Training a foundation image model and deploying it to millions of users

“a bottle of wine on a marble table”

200 GPU hours

“a bottle of wine on a marble table”

1300 GPU hours

“a bottle of wine on a marble table”

11 000 GPU hours

“a bottle of wine on a marble table”

50 000 GPU hours

About Photoroom

Photoroom

We help sellers create images that sell

- Trusted by Amazon, Decathlon, Doordash
- 200+ million downloads
- Available on Mobile, Web, and API

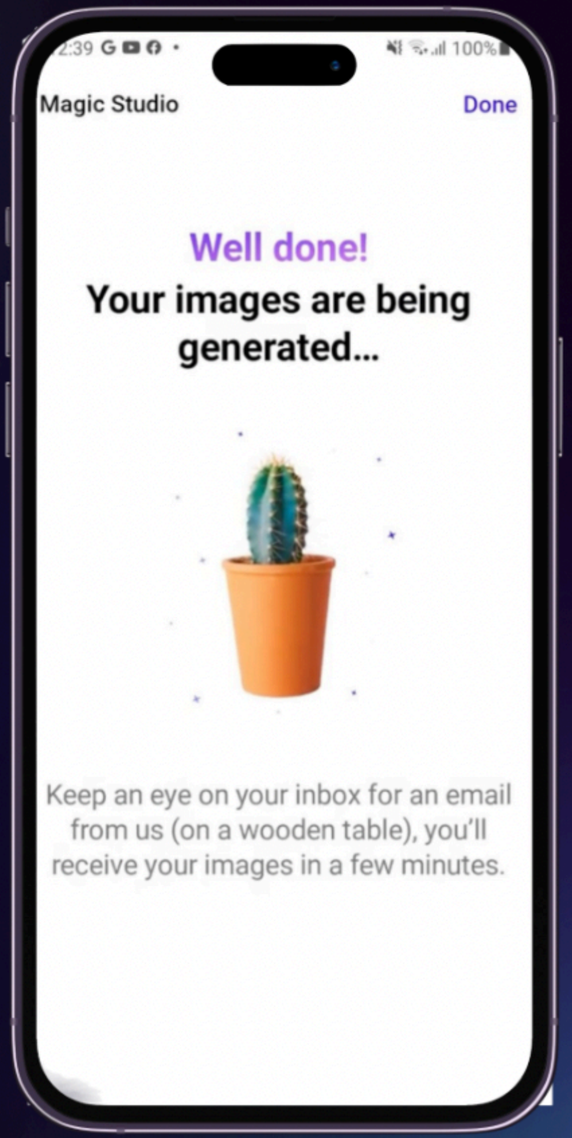
Lyline LIM

Head of Impact & Sustainability

at Photoroom

Why train a model from scratch?

