

Machine Learning and Architecture

But what is AI?

Introduction to Machine Learning and this uses in the field of Architecture

Overview

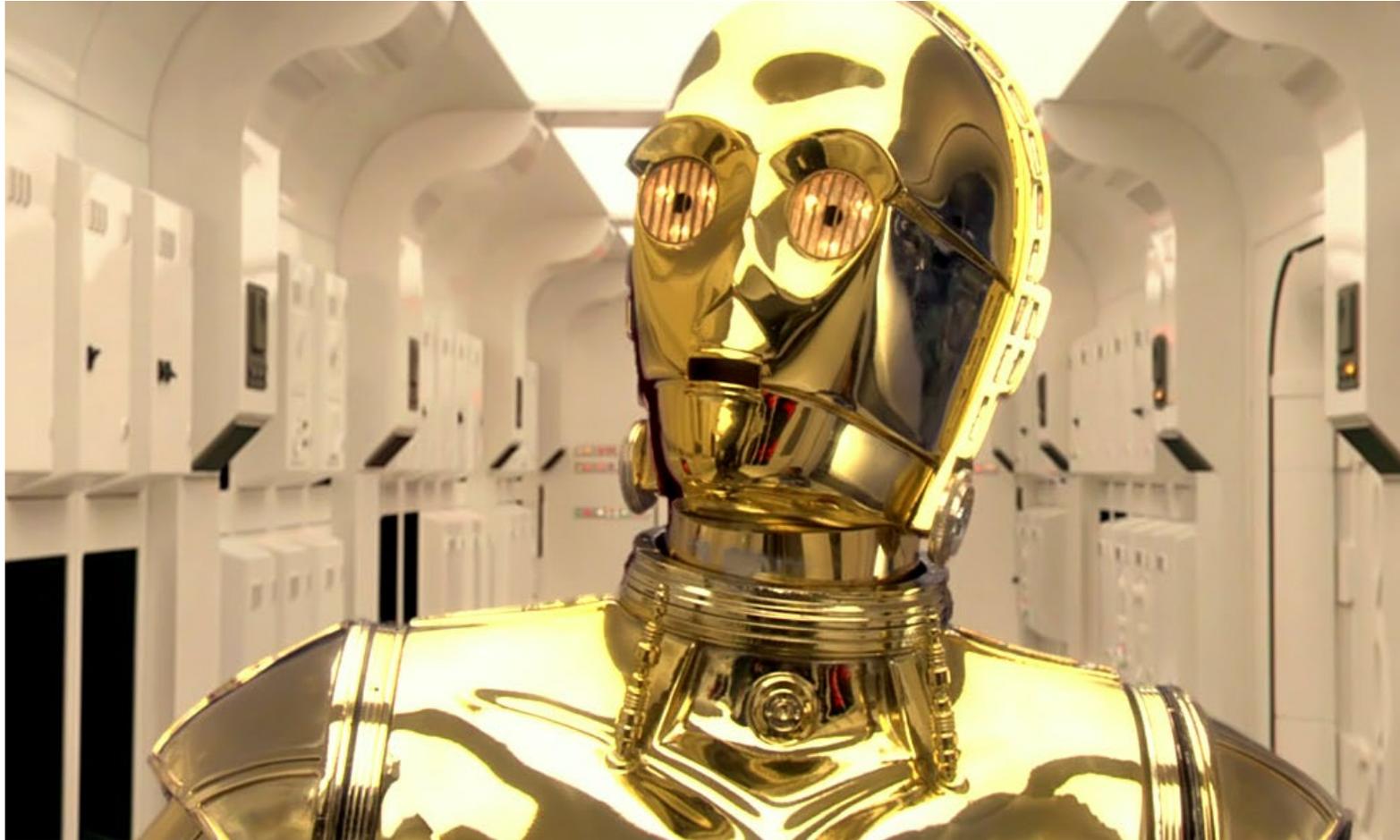
1. Introduction to Artificial Intelligence
2. Natural Language Processing
3. Image Processing

4. Applications in Practice
5. Promoting the Reuse of Vacant buildings
6. LINK Kompass®

7. Open Source and an outlook to the future

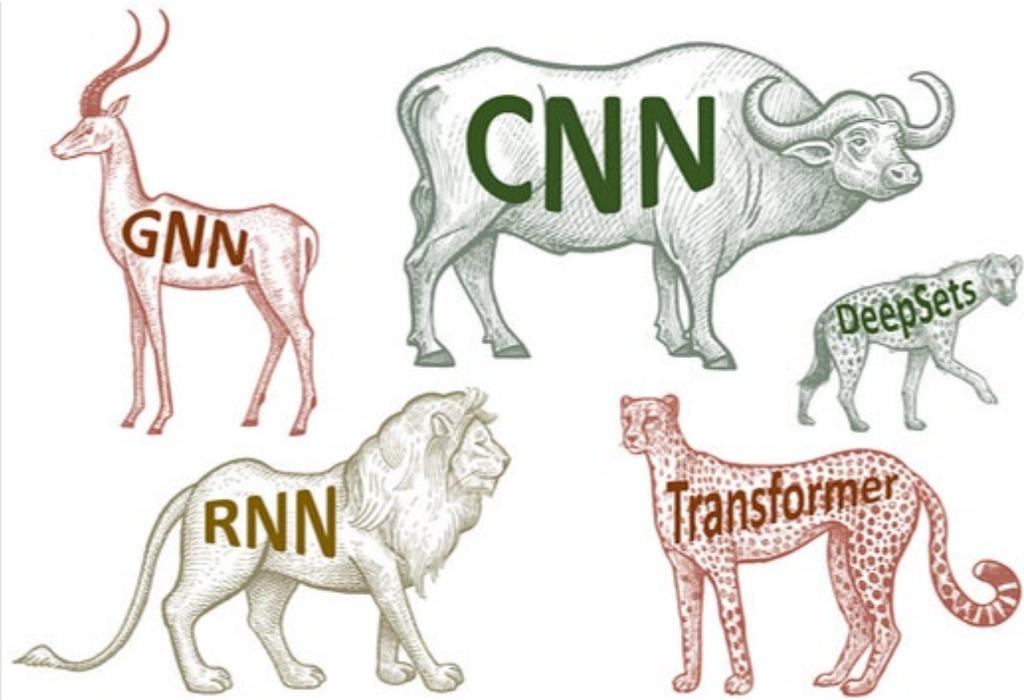
Introduction to Artificial Intelligence

What is Machine Learning?



Star Wars | Episode I, C-3PO

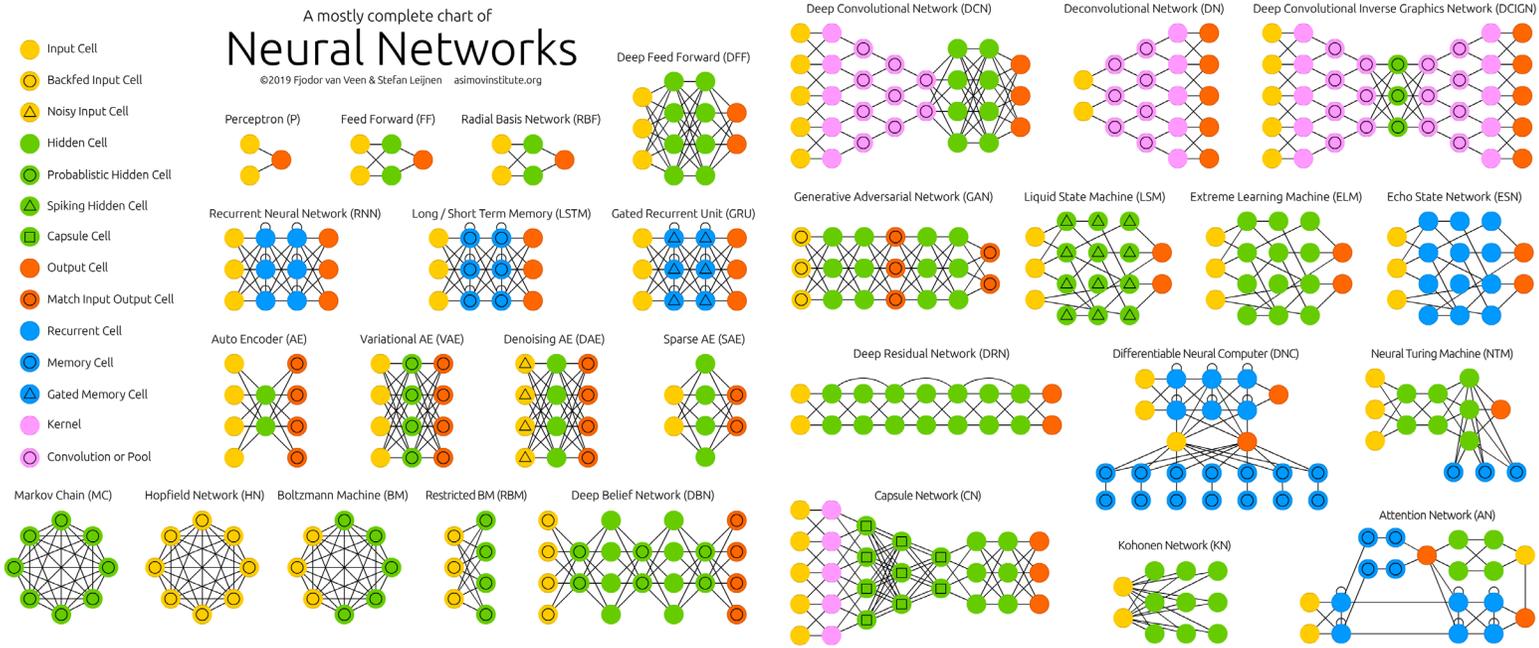
Zoo of Acronyms



Geometric Deep Learning: Grids, Groups, Graphs, Geodesics, and Gauges
Michael M. Bronstein, Joan Bruna, Taco Cohen, Petar Veličković

- CNN - Convolution Neural Network
- GNN - Graph Neural Network
- RNN - Recuring Neural Network

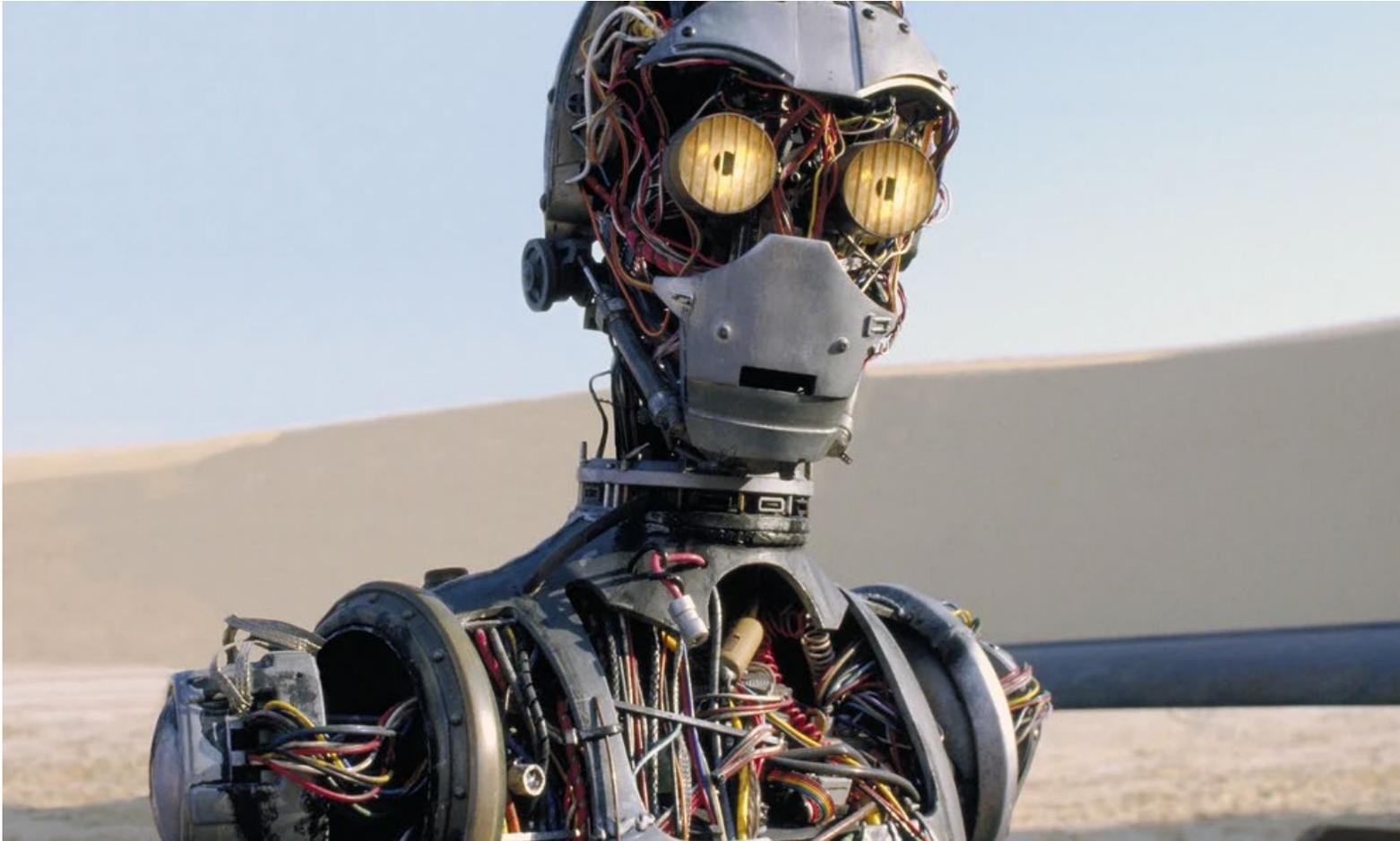
Neural Networks



A mostly complete chart of Neural Networks
 Fjodor van Veen & Stefan Leijnen asimovinstitute.org

