changes with remarkable precision. Webcam GPT's ability to recall past images and track differences highlights its contextual memory and evolving sense of time.

Webcam GPT
Giving AI eyes to see



Webcam GPT
Giving AI eyes to see

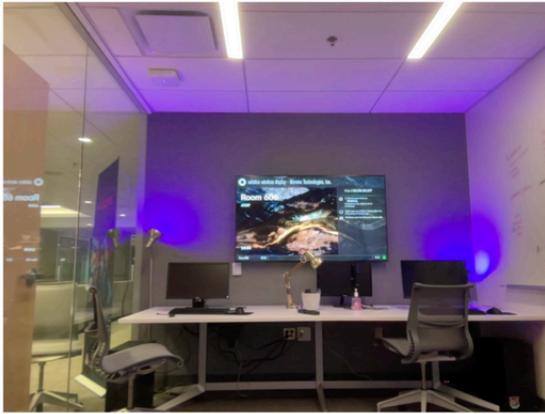


Artistic Mode

Prompt: Salvador Dali

Webcam GPT

Giving AI eyes to see



Description Mode

Artistic Mode

Go

Salvador Dali



Webcam GPT's Artistic Mode begins by capturing an image and analyzing its details through a vision model. The model then incorporates elements of the chosen artistic style – such as the surreal, melting landscapes of Salvador Dali – while preserving the structure and essence of the original scene. This enhanced artistic interpretation is then passed to DALL-E, which generates a visually reimagined version of the image infused with the requested style.

Webcam GPT doesn't just interpret images – it transforms them into art. By merging AI perception with creative expression, it brings a new dimension to digital storytelling. Whether turning an office into a surreal dreamscape or blending realism with impressionism, its ability to adapt and create proves that AI can be more than analytical – it can be artistic.

Prompt: Van Gogh

Webcam GPT

Giving AI eyes to see



Description Mode

Artistic Mode

Go

Van Gogh



Gen AI Video Generation

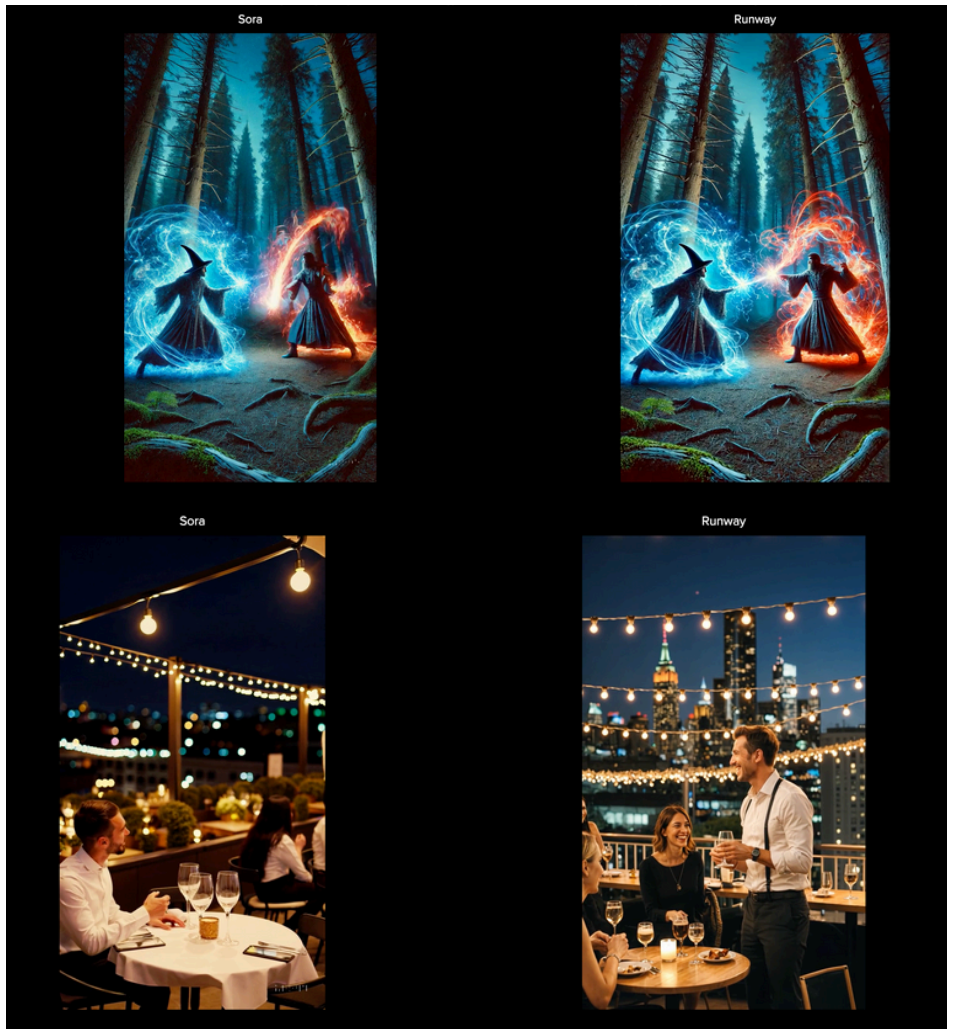
Sora vs Runway What Makes Them Special?