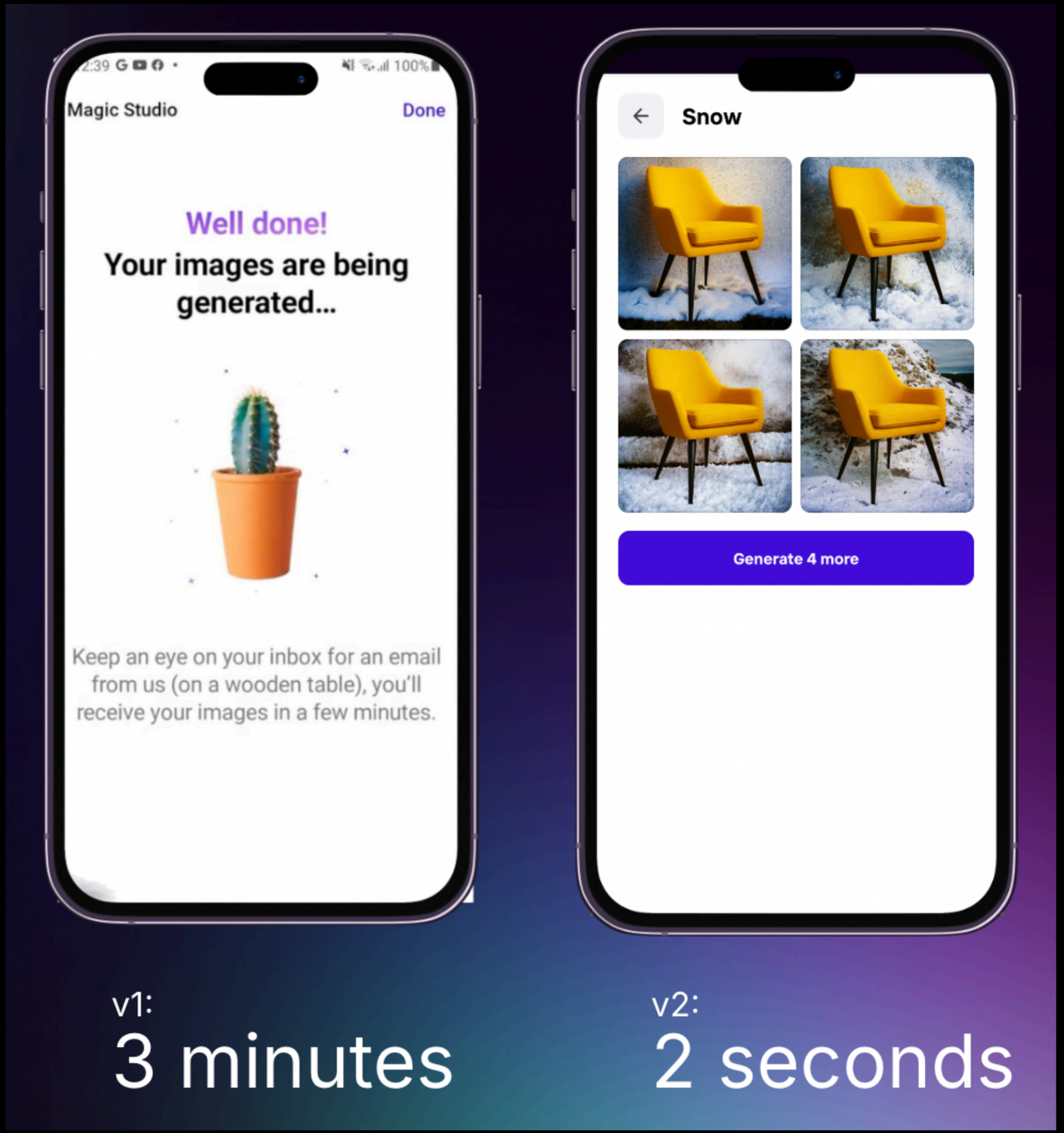
August 2022 - Launching our GenAI hackathon