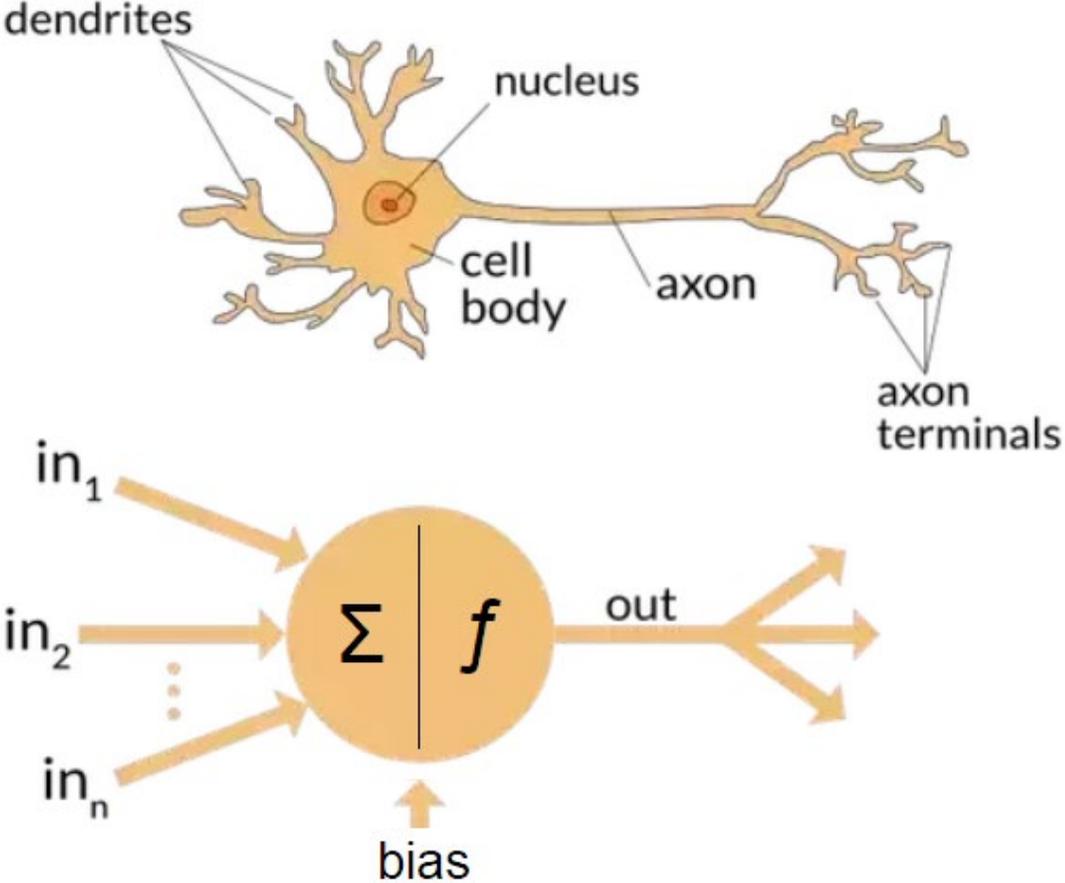


Looking under the hood



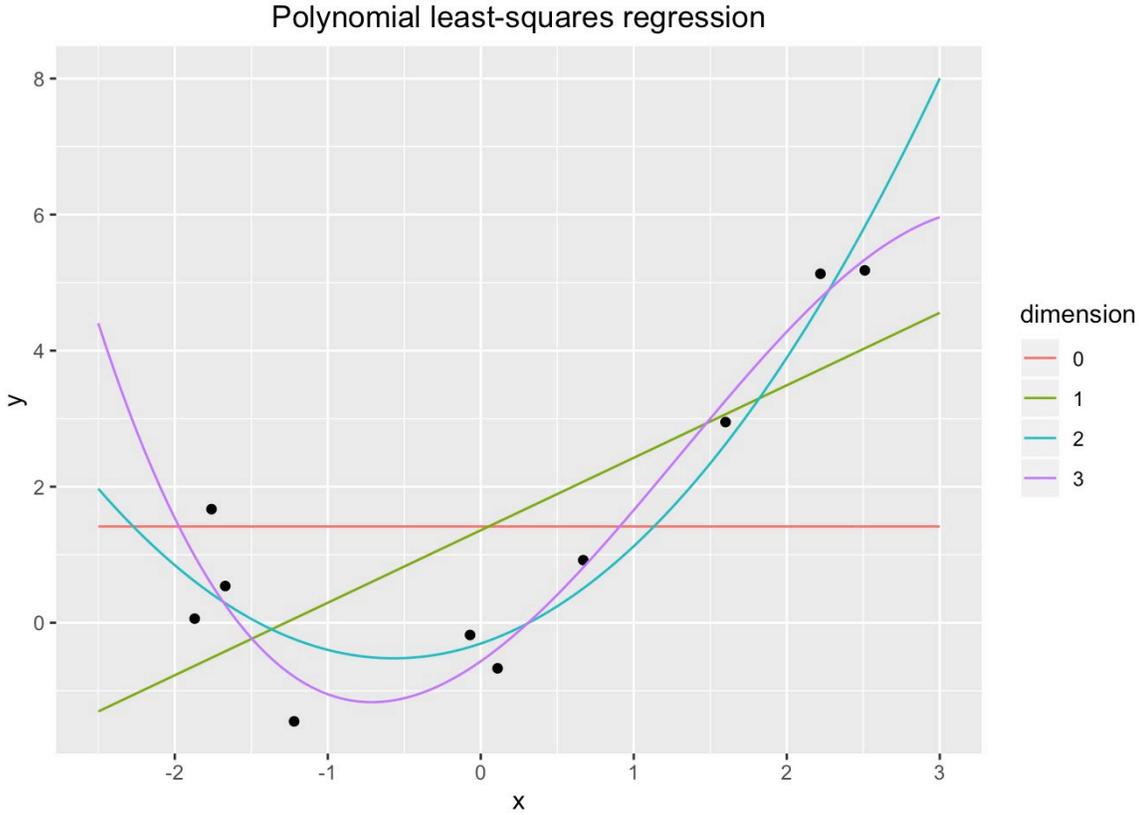
Star Wars Episode I, C-3PO

Neural Networks, Dendrite vs. Node



Comparison of biological neuron to a node in a machine learning model

But then what is Machine learning



It is “training” a function to output some specific values

$$f(x) = y$$

$$f(x) = Ax^2 + Bx + C$$

Specifically:

$$h_{\theta}(x) = \theta_0x_0 + \theta_1x_1 + \theta_2x_2 \dots \theta_nx_n$$

Or simply just:

$$h_{\theta}(x) = \theta T x$$

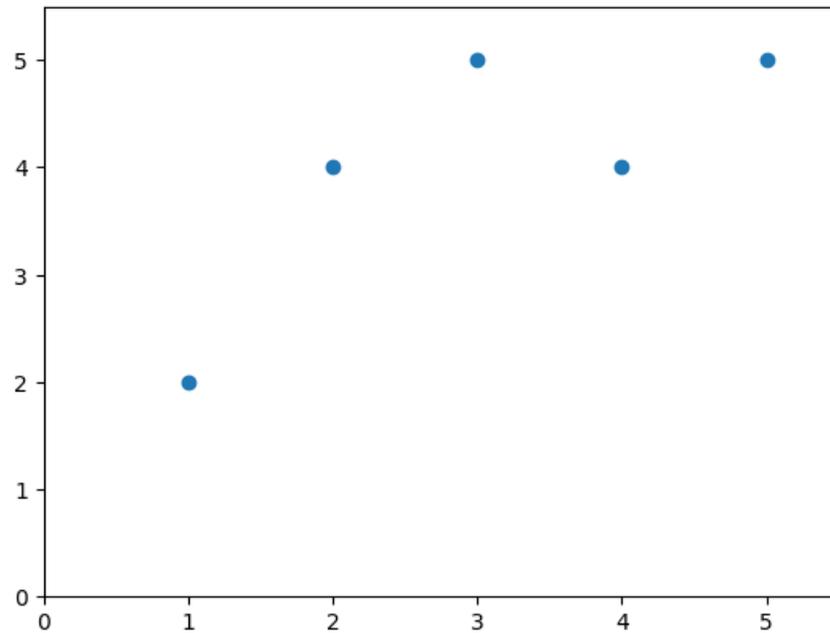
But then what is Machine learning



Chappie (2015), Chappie

Practical Example

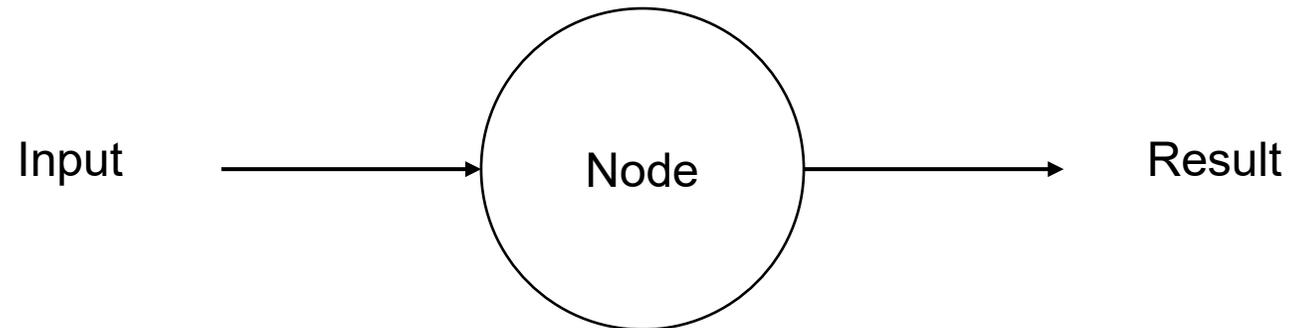
- First, we need some training data. That is data for which we have both the input as well as the result.



x	y
1	2
2	4
3	5
4	4
5	5

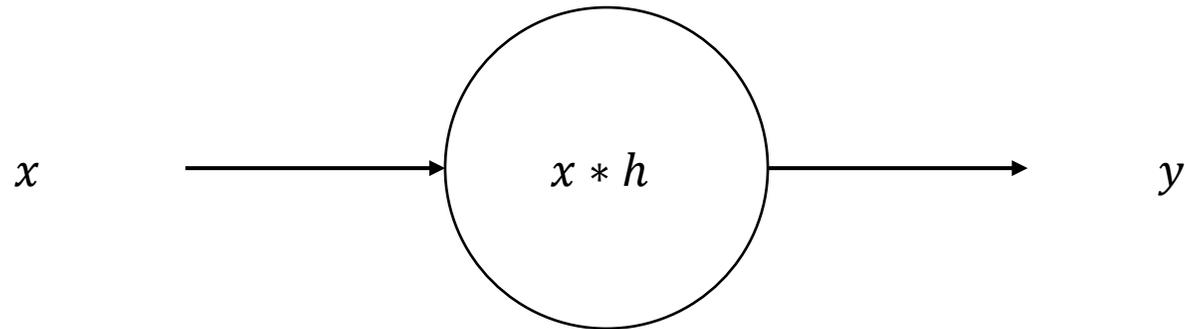
Practical Example

- Simple Network of one Node



Practical Example

- Simple Network of one Node



Practical Example

- We then input a value into our function
- And get a random result

$$x = 1$$
$$f(1) = 10$$

- That result we now compare to our correct value.

$$y = 2$$
$$(10 - 2)^2 = 64$$

Mean Squared Error (MSE)

Practical Example

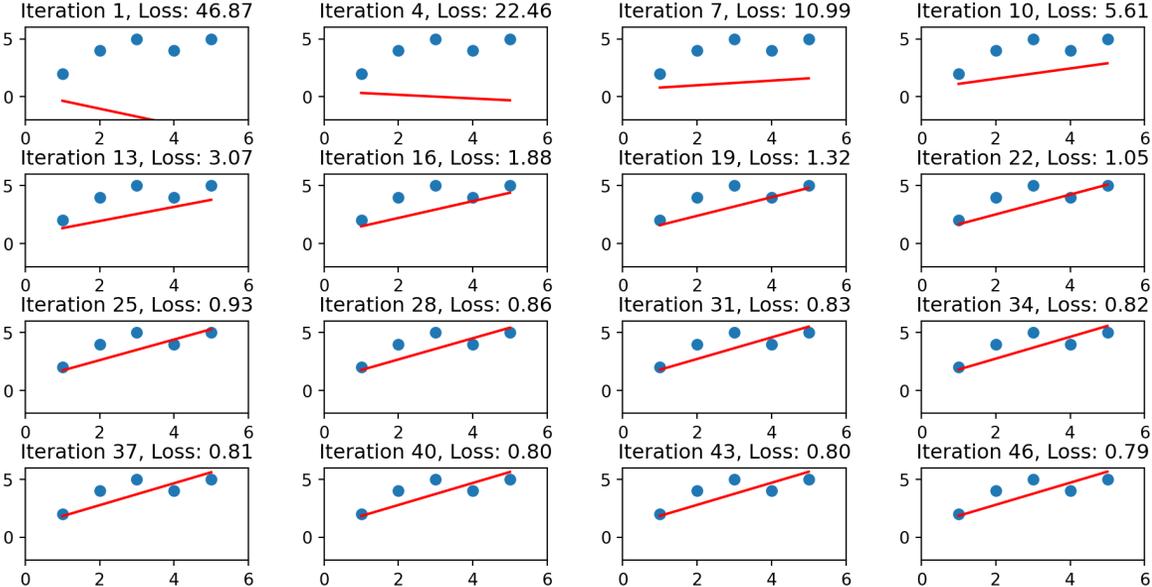
- That gives us a value for how well our model performs.

$$\textit{loss} = 64$$

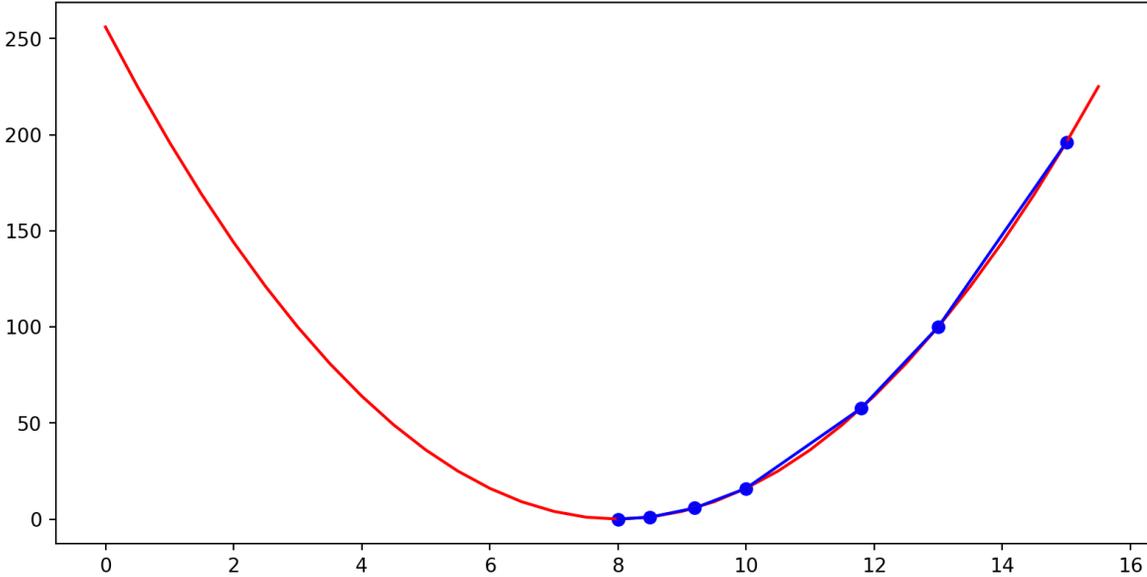
Not that great

Practical Example

Next, we adjust some values in our model and repeat the process (loss)



Simplified



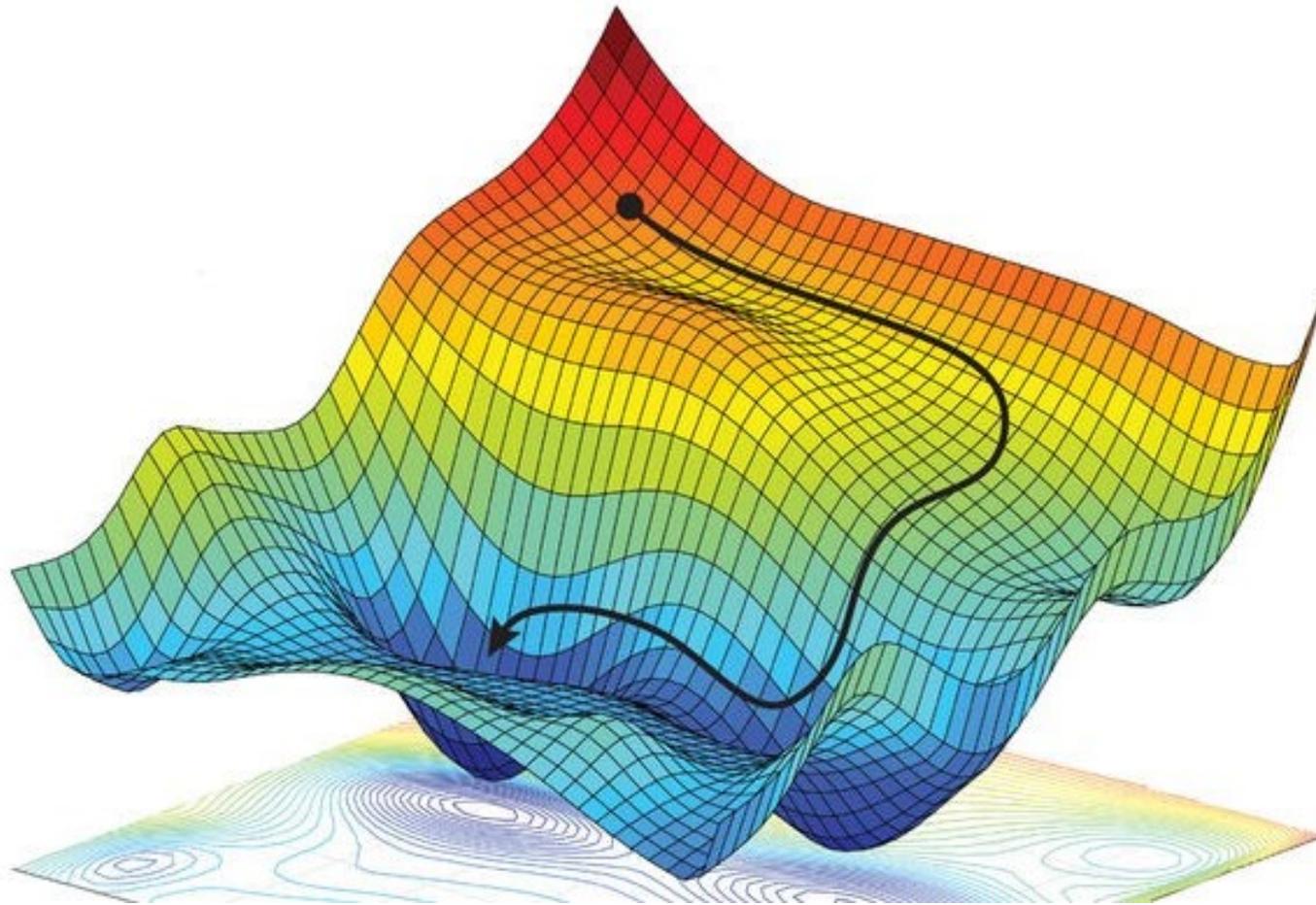
Loss over value for

So, for the function

$$f(x) = x * p = y$$

What is p ?

x	y
2	16
3	24



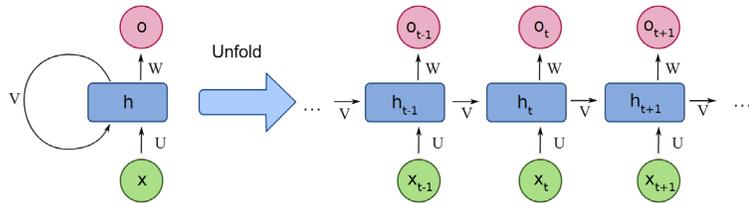
Gradient decent

Natural Language Processing (NLP)

Machine Learning and Architecture

Transformers

Attention is all you need (arXiv:1706.03762)



RNN Model Diagram
Predecessor to Transformers

Transformers innovation over RNN's and LSTM's is that the training can be parallelized rather than processing a single word at the time. It does this through adding **Position Encoding** and they added the concept of **Attention**.