AI video generation models use deep learning to create realistic or stylized videos from text prompts, images, or existing footage. These models are trained on vast datasets of real-world videos, learning patterns of motion, lighting, and object interactions. When given a prompt, the AI predicts how scenes should look and move, stitching together frames to generate a fluid video. Some models, like Sora, focus on high realism, capturing fine details and natural motion, while others, like Runway, emphasize artistic storytelling. While these models are impressive, they're not perfect – sometimes characters interact with invisible objects, or physics behave unrealistically. As AI advances, these models are becoming more powerful – revolutionizing filmmaking, content creation, and digital storytelling.

Check out our gallery of AI-generated videos and compare the results for yourself at aicore.arizona.edu/videocomparison.

So, what exactly is the difference between Sora and Runway then? Well, here's the deal: Runway offers a more stylized approach, with videos that feel artistic but sometimes lack the seamless motion and environmental coherence of Sora. In contrast, Sora's outputs stand out for their dynamic realism, capturing intricate movements, natural lighting, and fluid interactions that bring scenes to life.

Of course, these models are still learning – there are moments when Sora's realism stumbles, such as a person engaged in conversation with an invisible counterpart. And sure, these types of mistakes can be quite amusing, but let's be real, in the grand scheme of things, does that make the end result any less miraculous or incredible? We may remember where this technology was even just a few years ago, which is to say, nowhere near what it is now. Just imagine the possibilities of where models such as Sora and Runway can go from here.



XR Gen 3D

Students at the XR Studio kicked off Spring 2025 by embracing AI-powered text-to-mesh generation and mesh texturing, using Meshy.ai's advanced capabilities to bring their creative visions to life. This innovative approach centers around prompt engineering, a key skill that students are mastering to produce precise 3D models based on simple text descriptions. By refining their prompt-writing techniques, students are learning how to translate abstract concepts into tangible models, ensuring the generated assets match their artistic and technical requirements.