v1:
3 minutes

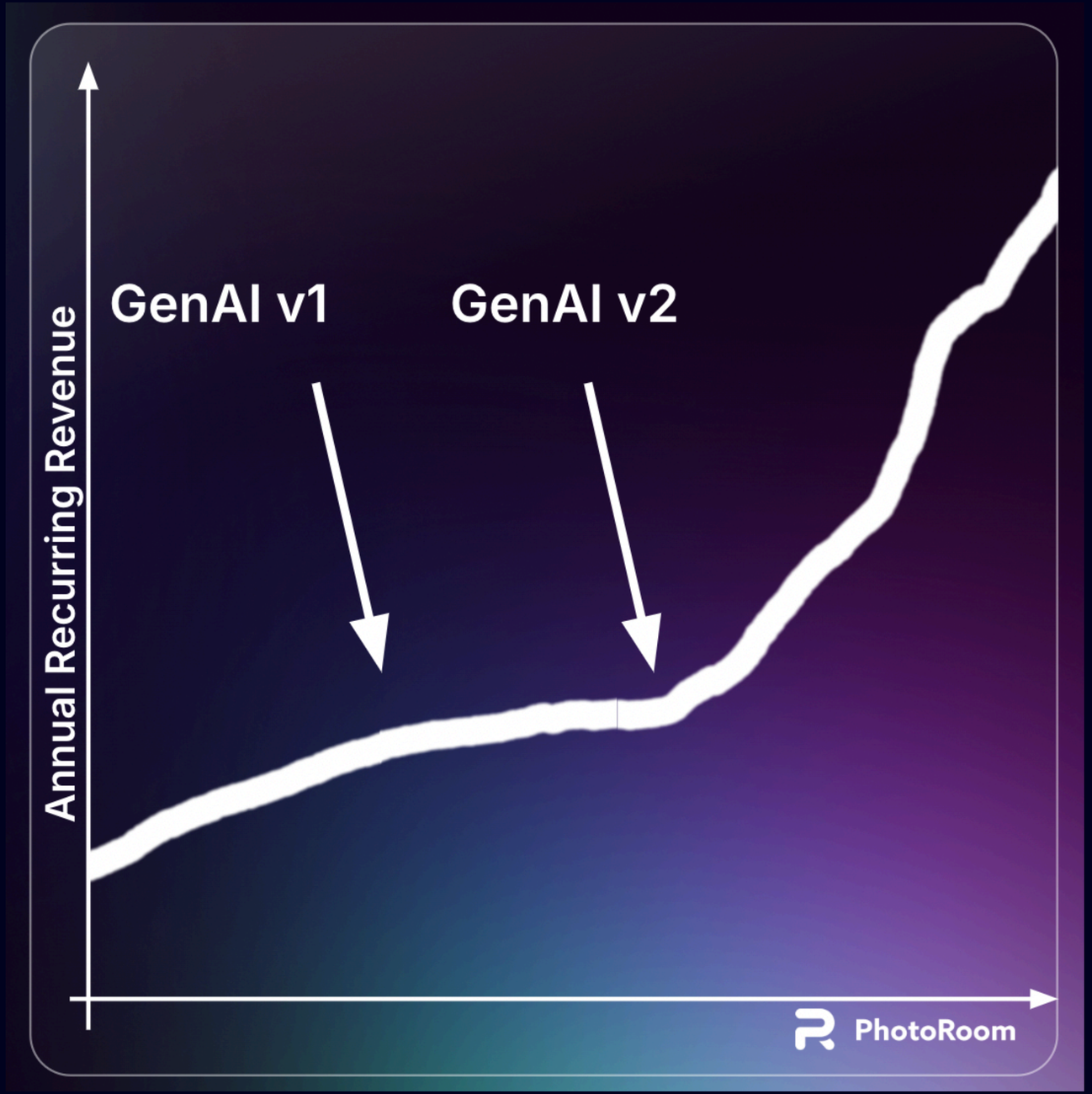
1. Quality matters

2. Speed matters



3. Accessibility matters

Gen AI impact on our business



Why train a model from scratch?

Quality

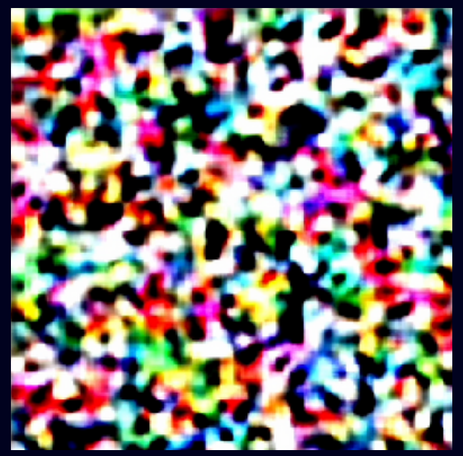
Speed

Control

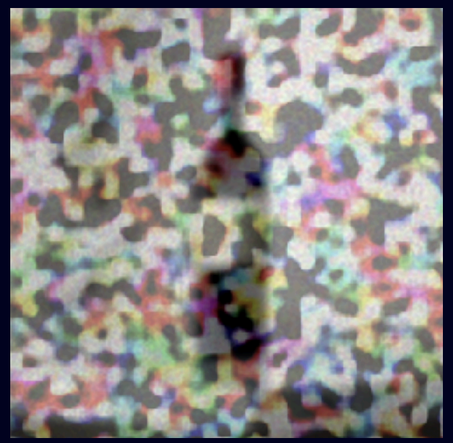
How does diffusion models work?

