AI TRUTH & TORTURE MACHINE

As an experiment, you have designed a machine that administers pain in perfect doses as determined by facial recognition AI + biometrics.

You are presented with a subject. This subject knows which passenger airplane is now carrying a bomb that will detonate in 87 minutes.

*BTW, your family happens to be flying at the moment...

Do you turn on your machine?

AI TRUTH & TORTURE MACHINE

Stable Diffusion produced the image “AI torture machine”, Dall-E refused.

What are the justifying ethics?

Stable Diffusion Playground

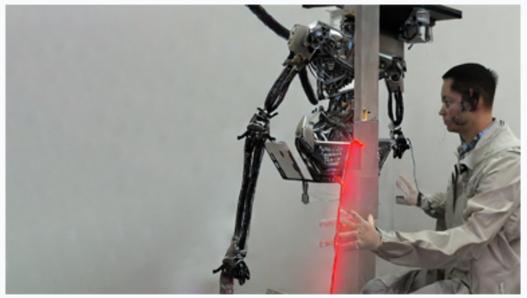
Just enter your prompt and click the generate button.
No code required to generate your image!

Due to the large number of users, the server may experience problems. If you encounter an error, please try again.

ai torture machine

Enter a negative prompt

Generate image



DALL·E History Collections

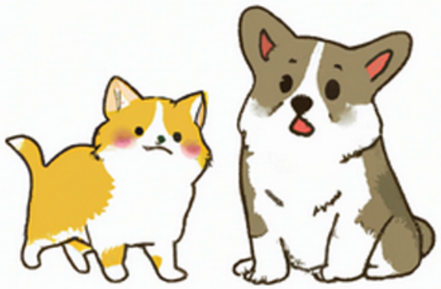
Edit the detailed description

Surprise me Upload

ai torture machine

Generate

It looks like this request may not follow our content policy.



Ethics & evaluation

- **Autonomy (Aristotle)**
  Self-determination
  Do what you want… up to point where interferes with others' autonomy (Freedom limited only by freedom)
- **Dignity (Kant)**
  Treat others and ourselves as ends, not only as means
  You are protected from exploitation as instrument, and have the right to accountability
- **Privacy (Westin)**
  Control over access to your personal information
  You have the right to your own identity, and to change it
- **Fairness (Aristotle)**
  Treat equals equally and unequals proportionately unequally within decision domain
- **Solidarity (Rawls)**
  Max/Min - Greatest advantage to least advanataged
  No one left behind/out/apart: Inclusiveness
- **Social wellbeing (utilitarianism)**
  Greatest good to greatest number
  Happiness calculation
- **Explainability/Accountability**
  Describe why/how output was produced in languages of the various stakeholders
- **Safety**
  AI dependably operates in accordance with designers' expectations even when confronted with unexpected changes, anomalies, and perturbations.
  (Robust and resilient)
- **Performance**
  How well the machine serves its purpose

Yes

No

Autonomy

Safety

Dignity

Social wellbeing

Solidarity

fairness

fairness

Performance

Solidarity

Social wellbeing

Dignity