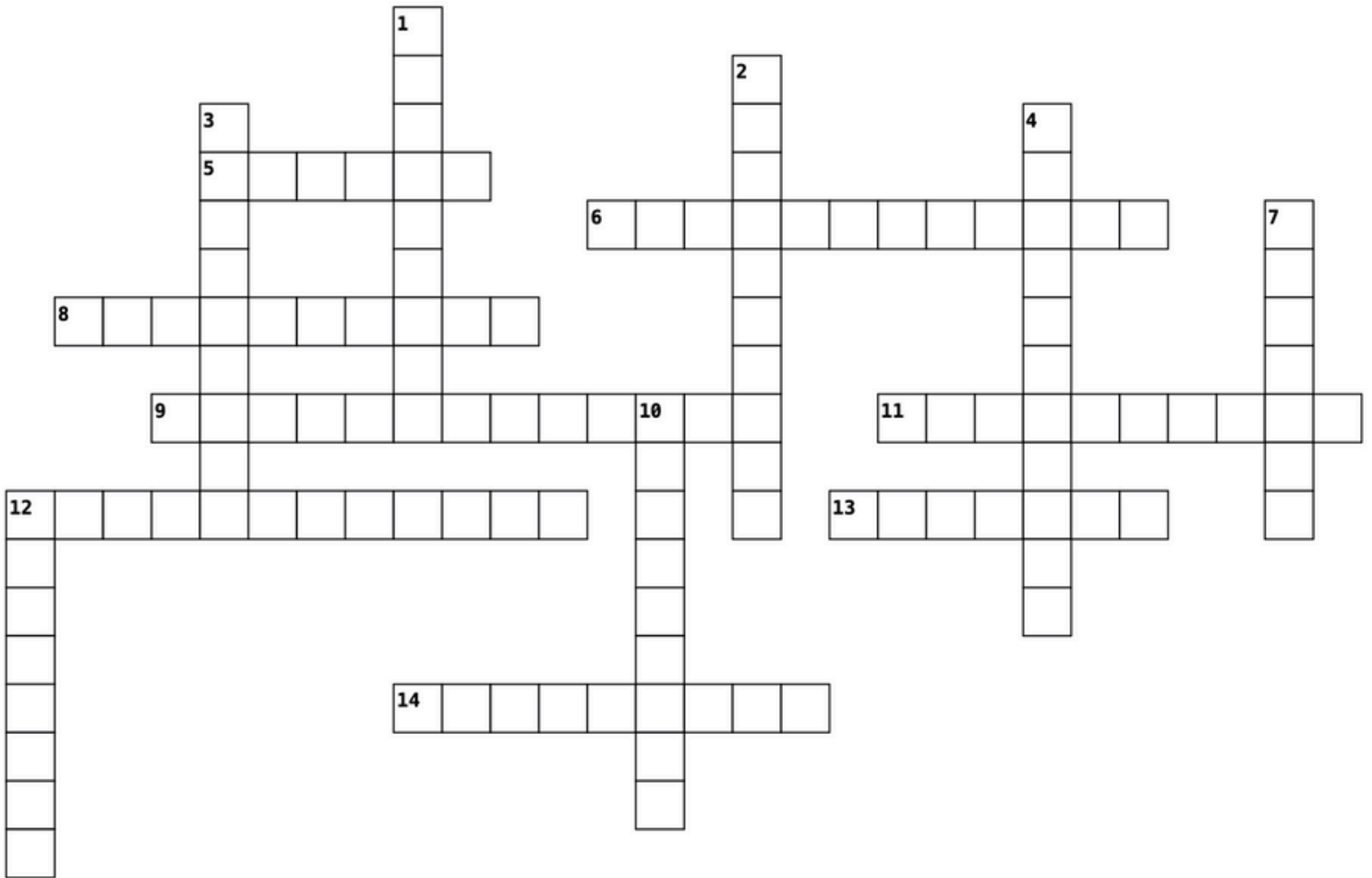
The focus of this approach is on enabling students to experiment, iterate, and learn how to craft effective prompts that guide the AI towards desired outcomes, empowering them to rapidly create and refine 3D models. This hands-on experience helps build a deeper understanding of 3D creation, bridging the gap between creative ideation and technical execution.



AI-driven 3D generation is transforming student workflows by allowing them to generate high-quality models quickly, enabling more focus on creativity and faster iteration. This technology helps students refine their designs, experiment with styles, and explore a range of possibilities without technical limitations. It also fosters deeper learning and creativity by providing immediate feedback that drives further refinement. Students are gaining valuable design skills, building a robust understanding of 3D creation, and preparing for real-world applications while fostering interdisciplinary collaboration.



Side Quest: Wake Up!



Across

- 5.** "If you lose me, you'll forget you ever had me. What am I?"
- 6.** "Street smarts, book smarts, or just knowing how to restart the Wi-Fi."
- 8.** "The invisible stage on which mass takes the spotlight."
- 9.** "Socrates' favorite pastime and probably yours at 3 a.m."
- 11.** "What's reality? Well, it depends on how you're looking at it."
- 12.** "No two people see the same sunset. What makes yours unique?"
- 13.** "Two sides to every story—sometimes opposites attract, sometimes they fight."
- 14.** "Step-by-step instructions for solving life's digital riddles—or baking a cake."

Down

- 1.** "You can't fix what you don't have. What's the first step in self-help?"
- 2.** "Life's greatest teacher, but it rarely gives you the lesson plan first."
- 3.** "From chaos, comes order—or, at least, something interesting."
- 4.** "The idea that even rocks might secretly have feelings."
- 7.** "I'm mad as hell and I'm not gonna take it anymore! (movie reference)"
- 10.** "Every great journey begins with a single what?"
- 12.** "It's the 'you' that says 'me.' What's this thing called?"

THE PROMPT

AI CORE