How diffusion works in 1 minute

"a bottle of wine on a marble table"



"a bottle of wine on a marble table"



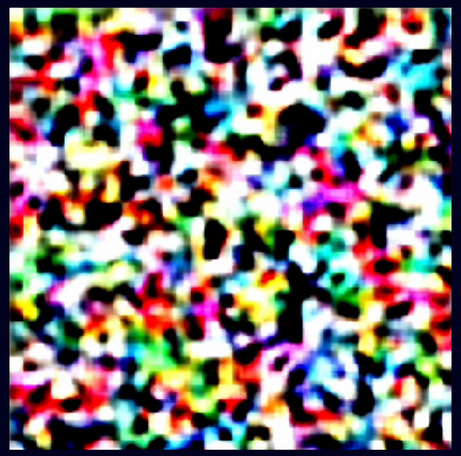
"a bottle of wine on a marble table"

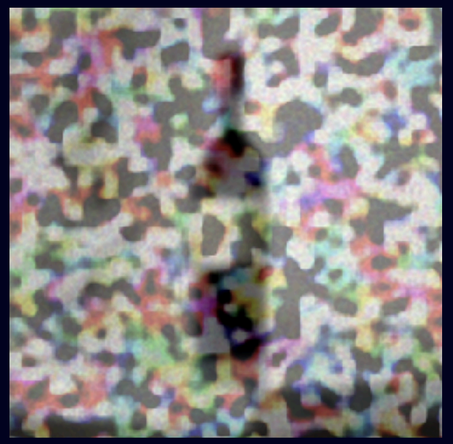
How diffusion trains in 1 minute

“a bottle of wine on a marble table”

“a bottle of wine on a marble table”

“a bottle of wine on a marble table”



