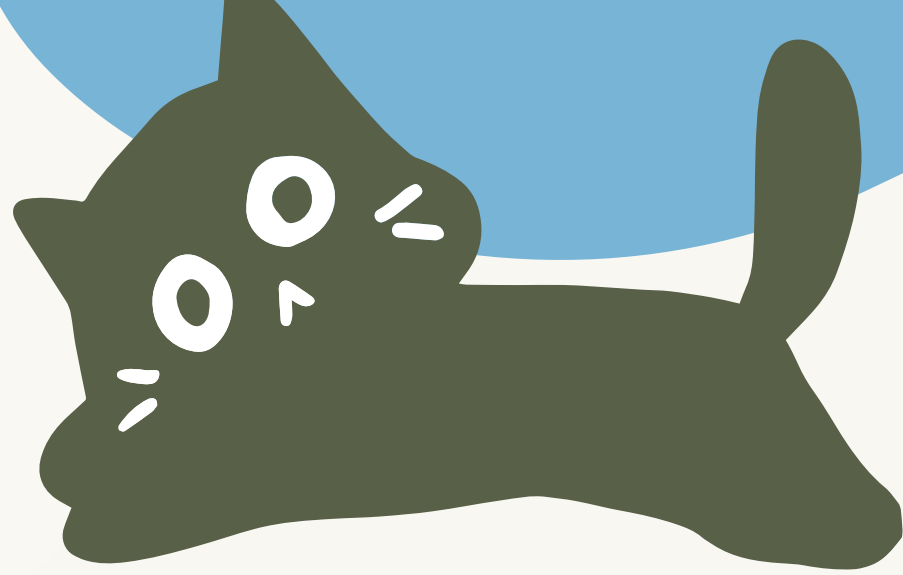


AGI, Me, Content Creation (individual research)

Nanxuan (Cherry) Zhao

@ CVPR 26 AC Workshop

(5min)



Quick Intro

Studied/Graduated @ Hong Kong

Assistant Professor @ University of Bath

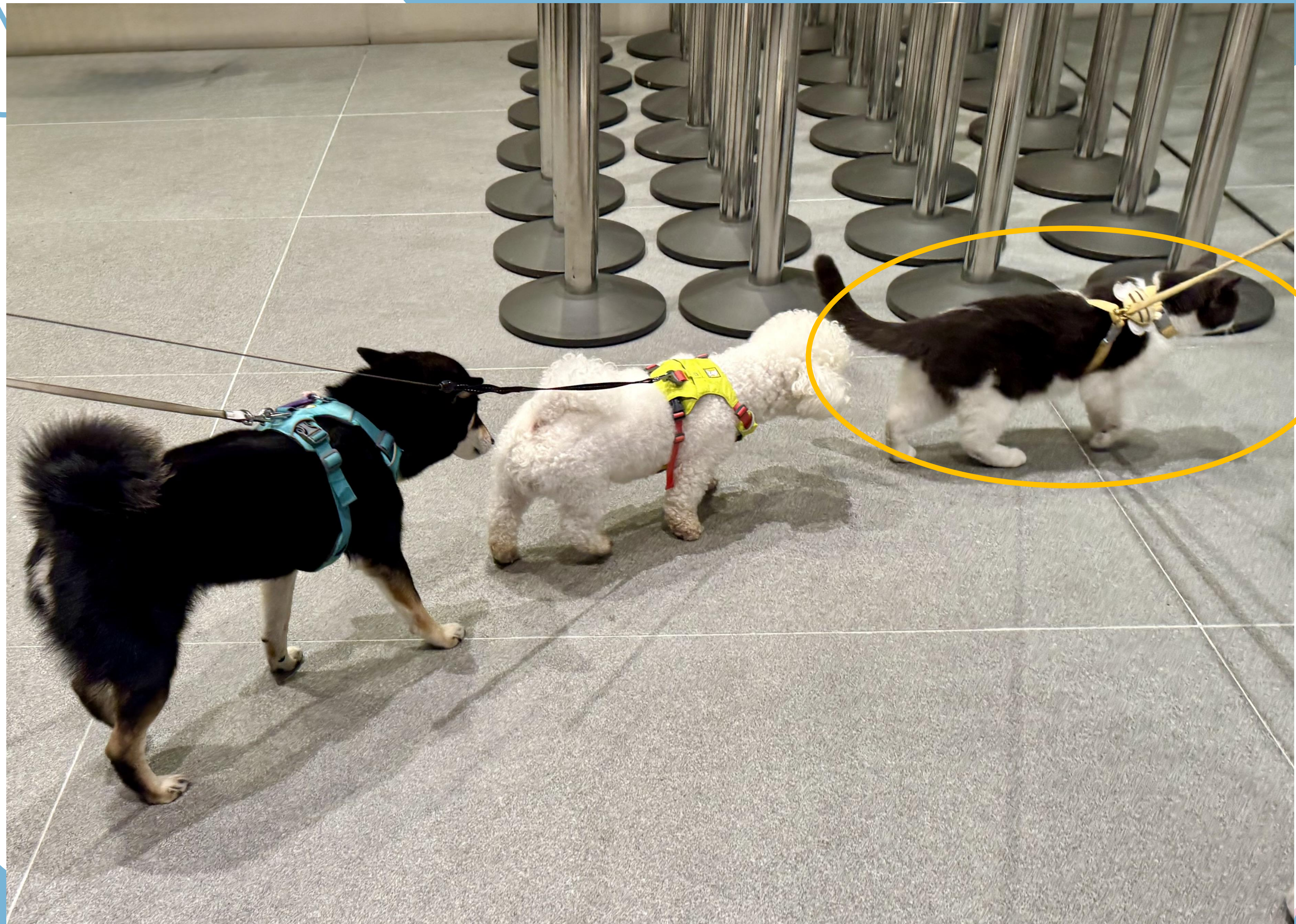
Research Scientist @ Adobe Research (based in the Bay area 😞😊)