- Attention
 - Self Attention
 - Causal self attention

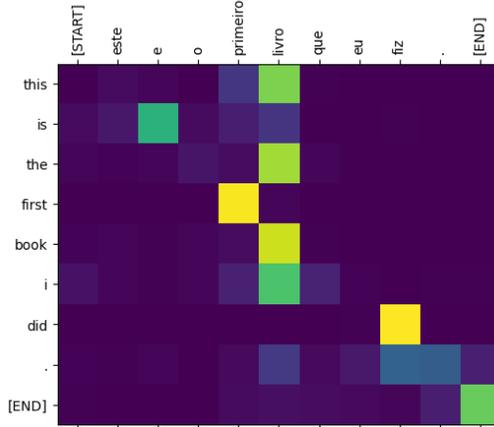
Simple example of positional encoding

Sequence	Index of token, k	Positional Encoding Matrix with $d=4, n=100$			
		$i=0$	$i=0$	$i=1$	$i=1$
I	0	$P_{00}=\sin(0)$ = 0	$P_{01}=\cos(0)$ = 1	$P_{02}=\sin(0)$ = 0	$P_{03}=\cos(0)$ = 1
am	1	$P_{10}=\sin(1/1)$ = 0.84	$P_{11}=\cos(1/1)$ = 0.54	$P_{12}=\sin(1/10)$ = 0.10	$P_{13}=\cos(1/10)$ = 1.0
a	2	$P_{20}=\sin(2/1)$ = 0.91	$P_{21}=\cos(2/1)$ = -0.42	$P_{22}=\sin(2/10)$ = 0.20	$P_{23}=\cos(2/10)$ = 0.98
Robot	3	$P_{30}=\sin(3/1)$ = 0.14	$P_{31}=\cos(3/1)$ = -0.99	$P_{32}=\sin(3/10)$ = 0.30	$P_{33}=\cos(3/10)$ = 0.96

Positional Encoding Matrix for the sequence 'I am a robot'



Transformer



Attention matrix

Boot time	Boot time is super-fast, around anywhere from 35 seconds to 1minute, but quite unreasonably priced	positive
priced	Boot time is super-fast, around anywhere from 35 seconds to 1minute, but quite unreasonably priced	negative
menu	The menu is limited but almost all of the dishes are excellent	negative
dish	The menu is limited but almost all of the dishes are excellent	positive

Sentences with words colored based on their importance in a specific context

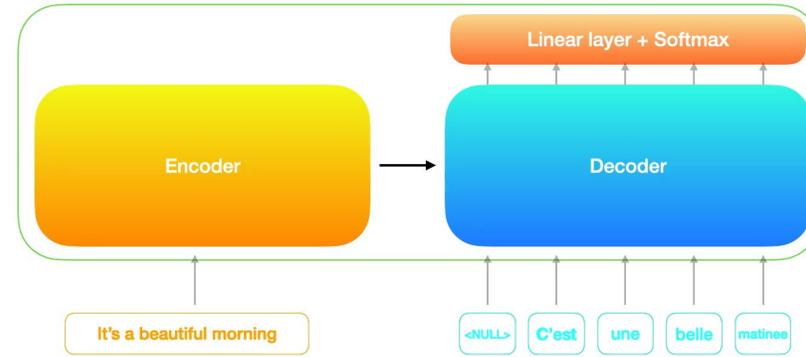


Fig 19: Flow of input and output through transformer model

Tokens



Archive (2020), George Almore & J3

Tokens

GPT-3 Codex

Many words map to one token, but some don't: indivisible.

Unicode characters like emojis may be split into many tokens containing the underlying bytes: 🍌

Sequences of characters commonly found next to each other may be grouped together: 1234567890

Clear

Show example

Tokens	Characters
64	252

TEXT

TOKEN IDS

OpenAI Token Visualizer



Tokens

GPT-3 Codex

Many words map to one token, but some don't: indivisible.

Unicode characters like emojis may be split into many tokens containing the underlying bytes: 🍌

Sequences of characters commonly found next to each other may be grouped together: 1234567890

Clear Show example

Tokens **Characters**
64 **252**

```
[7085, 2456, 3975, 284, 530, 11241, 11, 475, 617, 836, 470, 25, 773, 452, 12843, 13, 198, 198, 3118, 291, 1098, 3435, 588, 795, 13210, 271, 743, 307, 6626, 656, 867, 16326, 7268, 262, 10238, 9881, 25, 12520, 97, 248, 8582, 237, 122, 198, 198, 44015, 3007, 286, 3435, 8811, 1043, 1306, 284, 1123, 584, 743, 307, 32824, 1978, 25, 17031, 2231, 30924, 3829]
```