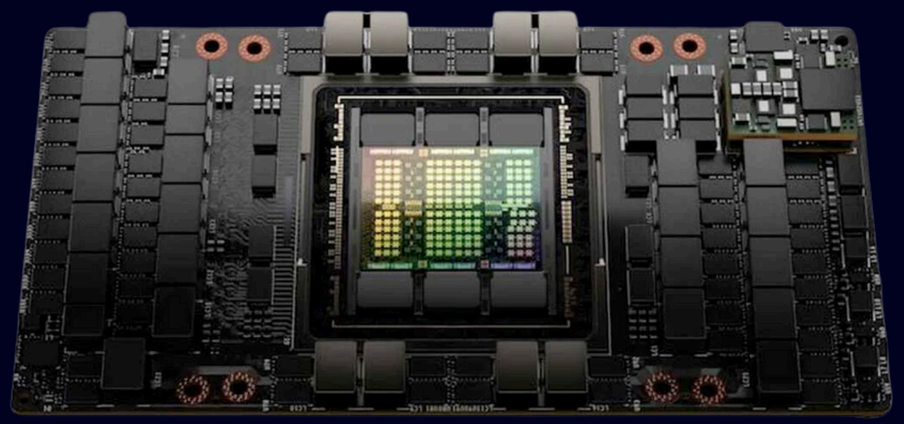


3 key ingredients of an image model

Compute

Algorithms

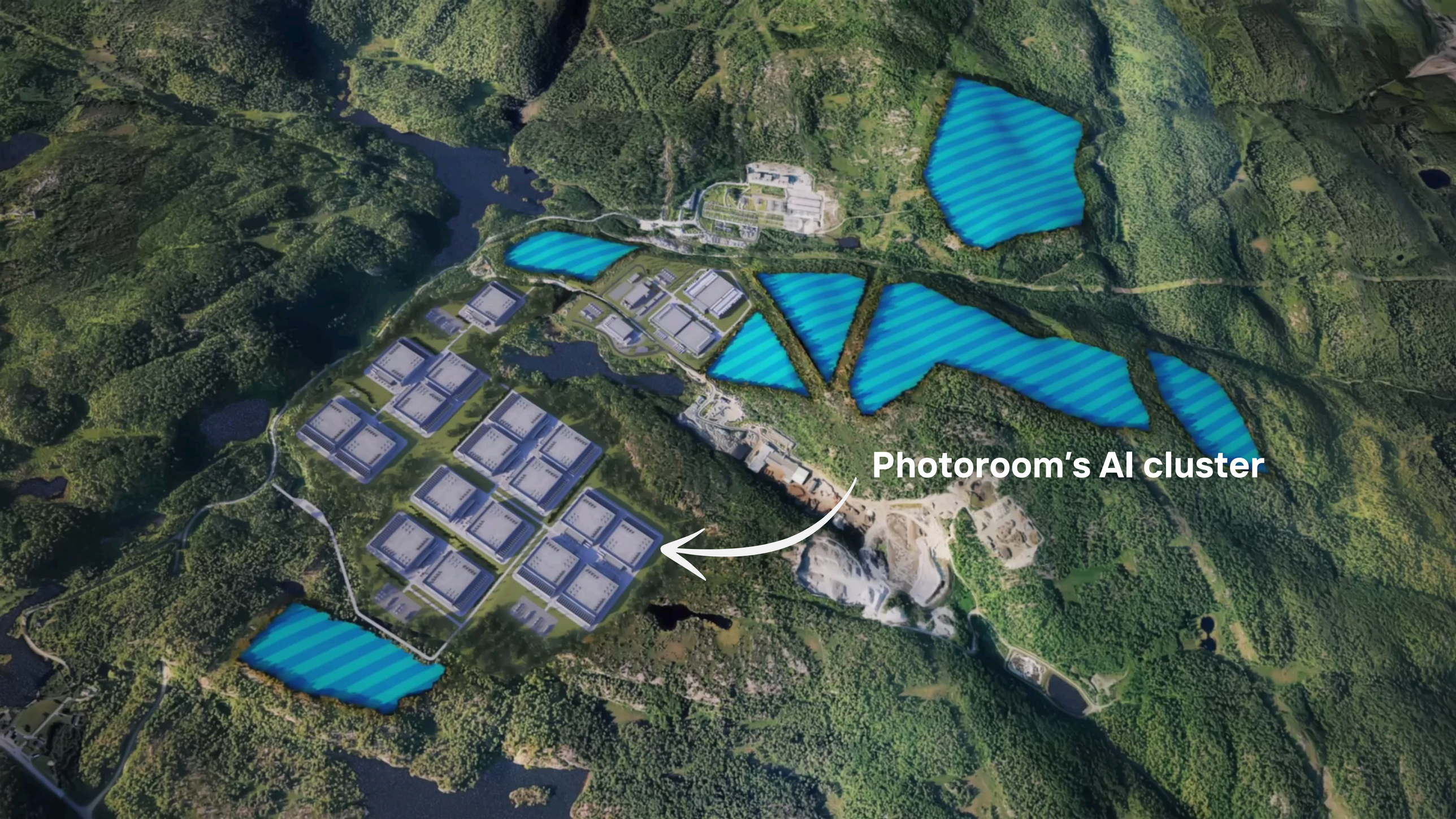
Data



700W per GPU

$30-$40k

8 per machine


Photoroom’s AI cluster

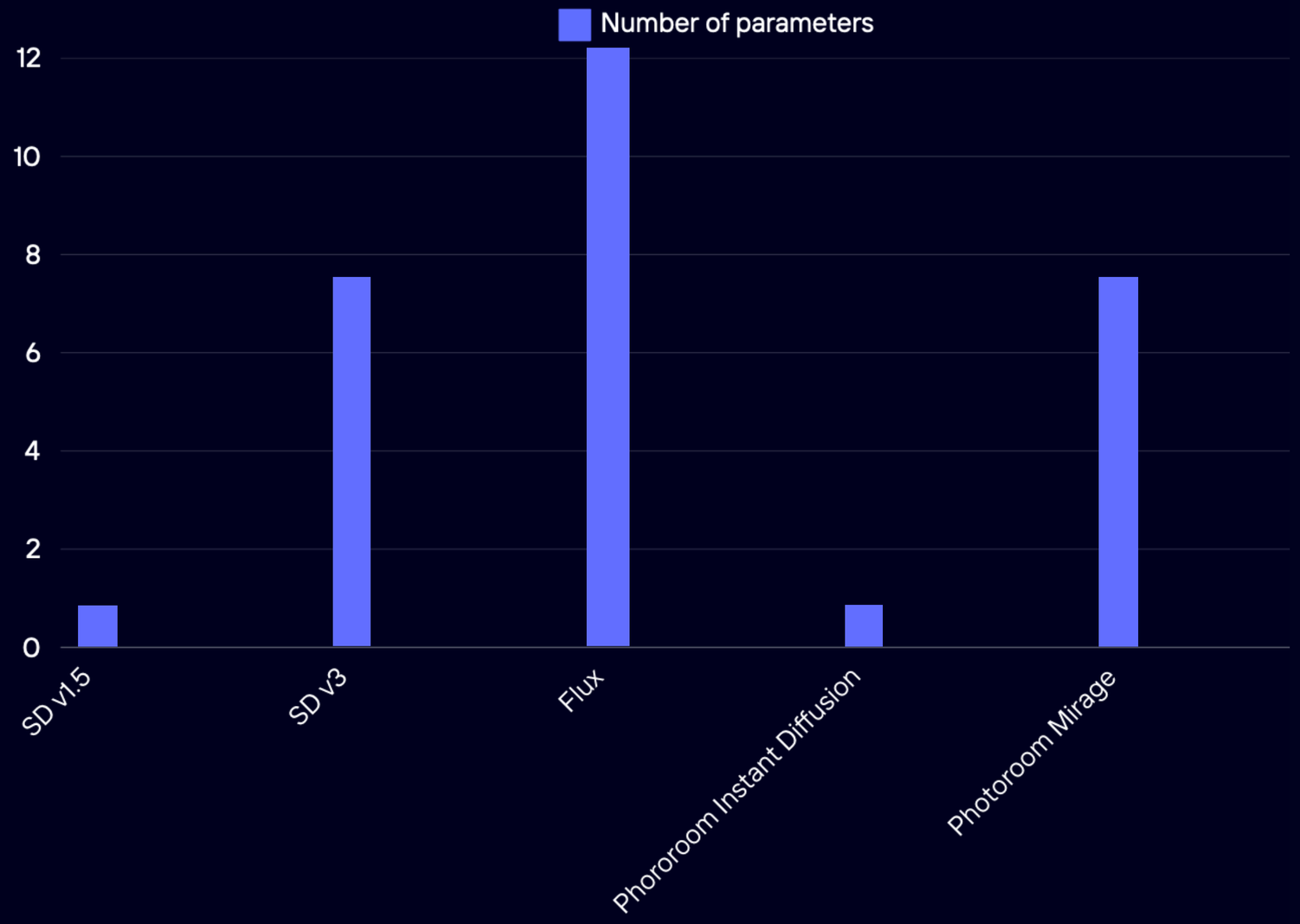
3 key ingredients of an image model

Compute ✅

Algorithms