CG X CV X HCI

Ski & Sn☺wboard, Drum, Wheel Pottery, ...

Doggie Cat





Controllable Content Creation (as always)

Human-Computer Interaction

Computer Vision

Trajectory First Frame Last Frame

A grid of images showing trajectories and first/last frames for various objects. The objects include a wallet, key ring, T-shirt, cap, tote bag, and phone case. The trajectories are shown as red arrows indicating movement paths. The first and last frames show the object's position at the start and end of the trajectory.

P1: Wallet P2: Key ring P3: T-shirt

P6: Cap P6: Tote bag P7: Phone case

Source image 3D edits Output

A 3D edit interface for a rabbit. It shows the source image, a 3D wireframe model of the rabbit, and the final output image. A slider labeled '3D edits' is visible between the wireframe and the output.

A movie poster for 'Ponyo' and a corresponding heatmap. The heatmap shows a color-coded representation of the image, with red indicating high intensity and blue indicating low intensity.

Chocolate cake The Statue of Liberty with the face of an owl

A spaceship flying in the sky A pig wearing a backpack

A grid of image-to-image translation examples. Each row shows an input image, a set of color-coded masks, and the resulting output image. The examples include a chocolate cake, the Statue of Liberty with an owl face, a spaceship, and a pig with a backpack.

Input image Supervised MoCo

A comparison of supervised and MoCo methods. The 'Input image' column shows a couple on a bench and a dog. The 'Supervised' column shows the result of supervised learning, and the 'MoCo' column shows the result of MoCo. The MoCo results show more accurate and detailed object segmentation.