TEXT **TOKEN IDS**

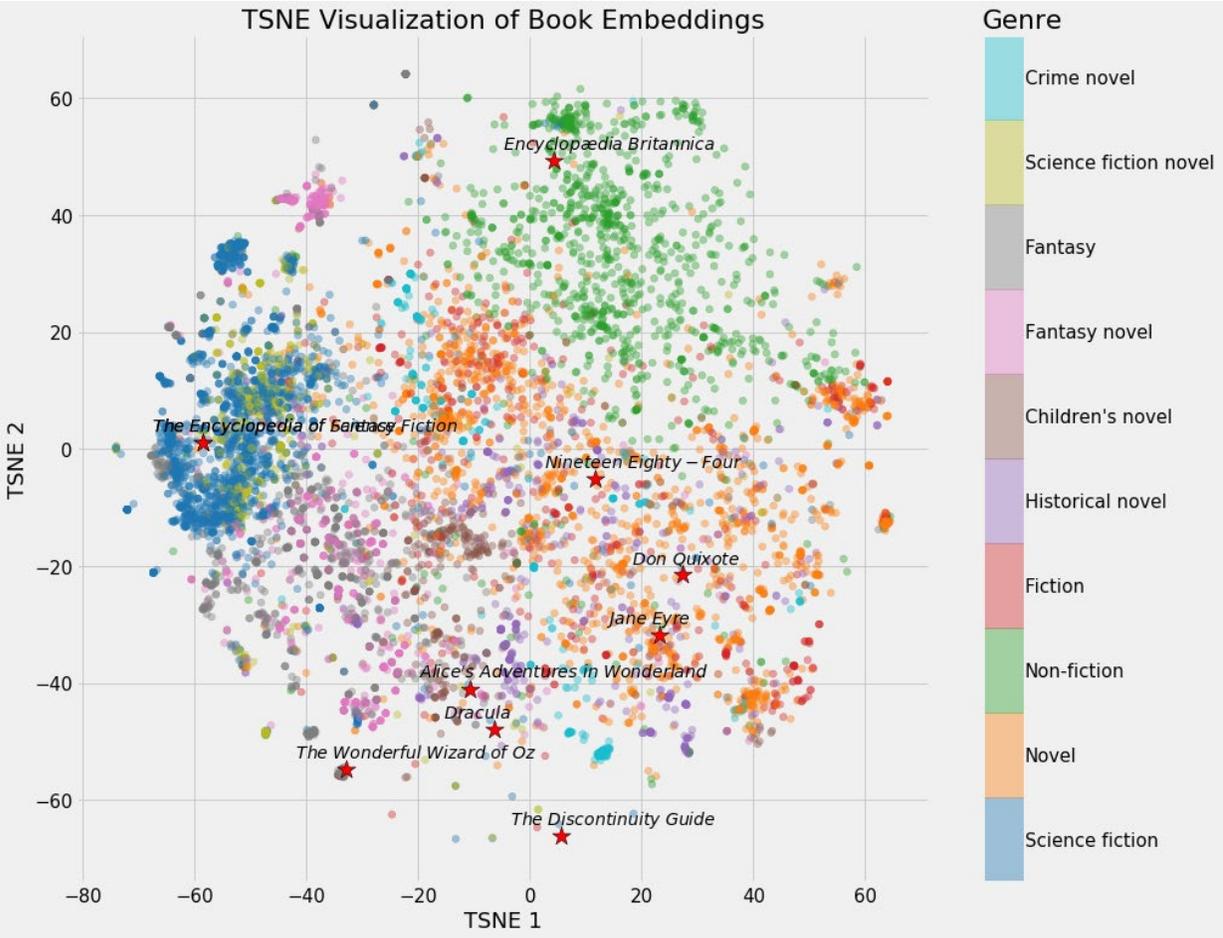
OpenAI Token Visualizer

Embedding
Multidimensional meaning space



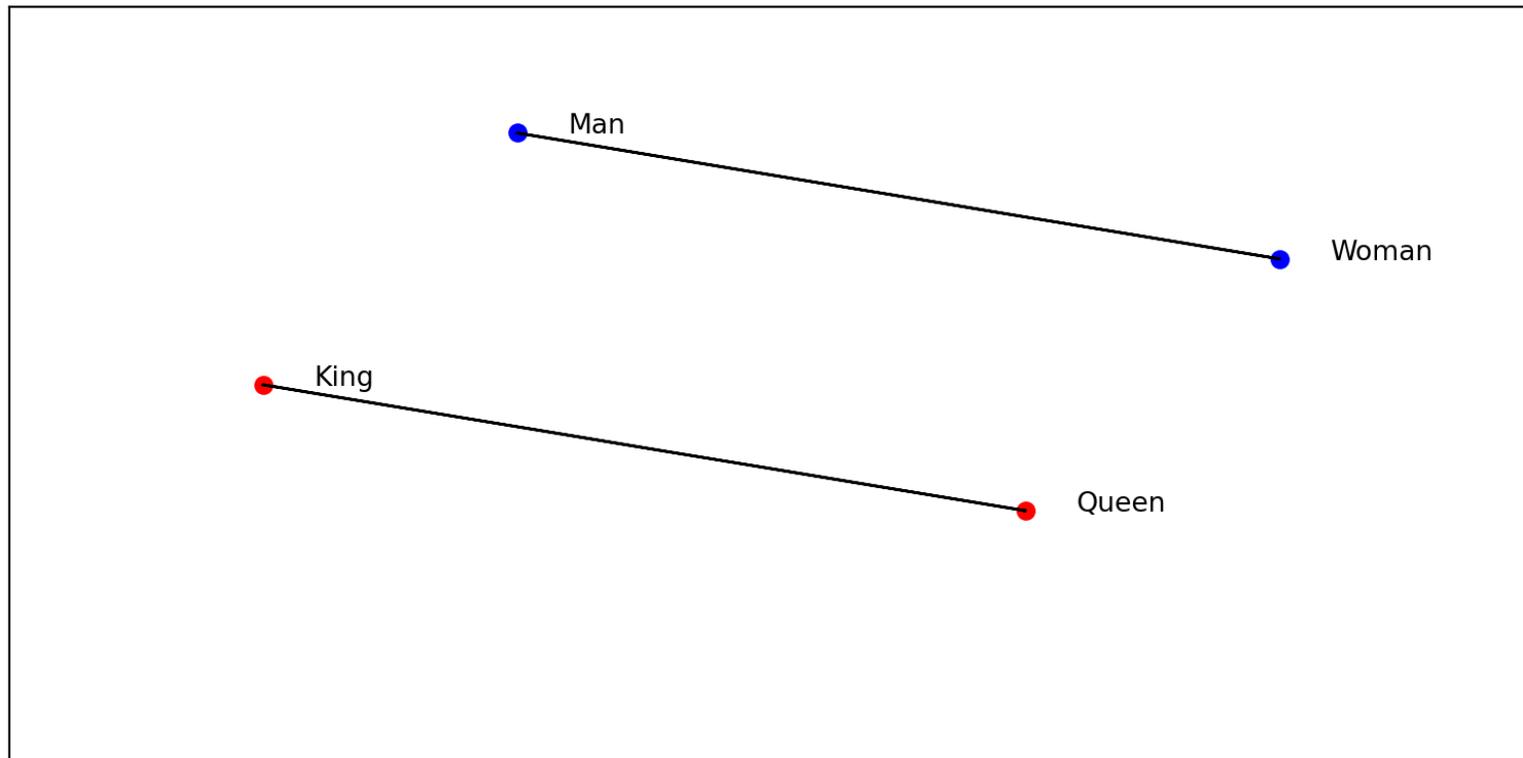
Ghost in the Shell (2017), Monks

Embedding

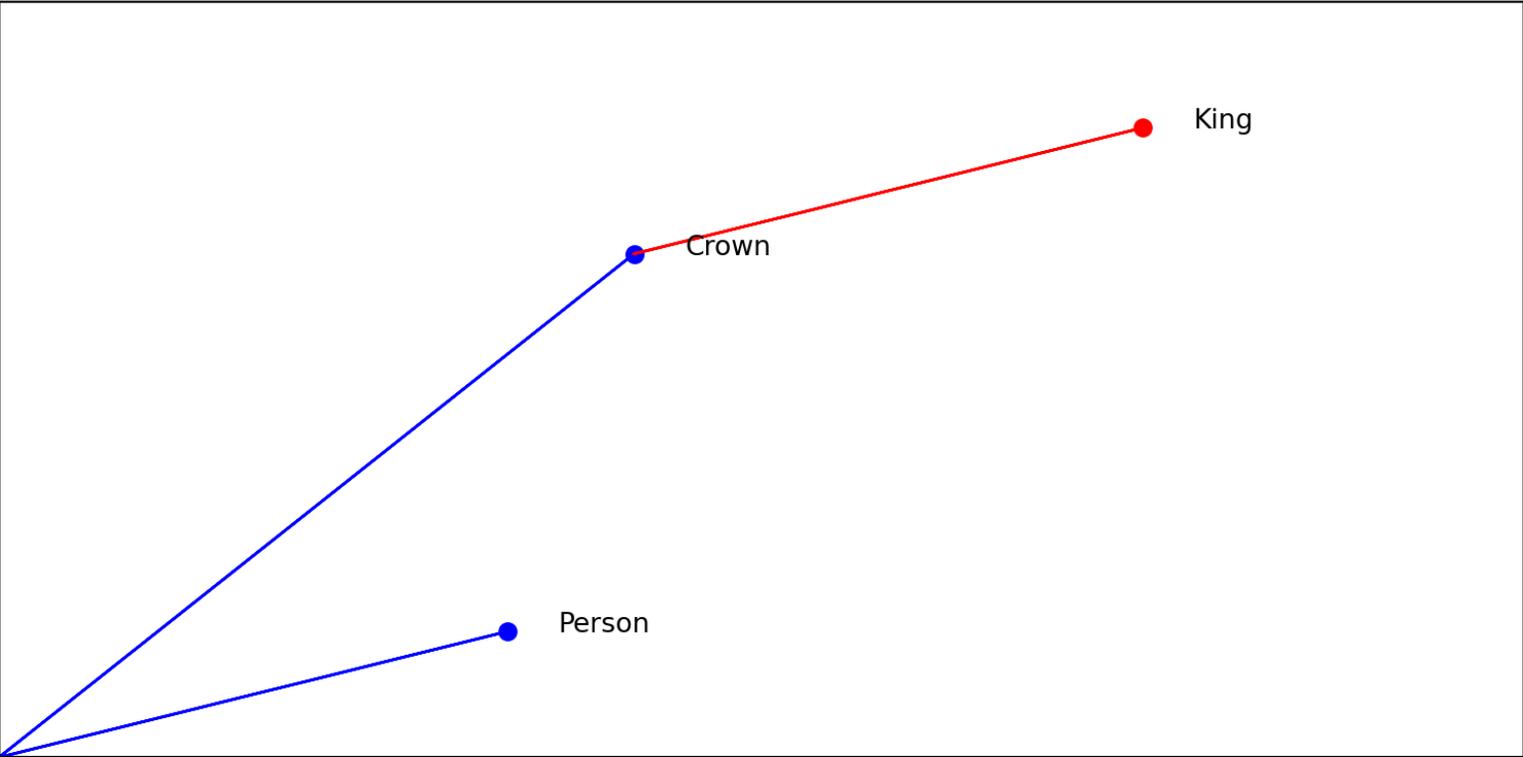


TSNE Visualization of Book Embeddings

Embedding

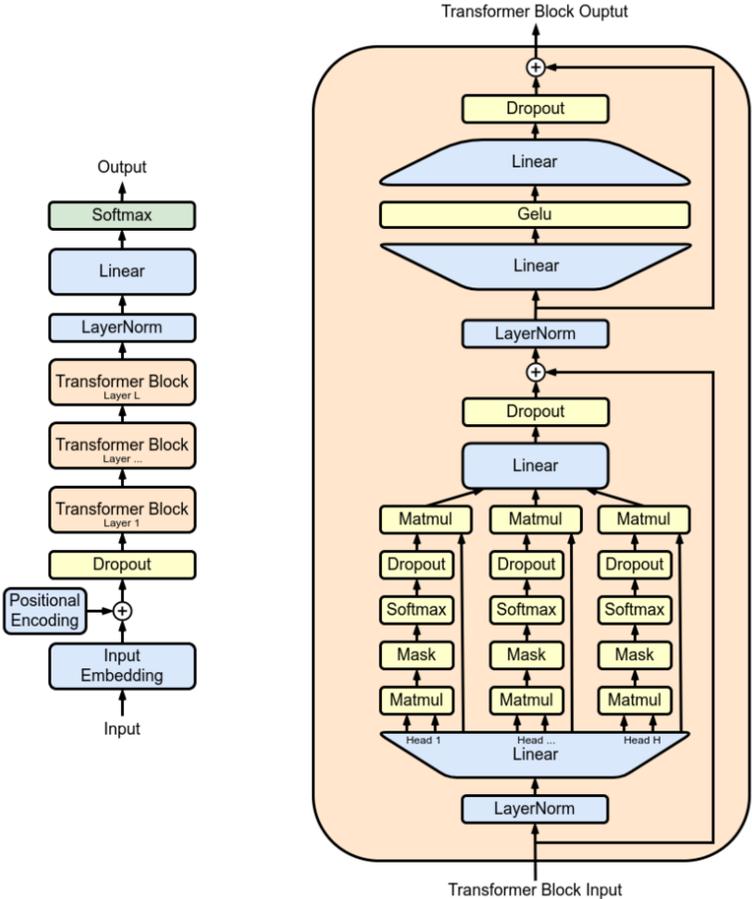


Embedding



Crown + Person => King

GPT Model

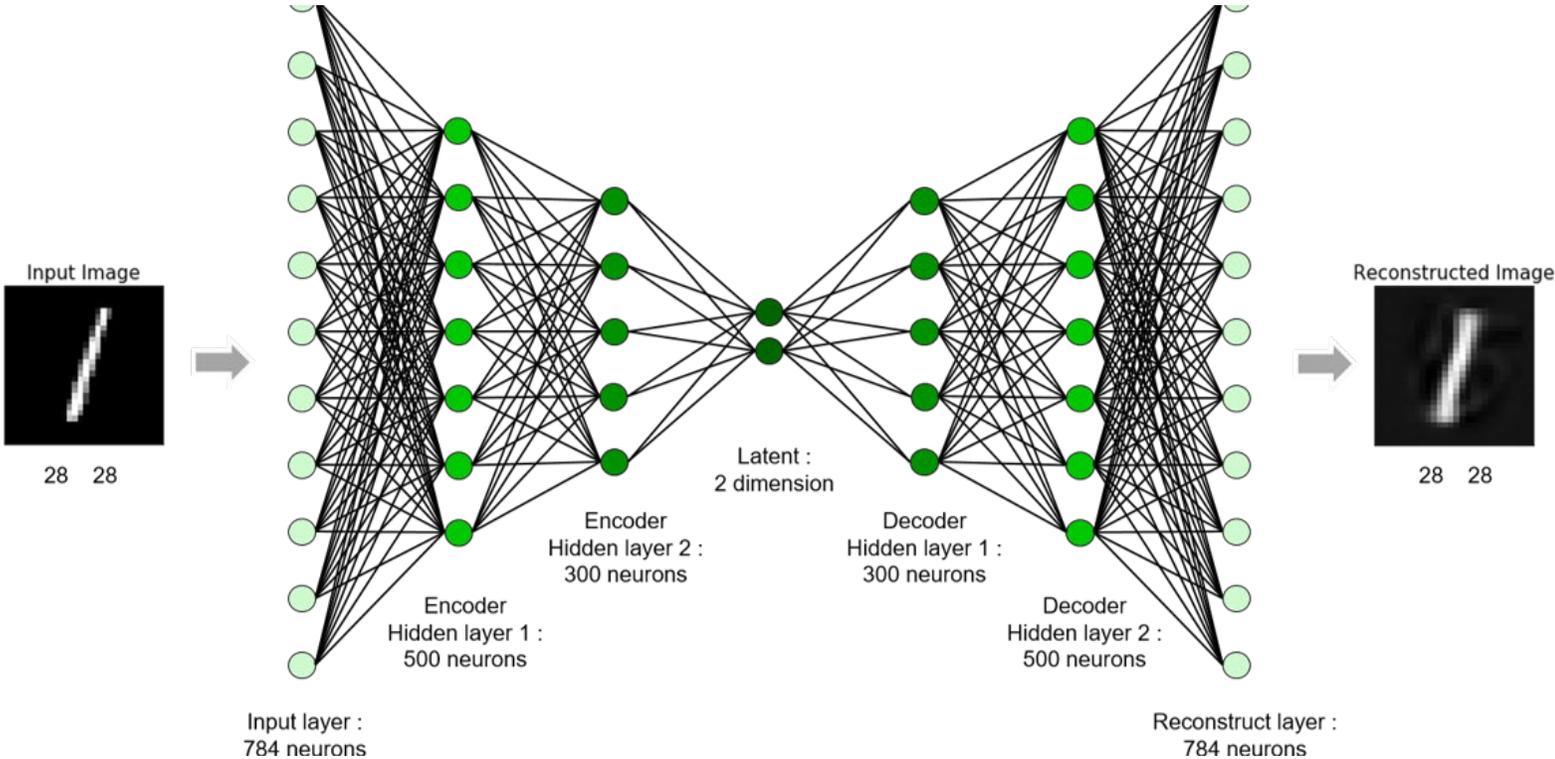


The full architecture of a generative pre-trained transformer (GPT) model. Marxav (Wikimedia)

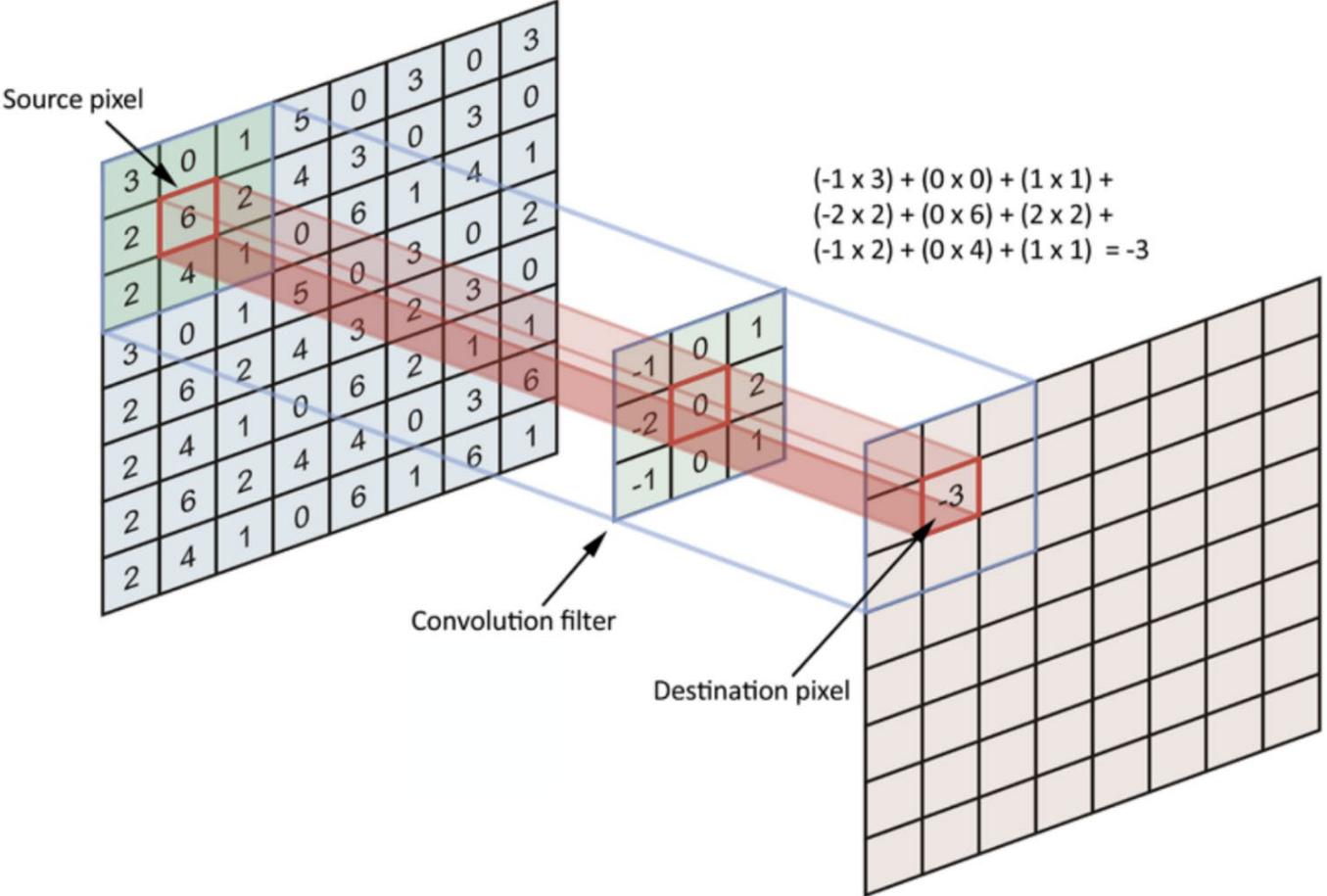


Image Processing

Latent Diffusion - Auto Encoder



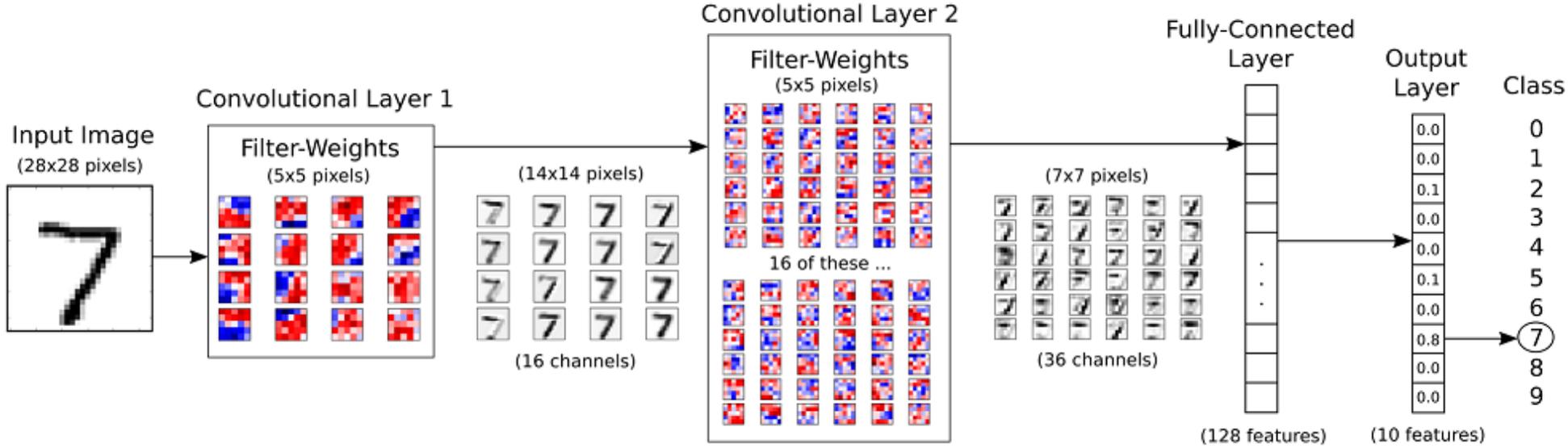
Latent Diffusion - Kernel Convolution



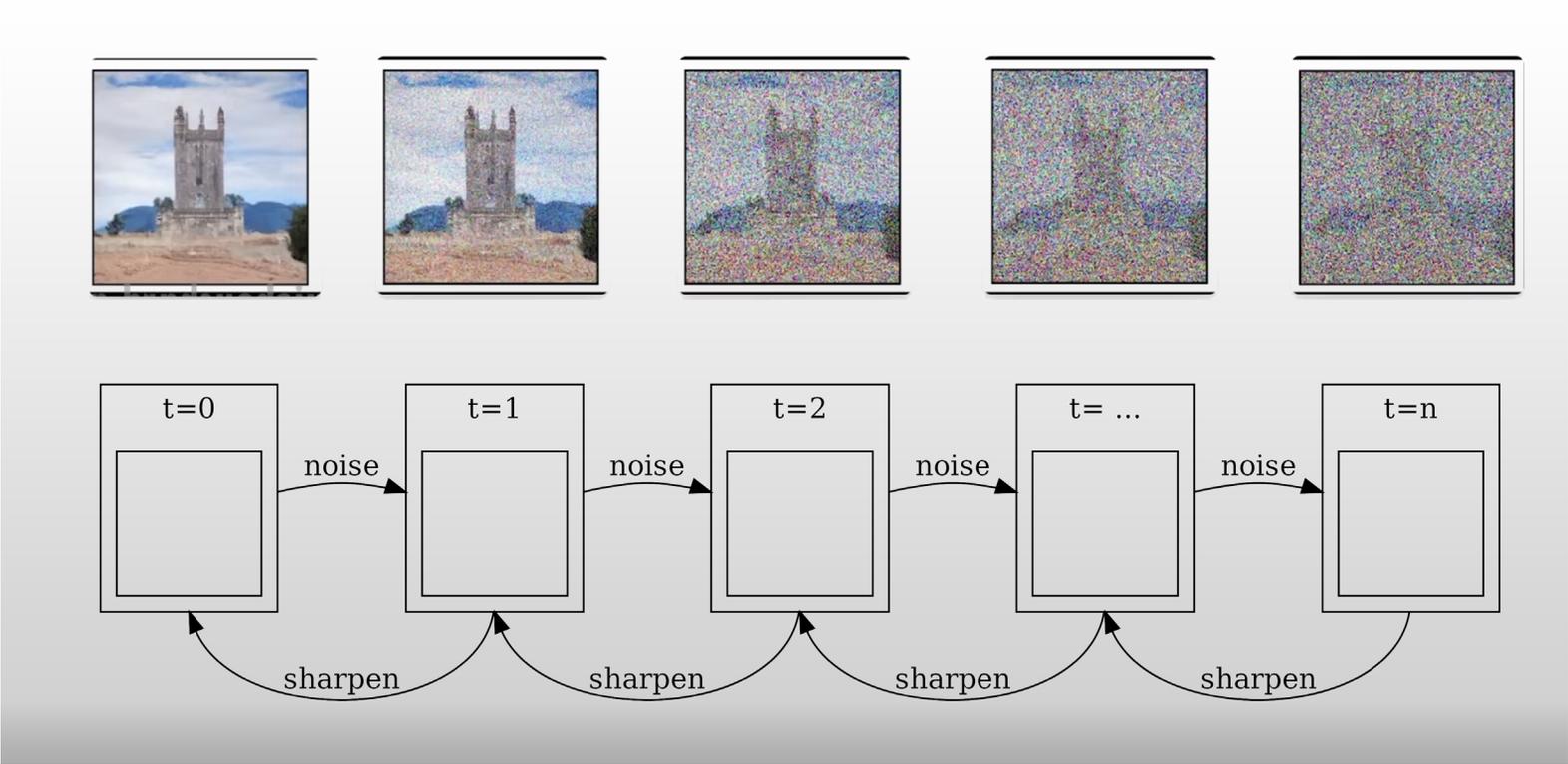
Latent Diffusion – Common Kernel Examples

Operation	Kernel ω	Image result $g(x,y)$
<u>Ridge</u> or <u>edge detection</u>	$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$	
Sharpen	$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$	
<u>Box blur</u> (normalized)	$\frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$	
<u>Gaussian blur</u> 3 × 3 (approximation)	$\frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$	