Data

Algorithm: picking the right architecture



3 key ingredients of an image model

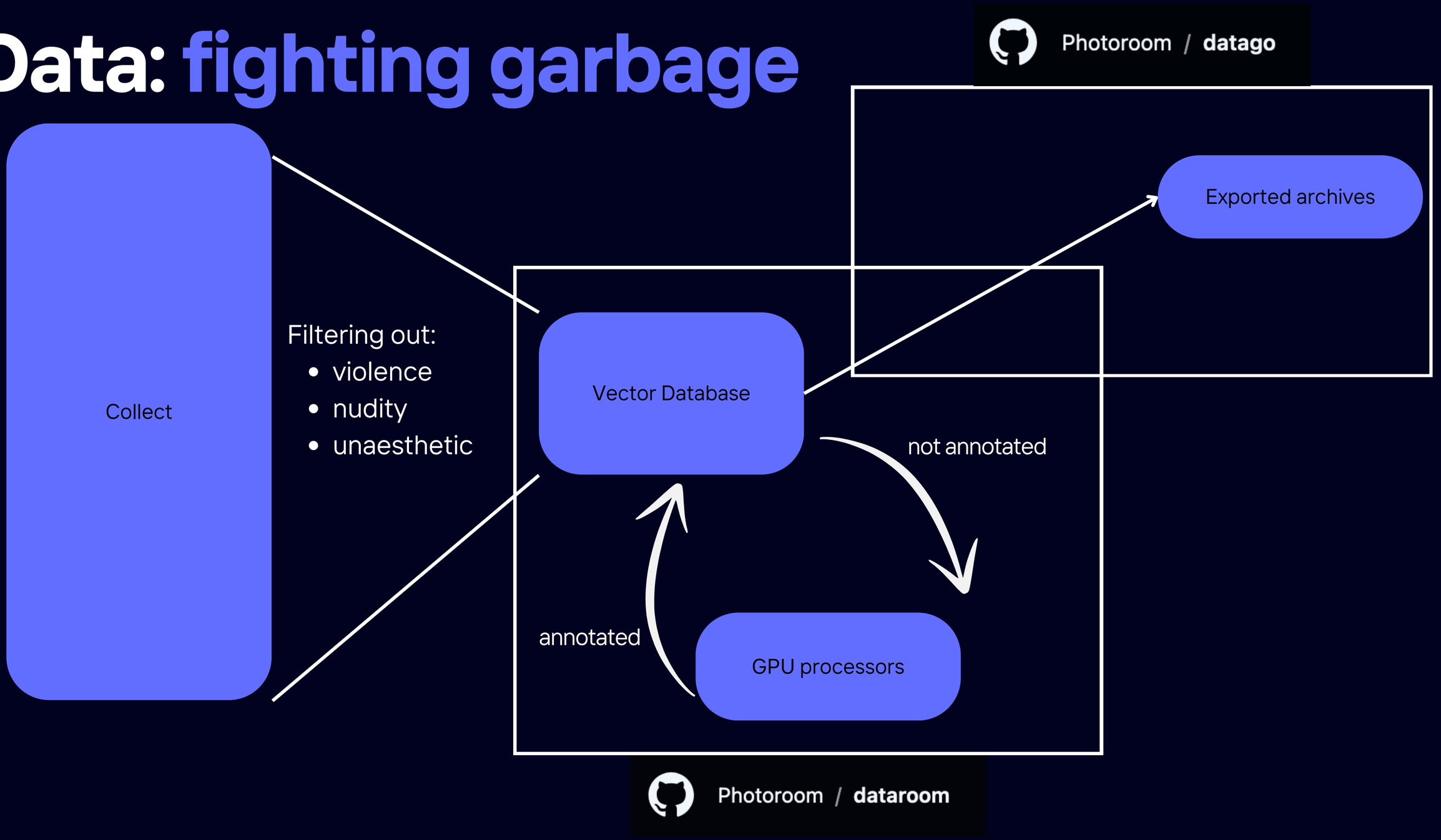
Compute ✅

Algorithms ✅

Data

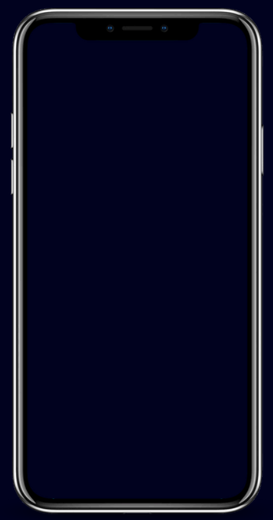
Garbage in
Garbage out

Data: fighting garbage

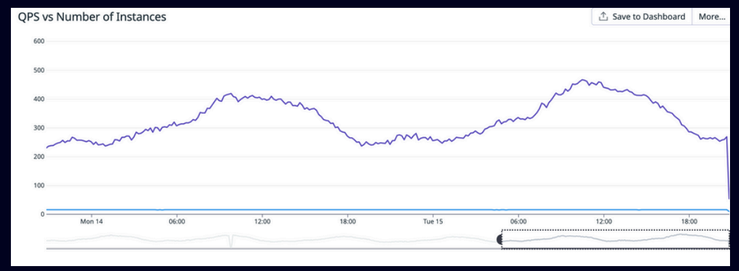


How to deploy to millions of users?

Inference: infrastructure & software

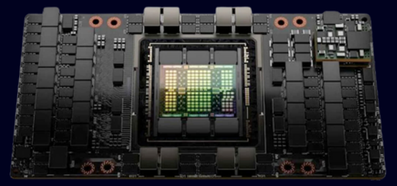


Request

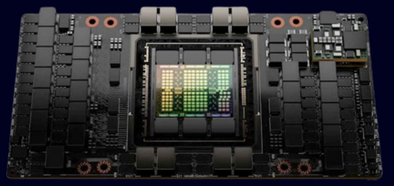


Request

Request



Request



We built our own model to answer customers needs

Quality

Speed

Control

Thank you!

Lyline LIM
Head of Impact & Sustainability
at Photoroom