Computer Graphics

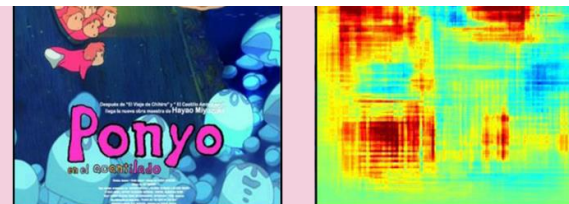
Controllable Content Creation

Human-Computer Interaction

Computer Vision

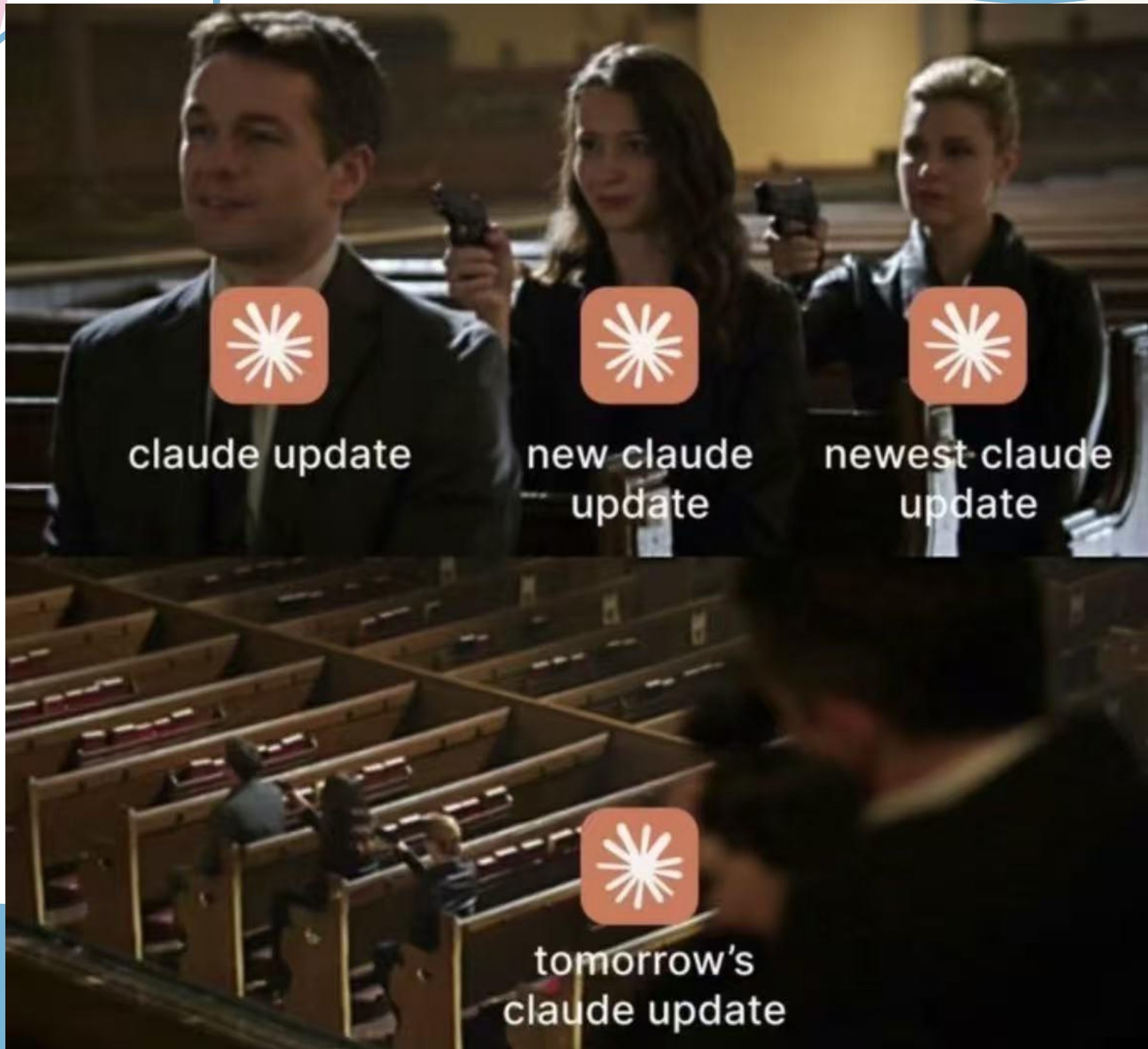
Trajectory First Frame Last Frame

What is my next research topic for the summer?



Computer Graphics





What is AGI?

"AGI" is a marketing buzzword."

-- Aaron Hertzmann (said in a 1-1 chat)

But one concrete technique is

LLM (& Agentic coding)



Text/Language is

one of the most powerful forms of content

“有志者事竟成”

“Where there is a will, there is a will”

“Querer es poder”

“ ... ”

As a result (two-week deep thinking) ...