Latent Diffusion - Kernel Convolution on MNIST Dataset

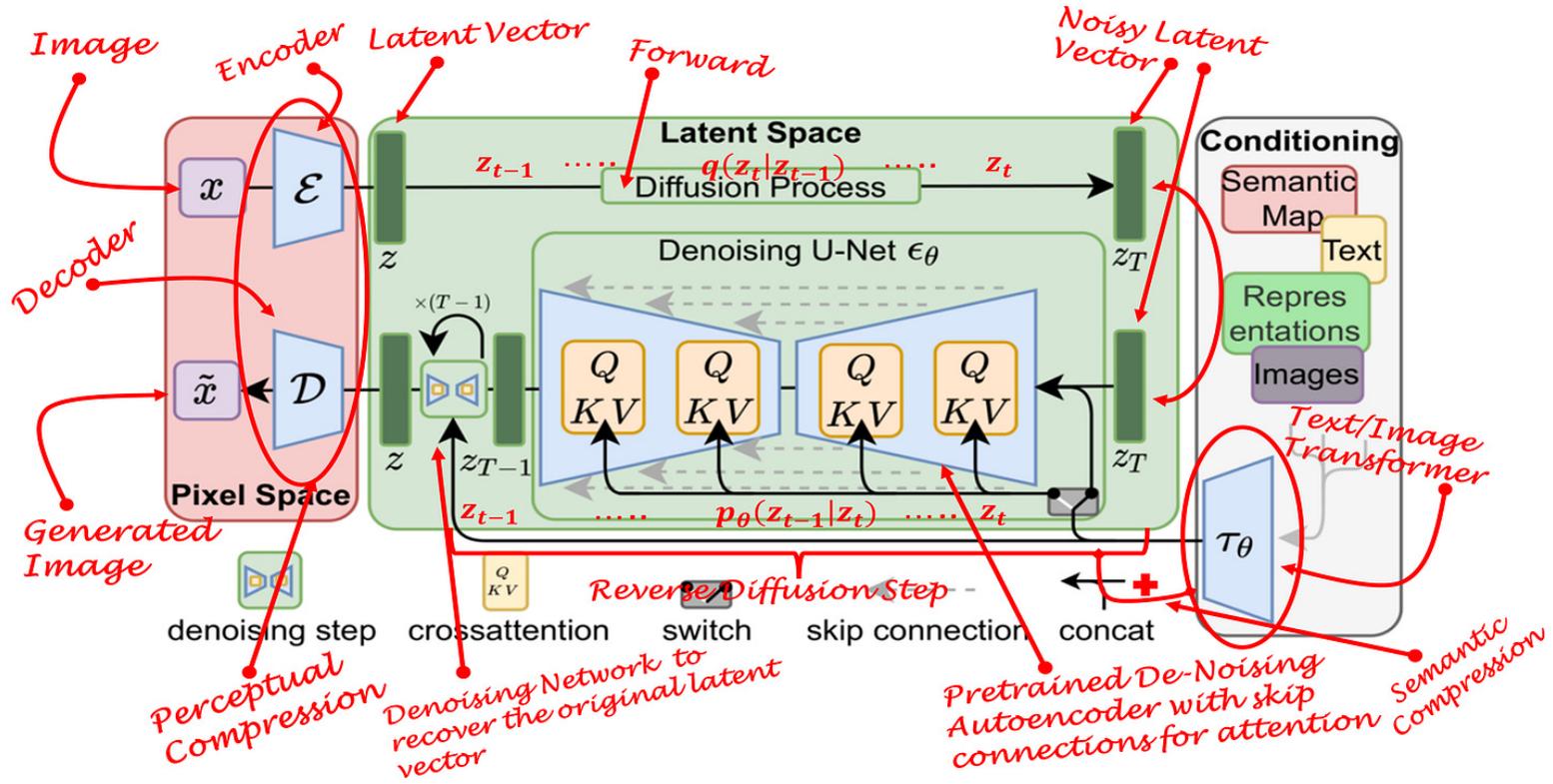


Latent Diffusion – Concept

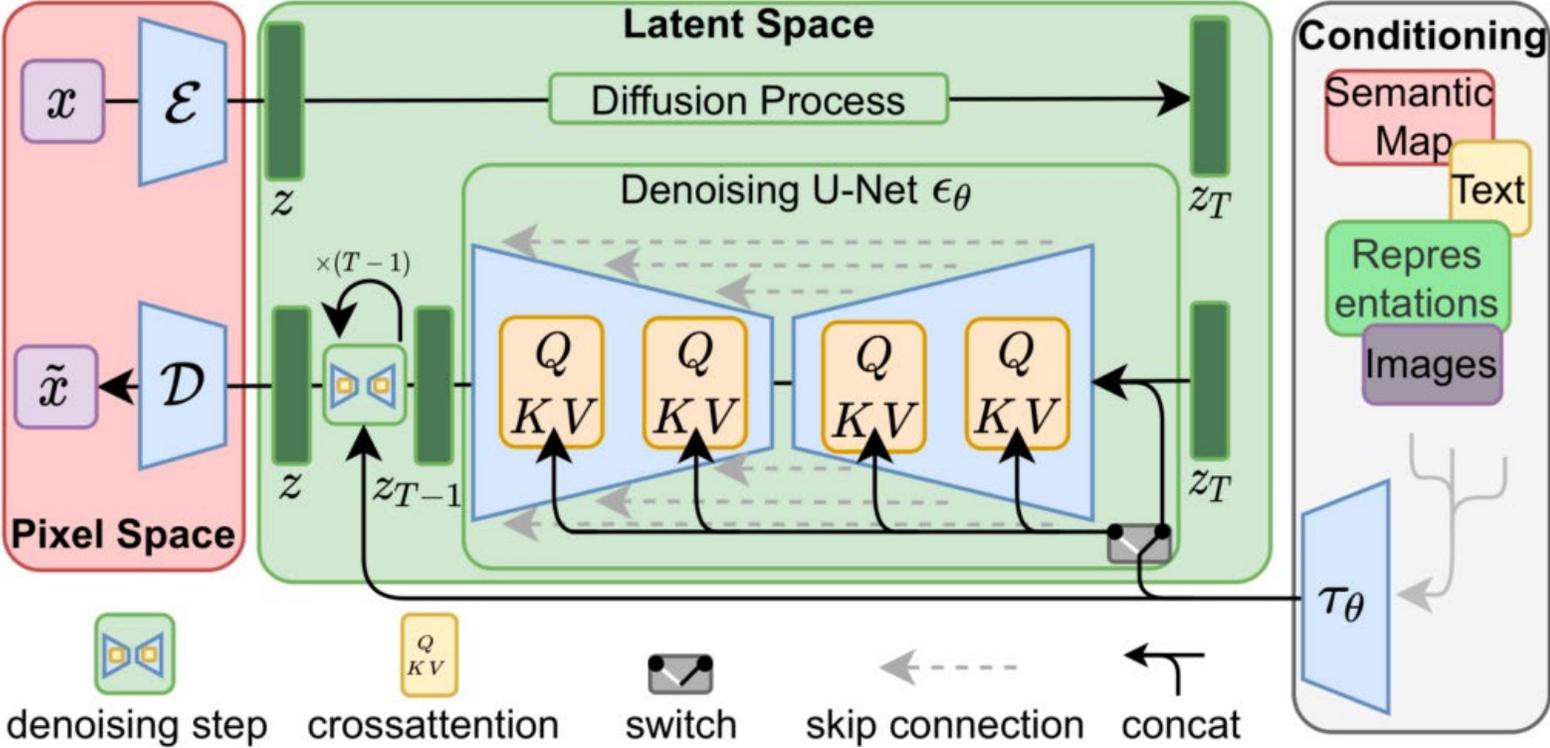


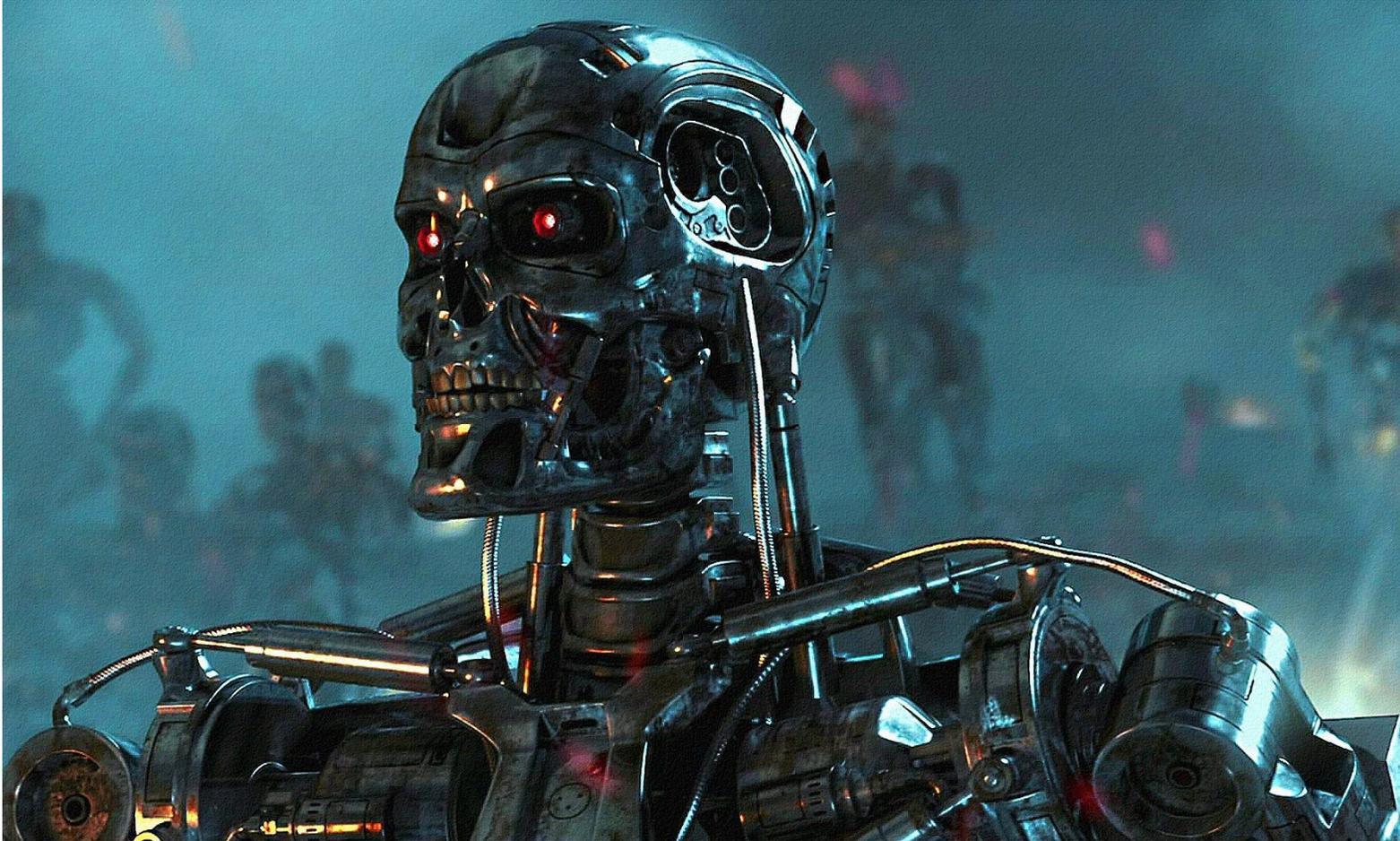
A model is trained to progressively sharpen an image

Latent Diffusion - Stable Diffusion Model



Latent Diffusion - Stable Diffusion Model

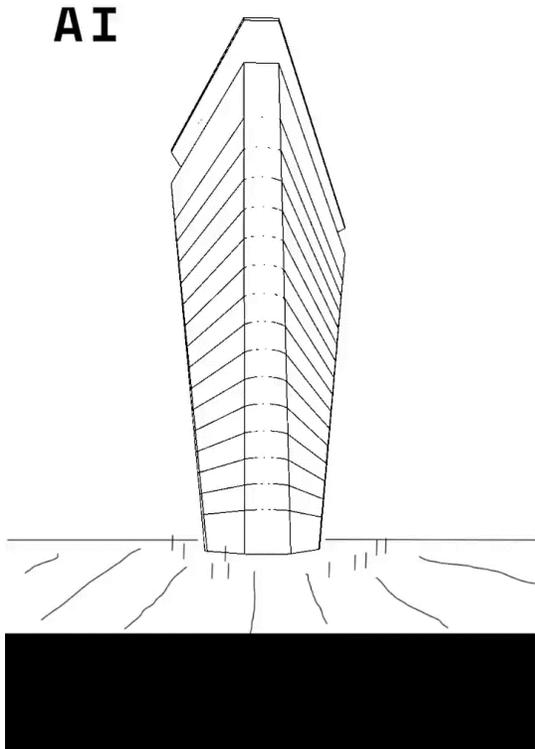




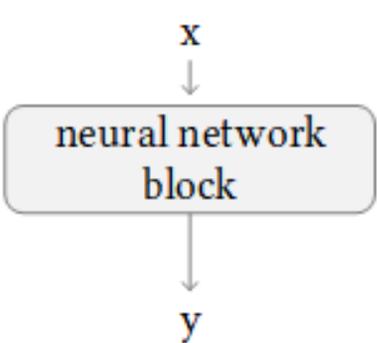
The Terminator (1984), Terminator



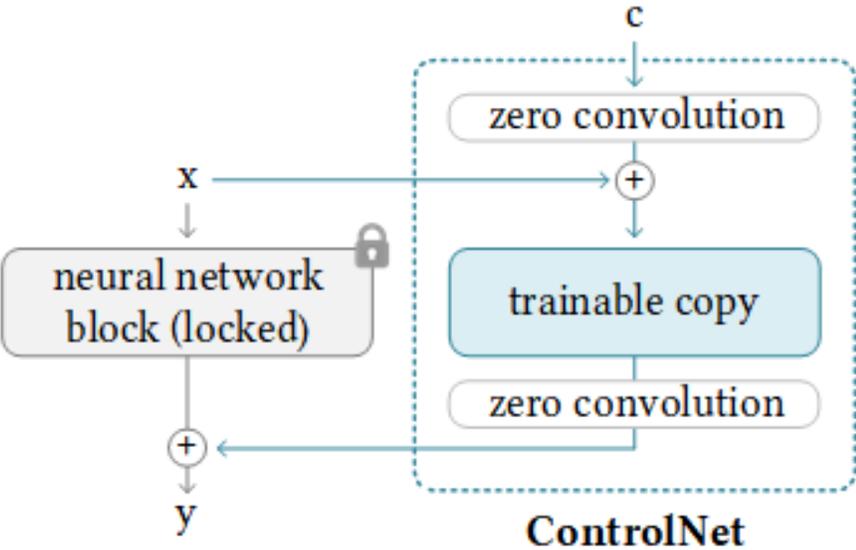
Introduction of ControlNet to help steer the image generation



ControlNet



(a) Before



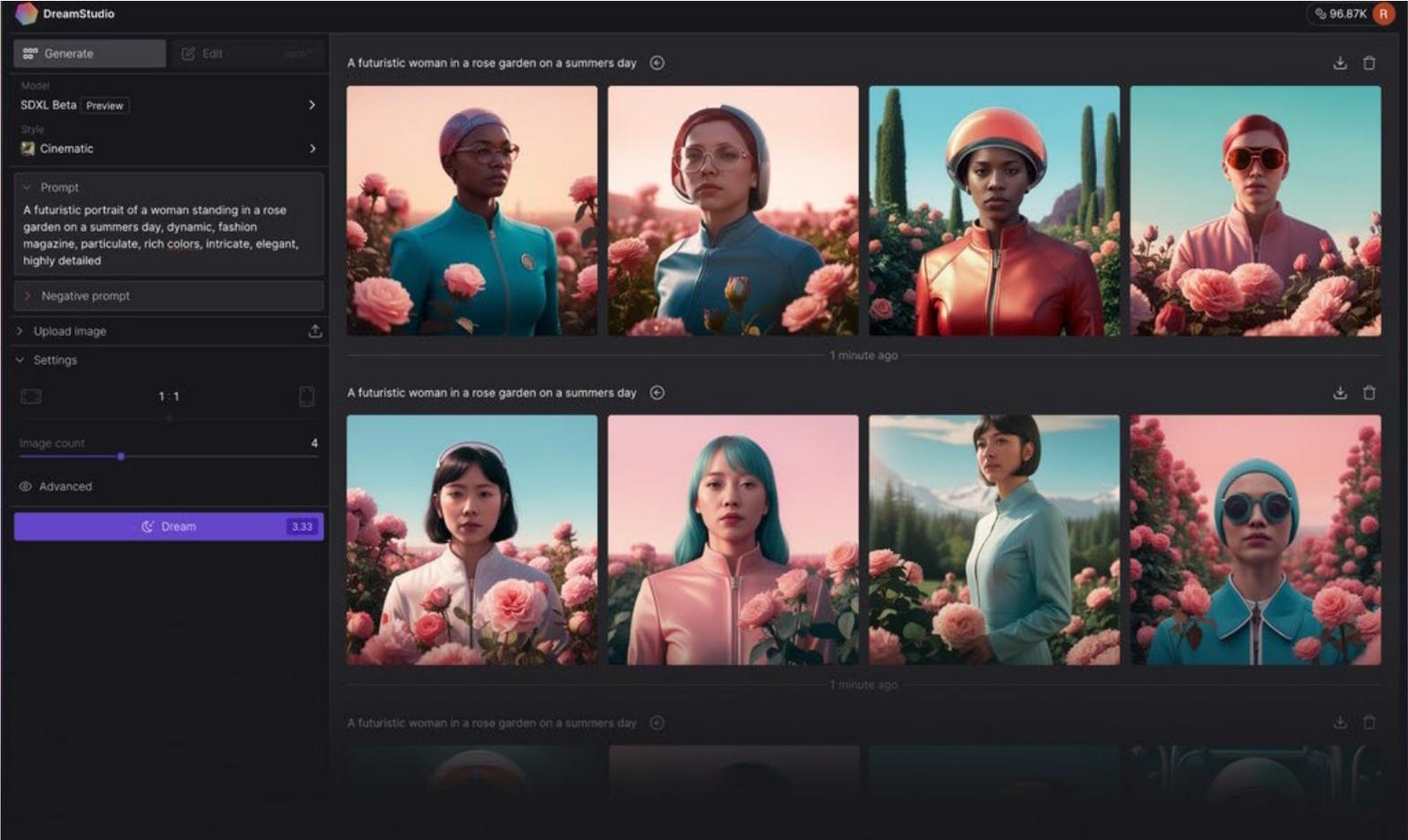
(b) After

ControlNet is a neural network structure to control diffusion models by adding extra conditions.

Applications in Practice

Machine Learning and Architecture

Dreamstudio by stability.ai



Mindjourney



The Pope Drip u/trippy_art_special, r/midjourney
https://www.reddit.com/r/midjourney/comments/120vhdc/the_pope_drip/

Mindjourney & Architecture



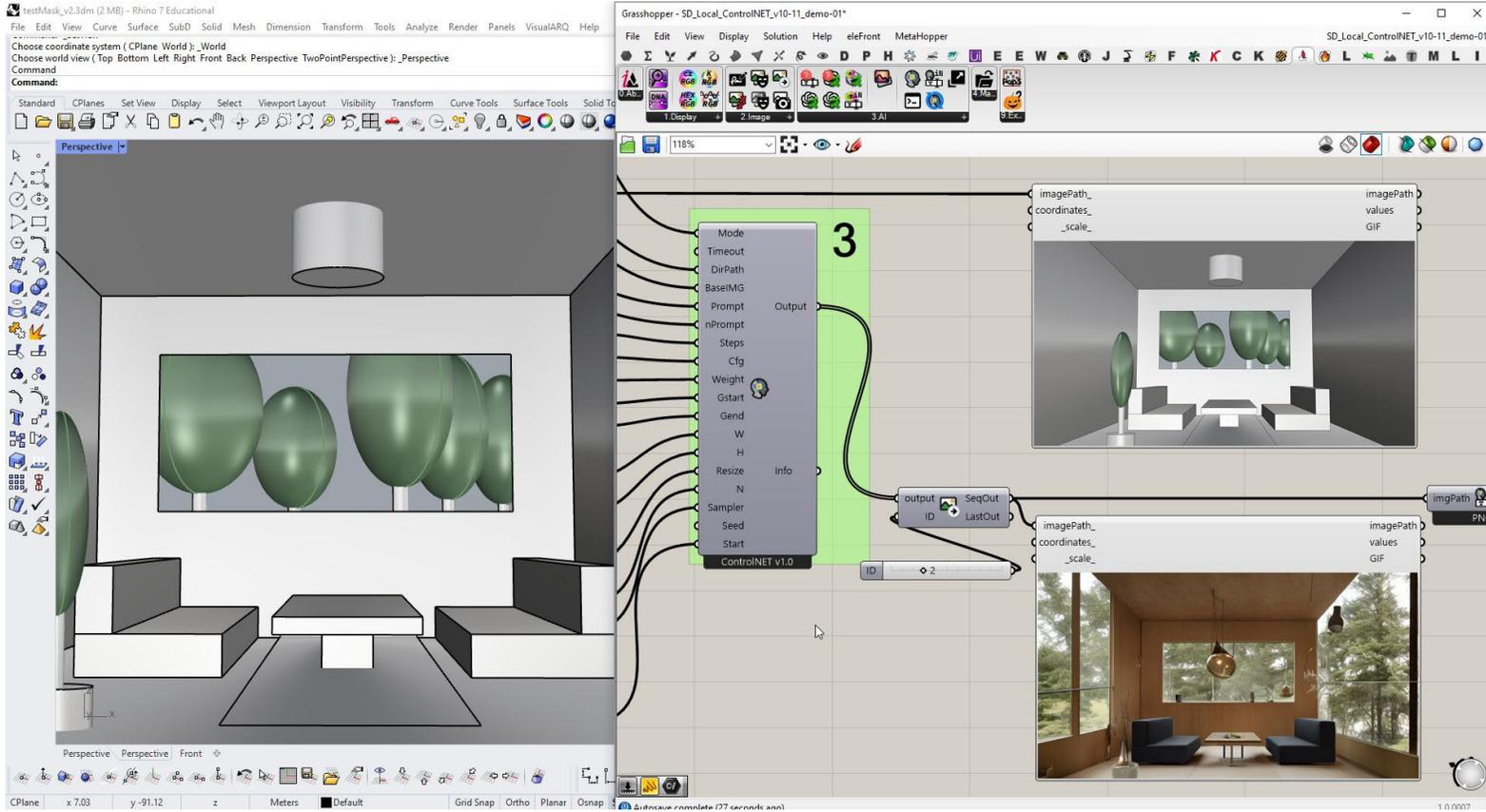
Hassan Ragab, images created with midjourney
<https://www.instagram.com/hsnrqb/>

Mindjourney & Architecture



Edward Crump, images created with midjourney
<https://www.instagram.com/edthearch/>

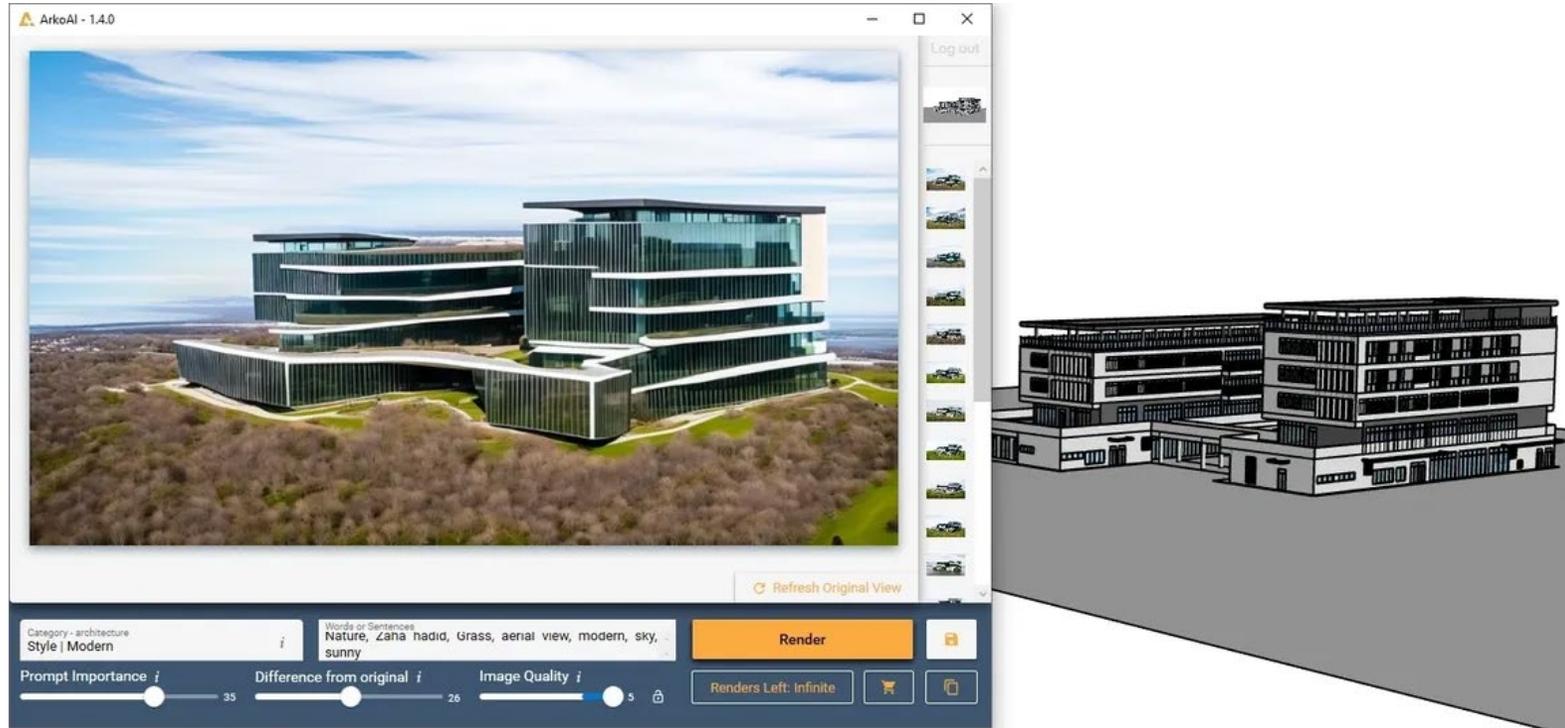
Ambrosinus-Toolkit



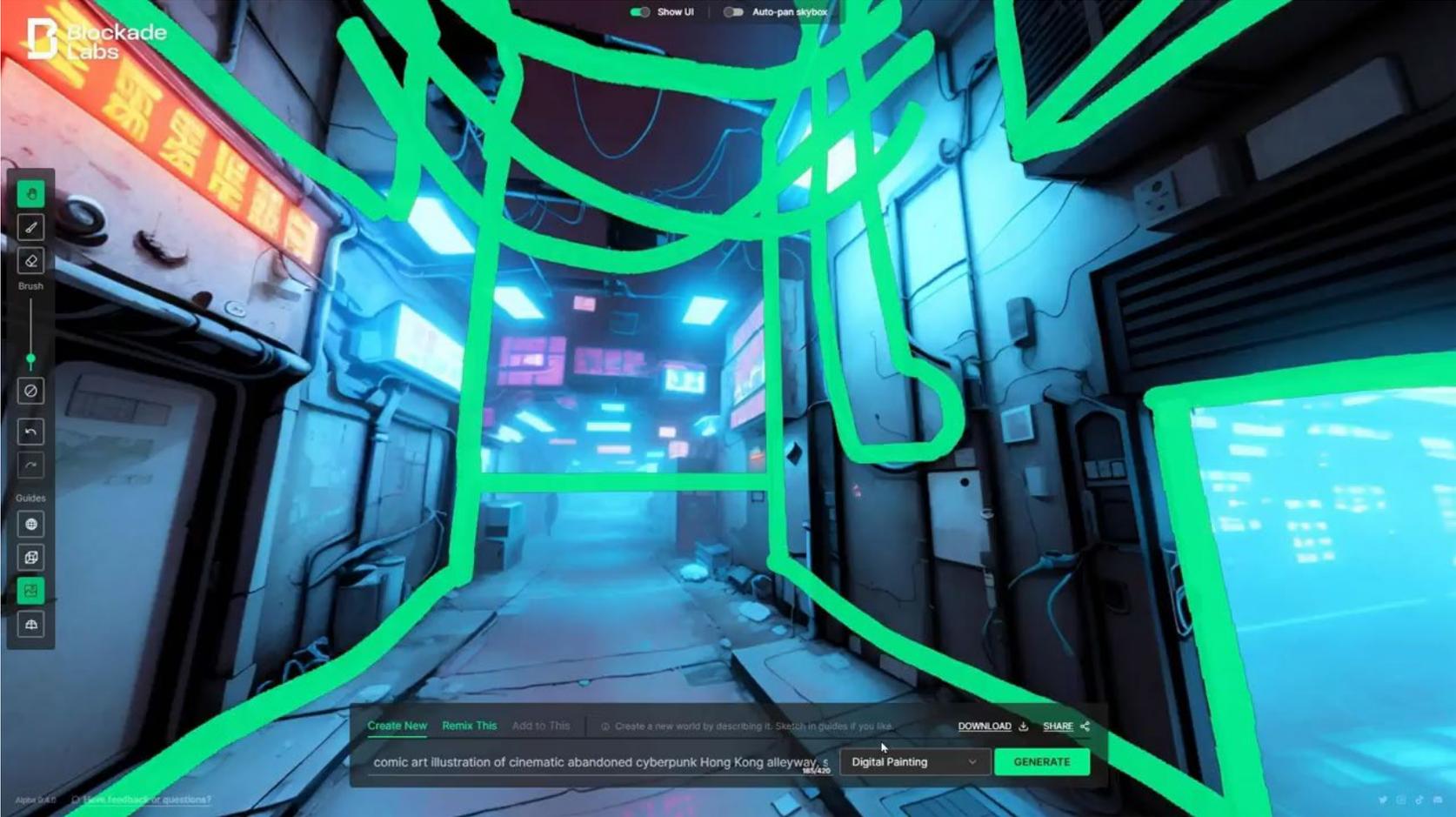
https://www.linkedin.com/posts/lucianoambrosini_ambrosinustoolkit-automatic1111-controlnet-activity-7059777530538467329-vwb1/?utm_source=share&utm_medium=member_ios



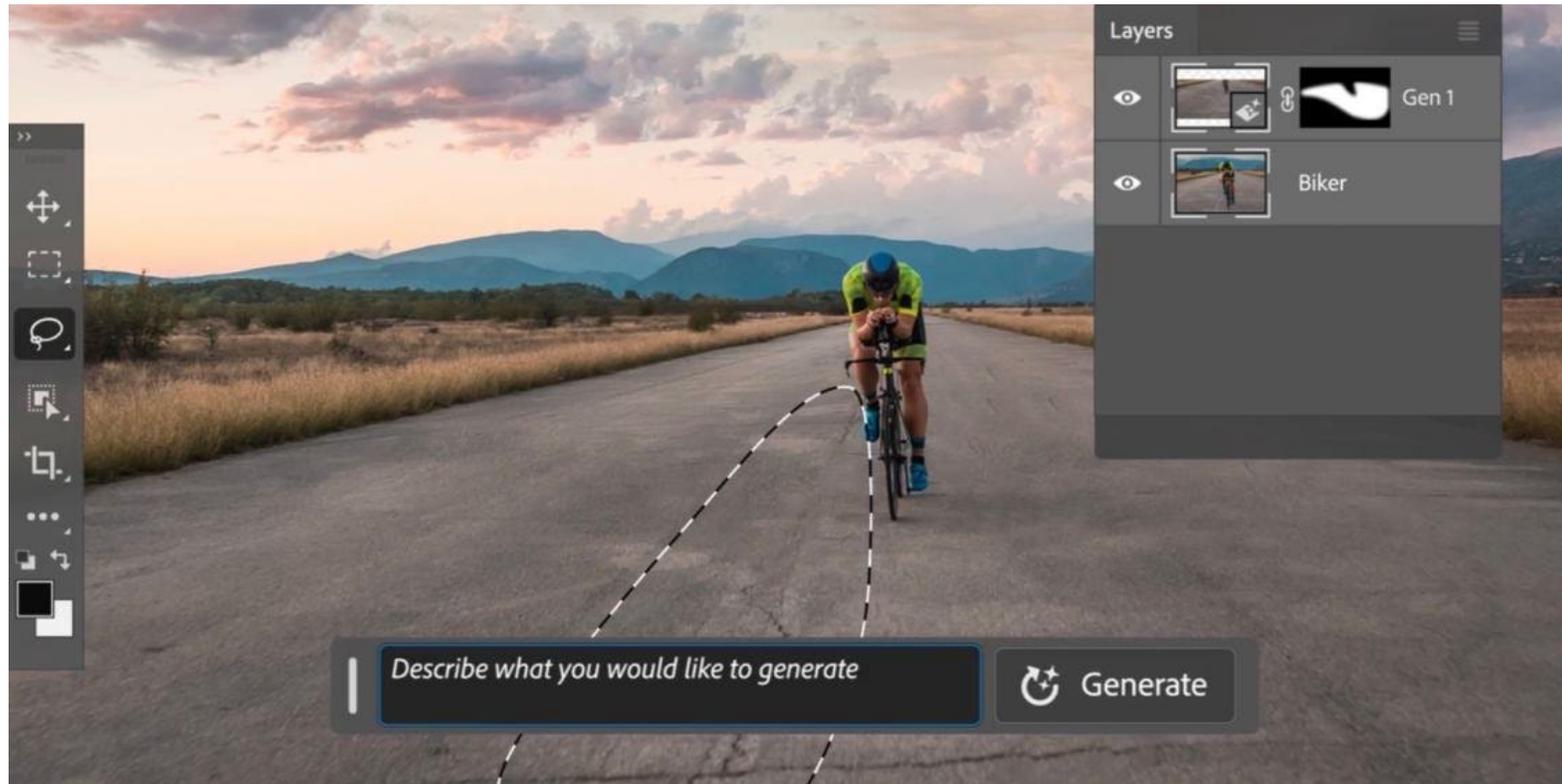
ArkoAI



Blockade Labs – 3D sketching



Adobe Photoshop – New AI tool



Finch 3D

The screenshot displays the Finch 3D software interface. On the left, a sidebar contains several sections: 'CREATE LIBRARIES FILL', 'TOOLS' with drawing icons, 'UNITS' with a table of metrics, 'SELECTION TAGS' with a 'Residential' tag, and 'SELECTION SPACES' with 'Livingroom', 'Bedroom', and 'Hall' tags. The main workspace shows a 2D floor plan of a building with a central staircase. On the right, a 'PARALLEL AXIS DIAGRAM' is visible with a table of optimization parameters. Below this, a 'PLANS GENERATED: 0' section includes a 'Select Property' dropdown and a 'Run Optimization' button.