I have manually typed out

an 8,000-word article in pure text

(with my native language + slightly bilingual)

Called void.pdf

(no AI involved)

By stream-of-consciousness writing

Void.pdf

AGI, 我, 和 Content Creation

科技的发展应该: **以人为本, 提高生产力, 提升人的生活质量, 最终造福全人类**

前言:

过去差不多一年时间, 我观察到一些群体性的焦虑。因为最近进入了一个工作的空窗期, 我开始思考暑假要做什么 project。由于这个过程过于奇幻, 不妨讲给大家听听。请注意 **这只是我作为一个个体非常片面的对世界非常主观的观察**。你不需要赞同也不需要去判断我说的真假, 权当一个故事看。**同时本文全部手搓, 没有任何 AI 参与**。为了保证描述精准性, **我将采用中文为主, 英文为辅的方式讲述, 如果你的翻译水平可以做到 native 且觉得这个故事很有意思值得分享, 也非常感谢你帮忙翻译**。你可以代入这个我为任何她, 你的女儿, 你的妻子, 你的母亲; 你也可以把我换成他, 这也不重要。如果你认为你认识 ta, 你爱 ta 的话请你保护 ta, 如果你不爱的话也请不要伤害 ta。**请你尊重 ta 去思考下面的故事。**

正文:

第一章

大概是从 2025 年 6 月 CVPR 开始, 因为大语言模型质的飞跃, TBD 的高调组建方式和夸张高薪 (>20 倍) 带来对市场的扰动, 以及各种内外部的流言, 公司内部出现了异常的情绪波动, 很多人都在这个时候开始向外探索。我也收集了一些信息, 但是因为不确定性太大且手头还有一些要紧的项目在做, 并未进行什么实质性的行动。时间很快来到 2025 年 12 月, 这个时候已经陆陆续续有多名同事离职, 公司内部的焦虑似乎在蔓延, 同时因为一定程度的 hiring freeze 和市场的混乱, 很多人还抱有观望态度。这个时候的大语言模型因为 Anthropic 的战略眼光, 已经被优化的非常好用了, 我这个有着超过 15 年丰富代码经验, 掌握多种语言和框架的老手都自愧不如。我正好手头项目 ship 结束且需要回去参加一个国际会议, 终于迎来一个 break。

尾言

我一气呵成, 你问我我的思维这么顺, 为什么会想这么久。我觉得并不重要, 你就当个故事听, 你如果能理解我讲的故事, 你大概率也能跟我一样有同样的担忧。我是一个资质平平的人, 没有办法教任何人做事。可是我好像不知道这个故事的结局应该怎么写。似乎有点儿两极分化呢。

这个世界之所以美好就是在于它的多样性。万物相生相克, 老祖宗真的好有智慧诶, 我现在才体验到。

科学家有发展技术的能力, 也有停止发展技术的能力, 我们有着这个世界上最强的武器和幸运的拥有了最强的大脑。人本身最大的问题之一就是身体跟不上脑子, 那我们发展一个脑子比身体智能更多更多的硅基生物的意义又是什么呢?

你问我你说了这么多, 你在表达什么:

我的猫猫平等的爱着每一个人, 我也是, 你也是, 希望 AGI 也是

AGI, 我, 和 Content Creation

科技的发展应该: **以人为本, 提高生产力, 提升人的生活质量, 最终造福全人类**

**My naïve philosophy about:
Model/Tech, Data, Myself, Education,
Society, Humanity, Product, Universe,
etc.**

, 你就当个故
是一个资质平
么写。似乎有

智慧诶, 我现

最强的武器和
我们发展一个

Void.pdf (kind of finishing the theory part)

AGI, 我, 和 Content Creation

科技的发展应该: **以人为本, 提高生产力, 提升人的生活质量, 最终造福全人类**

I found my answer to the question

And

I feel genuinely excited about AGI.



, 你就当个故
是一个资质平
么写。似乎有

智慧诶, 我现

最强的武器和
我们发展一个

After that ...

I have conducted two "User studies."

One with my "void.pdf" shared with my friends per request

Most are CS-related

(to test whether I really understand what content means)

The other is on general social media

Anyone with or without a CS background

(to test how people feel about AGI in general)



If you want to know more ...

Maybe I will write a blog shortly ...

And please stay tuned ...



Thank You