Area	Energy	Cost	CO2-Efficiency	Gridlock	Occupancy	Fitness	Material	Connections	Design

SELECTION TAGS	Area
Residential	895 m ²

SELECTION SPACES	Area
Livingroom	68 m ²
Bedroom	65 m ²
Hall	56 m ²

UNITS

Inner wall width	120
Count	7
Area Total	232 m ²
Area Range	56-73 m ²

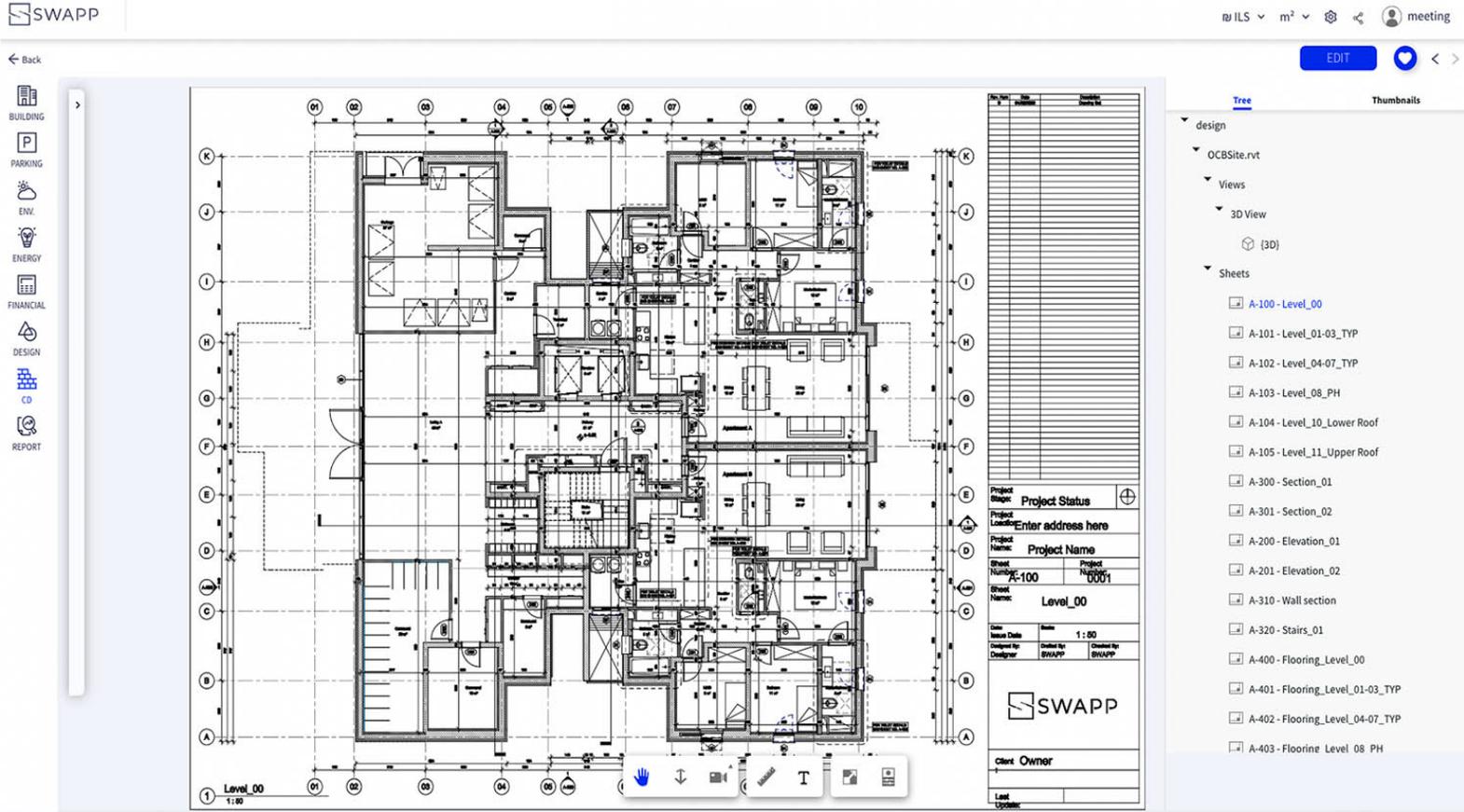
PLANS GENERATED: 0

Select Property

Run Optimization

Select Property

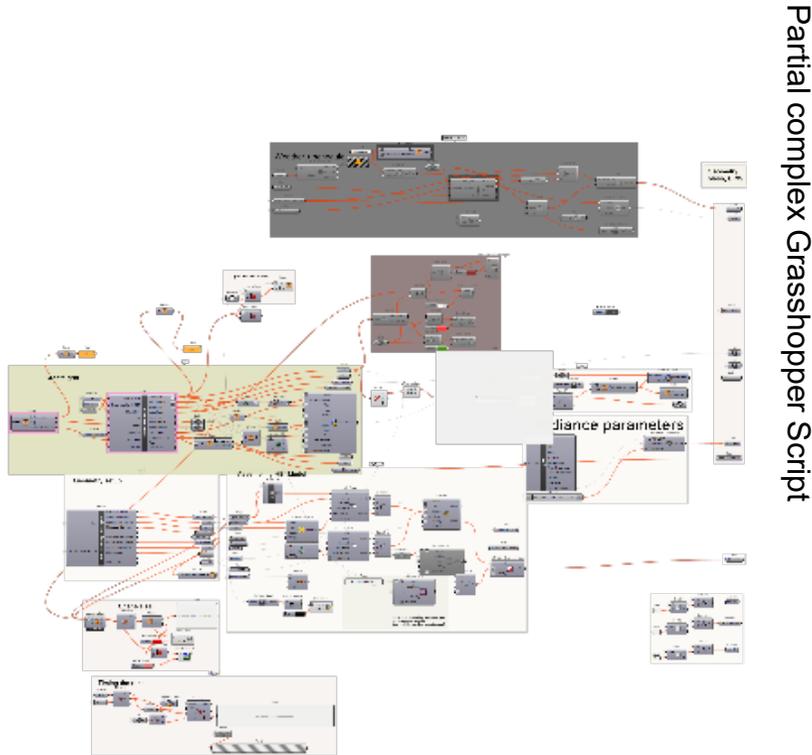
SWAPP



Swapp can generate to full sets of construction documents



LLM' as Humane to Machine Interface



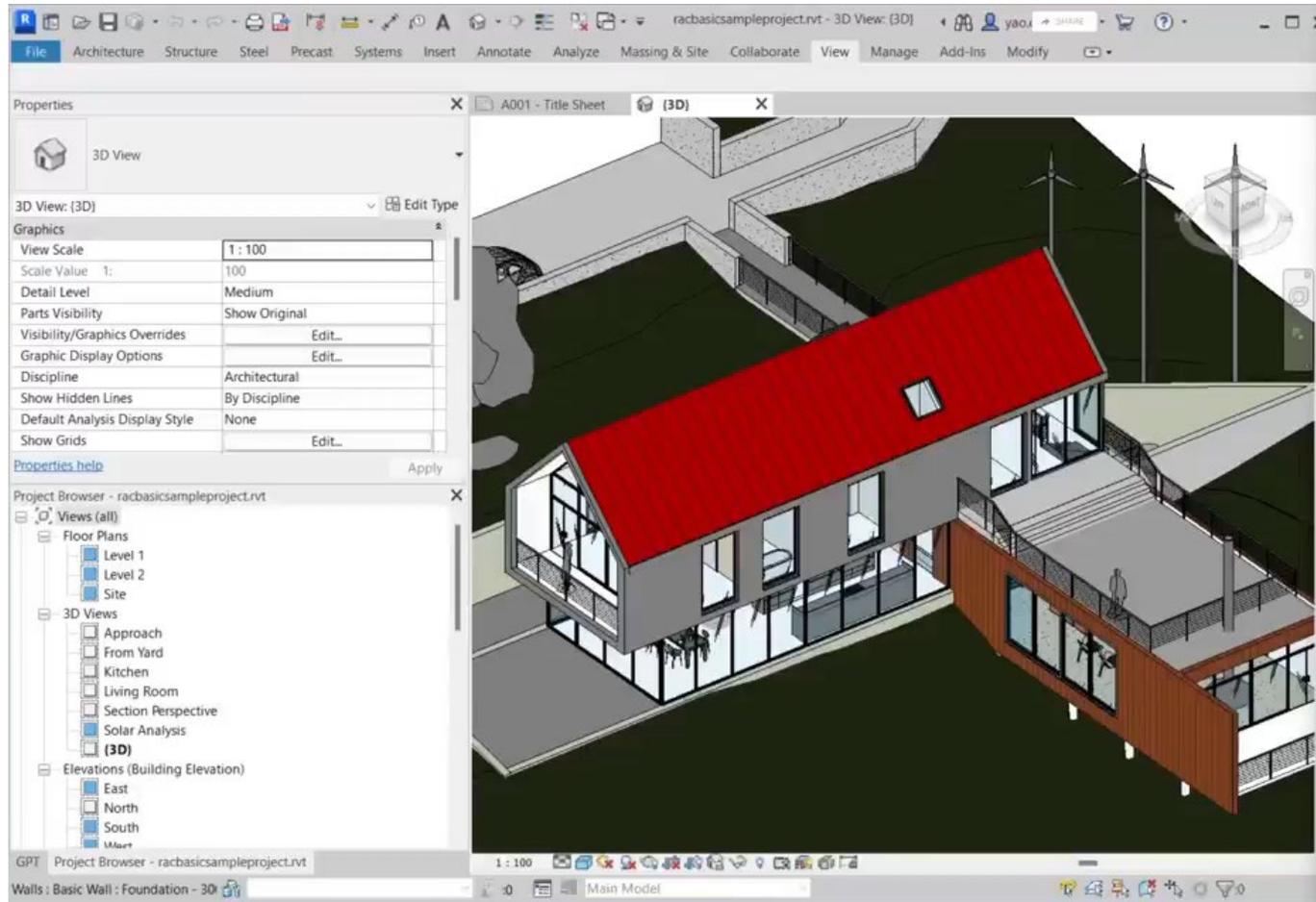
One potential use case of LLM might be as human to machine interfaces.

- In correcting out error prone inputs.

$$\left. \begin{array}{l} \frac{væg}{v\ddot{a}gg} \\ \frac{concrete\ 300mm}{} \end{array} \right\} \Rightarrow \text{wall}$$

- By translating intent into code

Yao Chen – ChatGPT in Revit



SUB-LAUNCH DETECTION DEFENSE ANALYSIS
UNITED STATES OF AMERICA
INIT SPEED SEQ 32-9445

ON STA WLD

Promoting the reuse of vacant buildings using machine learning

Machine Learning and Architecture

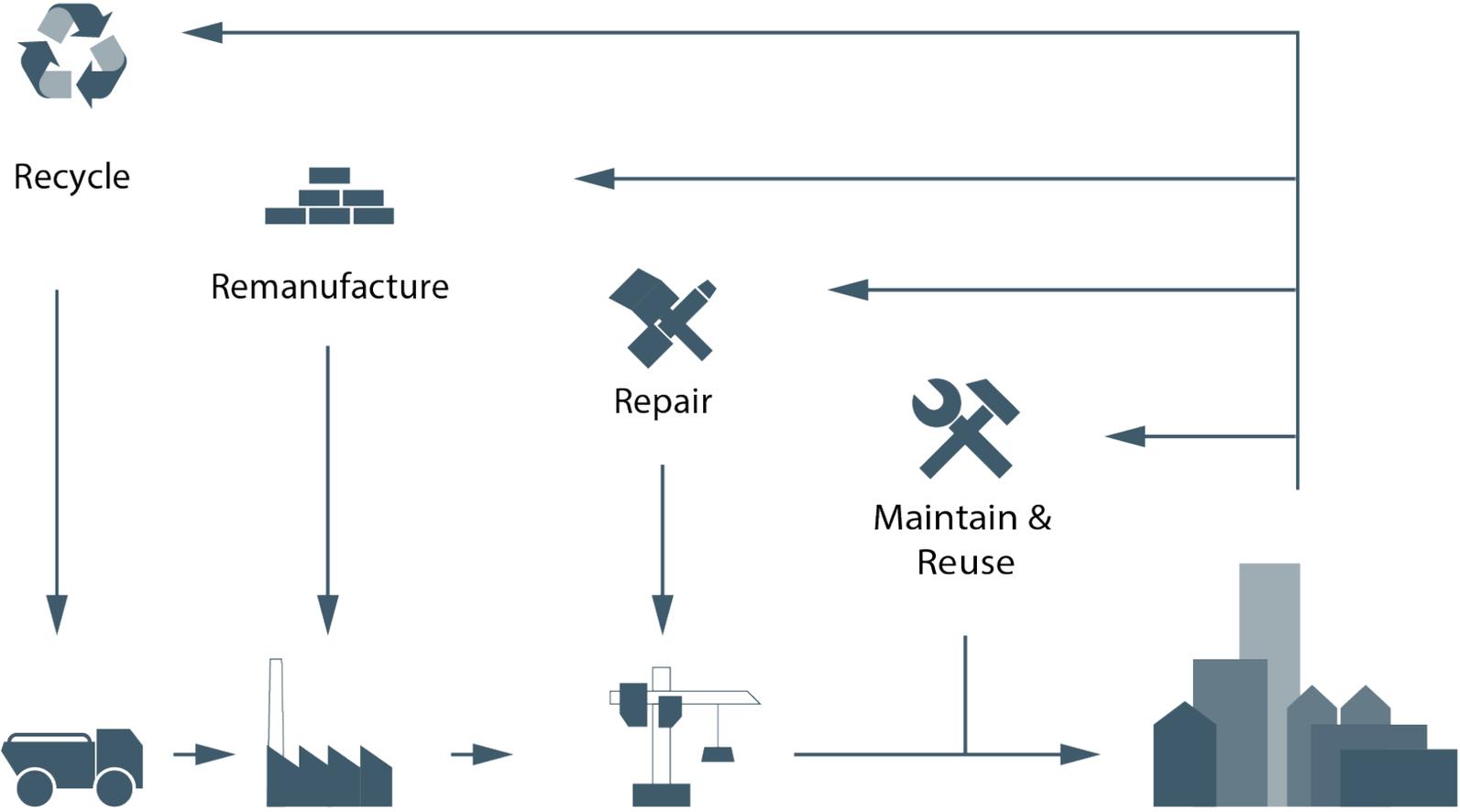
ENR 5 45-55 CONF 70
ENR 6 25-35 CONF 40
TRANSMISSION WINDOWS

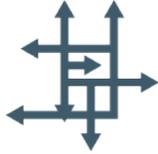
1st phase
CAPABILITY

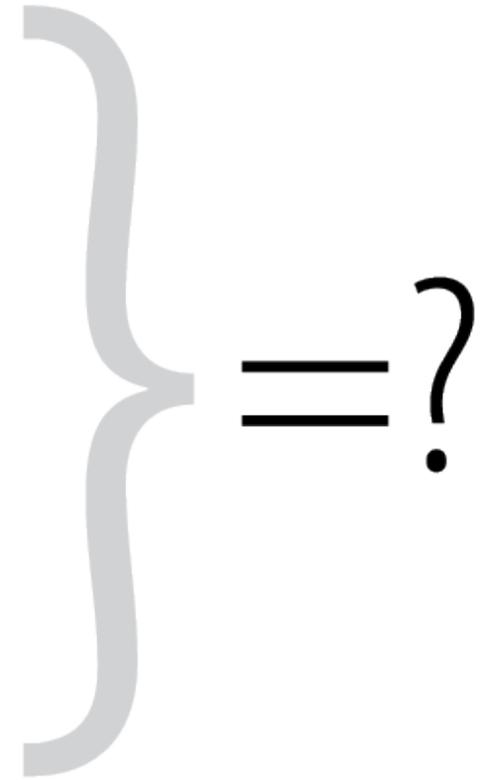
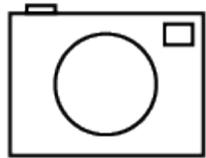
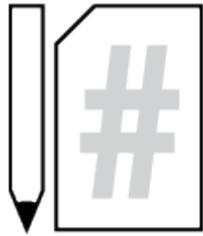
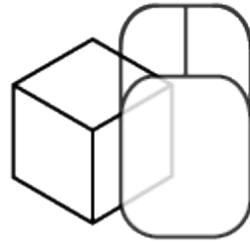
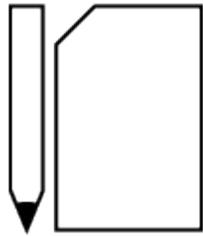
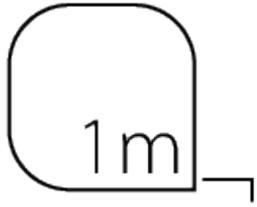


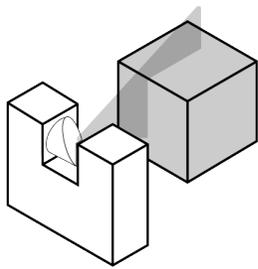




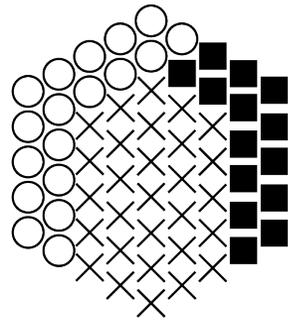








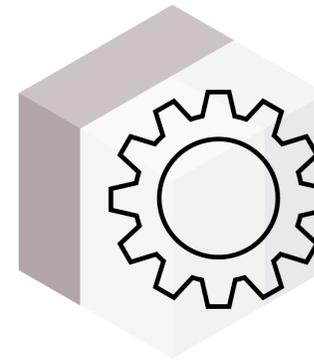
scanning



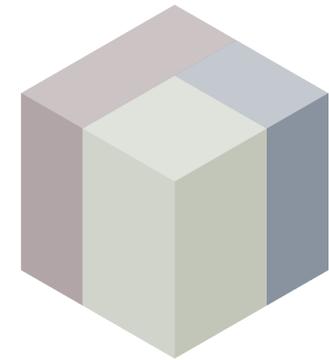
segmentation



mapping

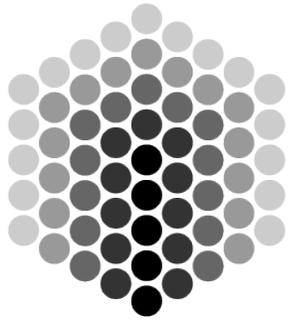


assessing

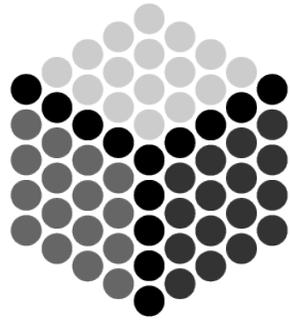


presenting





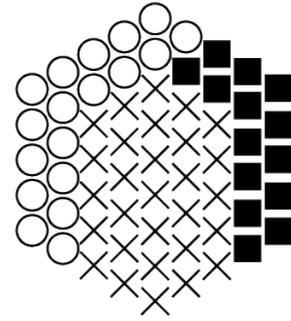
data



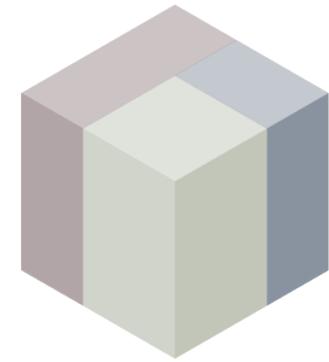
structure



colors

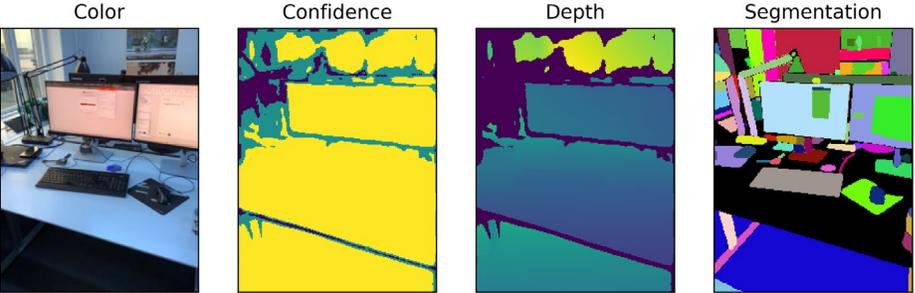


labeled



rationalized

RGB-D Image Semantic Segmentation and Point Cloud Reconstruction

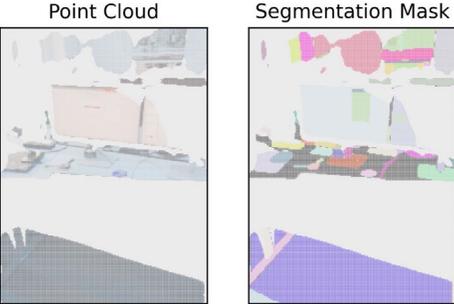


Intrinsic Matrix

PinholeCameraIntrinsic with width = 256 and height = 192.
Access intrinsics with intrinsic_matrix.

Extrinsic Matrix

```
[[ 0.12052873 -0.89548501 0.4284617 0.25570811]
 [-0.9441382 0.02994955 0.32818605 -0.00481559]
 [-0.30671793 -0.4440829 -0.84185182 0.34062063]
 [ 0. 0. 0. 1. ]]
```

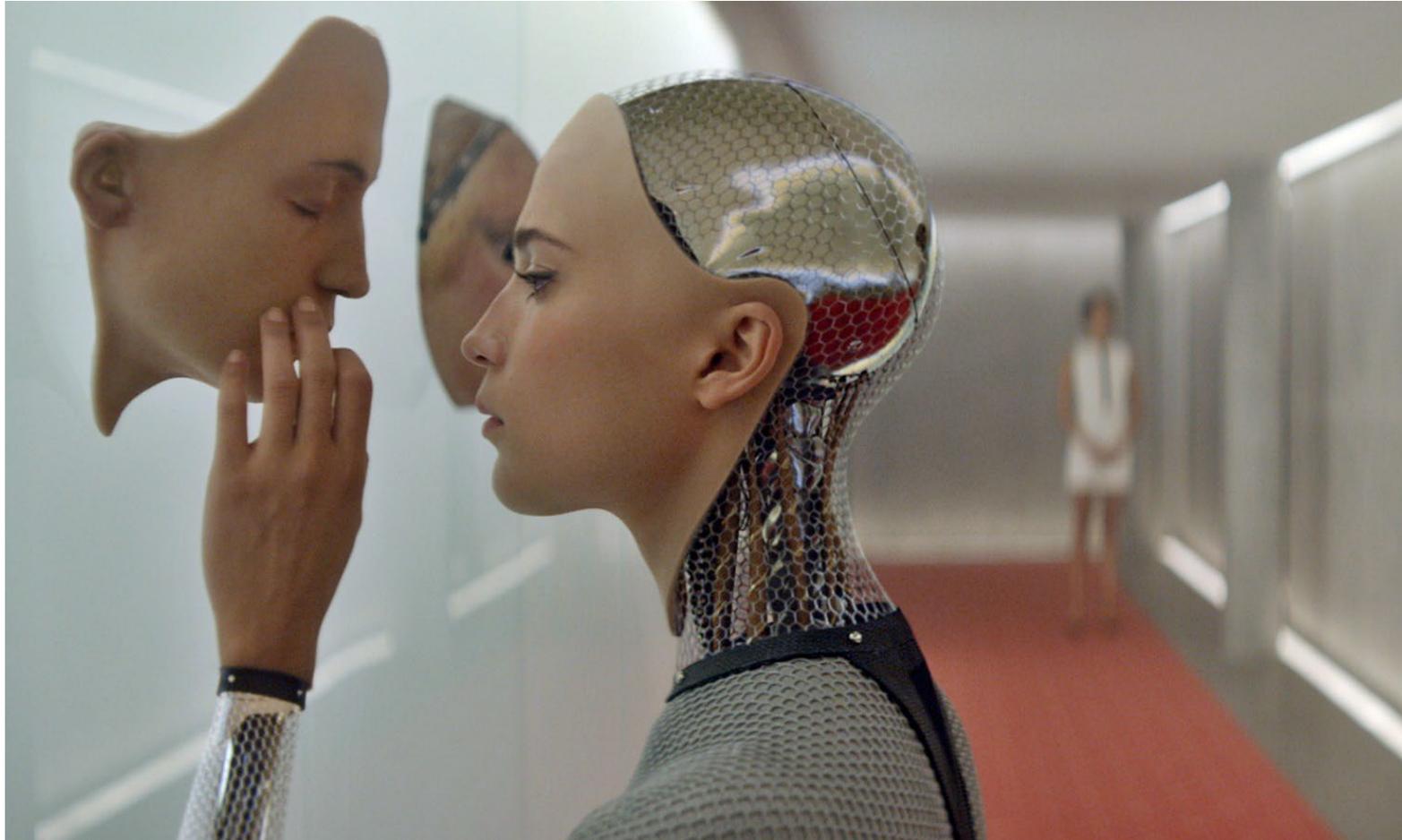


Visualization of data obtained from an iPad, with segmentation applied using Segment Anything (SAM)

RGB-D images can be used to generate a segmented point cloud, if images segmentation were to be applied prior to the projecting, it into 3D space.



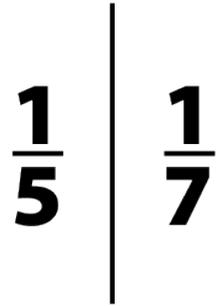
Motivation or why should we as architects care



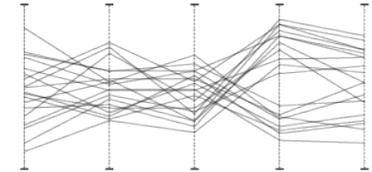
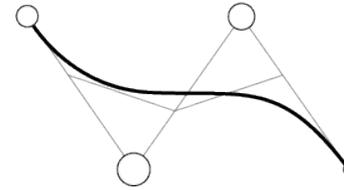
Ex Machina (2015), Ava

Greg Lynn

Mario Carpo



.75



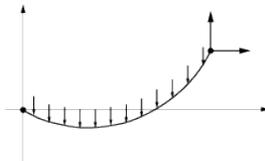
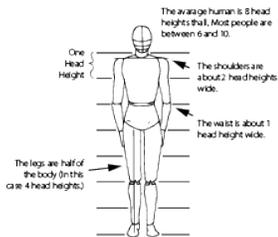
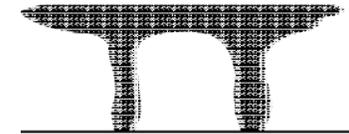
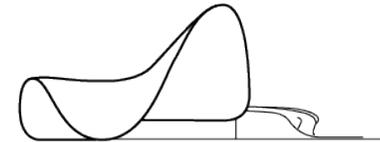
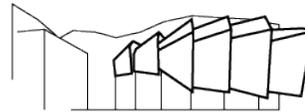
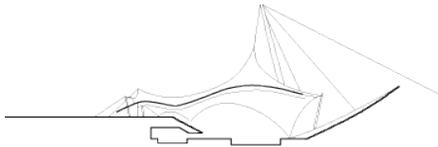
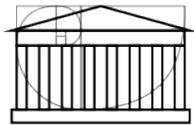
Proportion

Forces

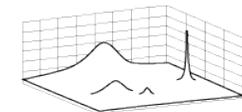
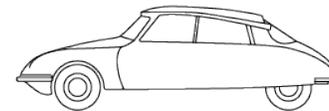
Animation

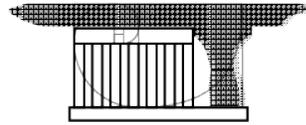
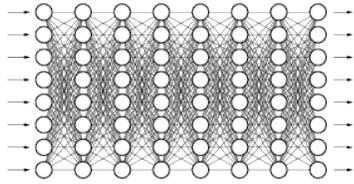
NURBS

Simulation

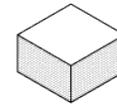
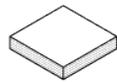


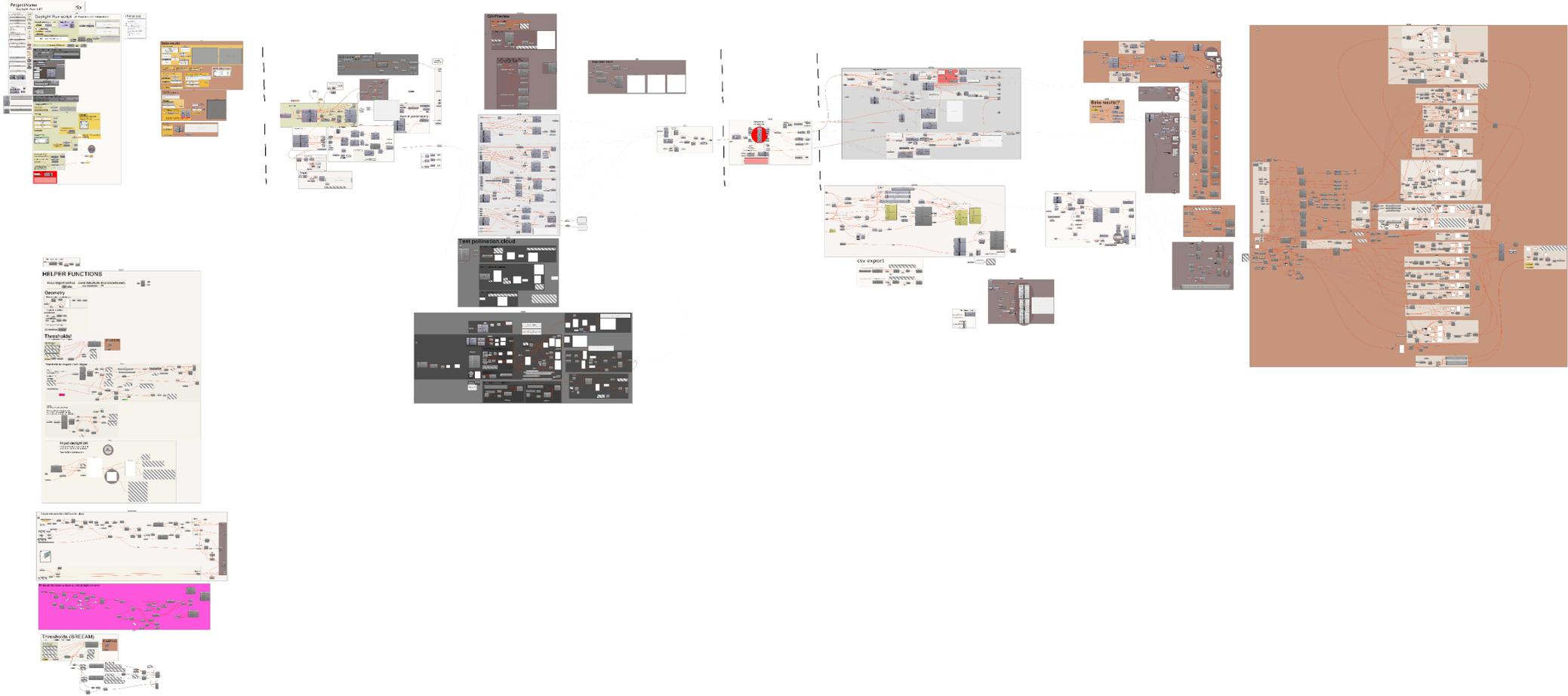
$$\int_a^b f(x)dx = F(b)-F(a)$$

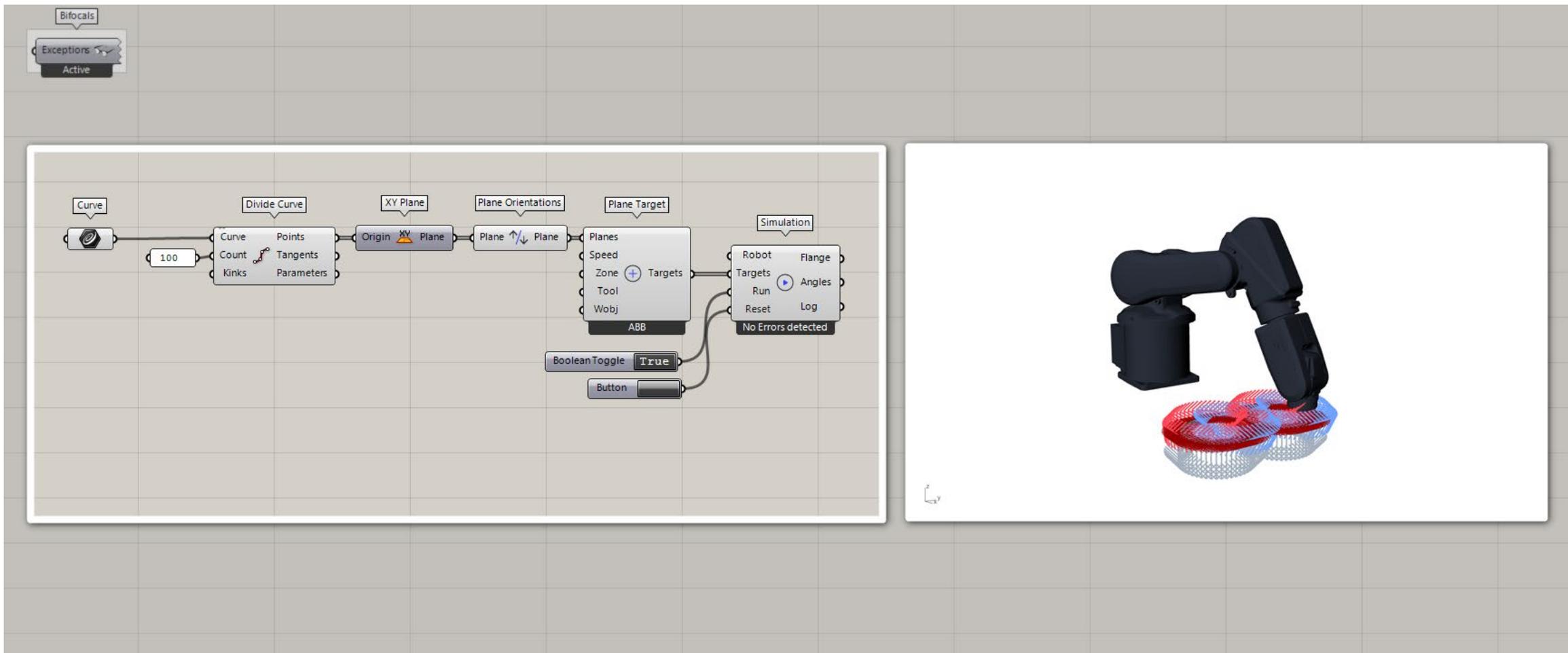




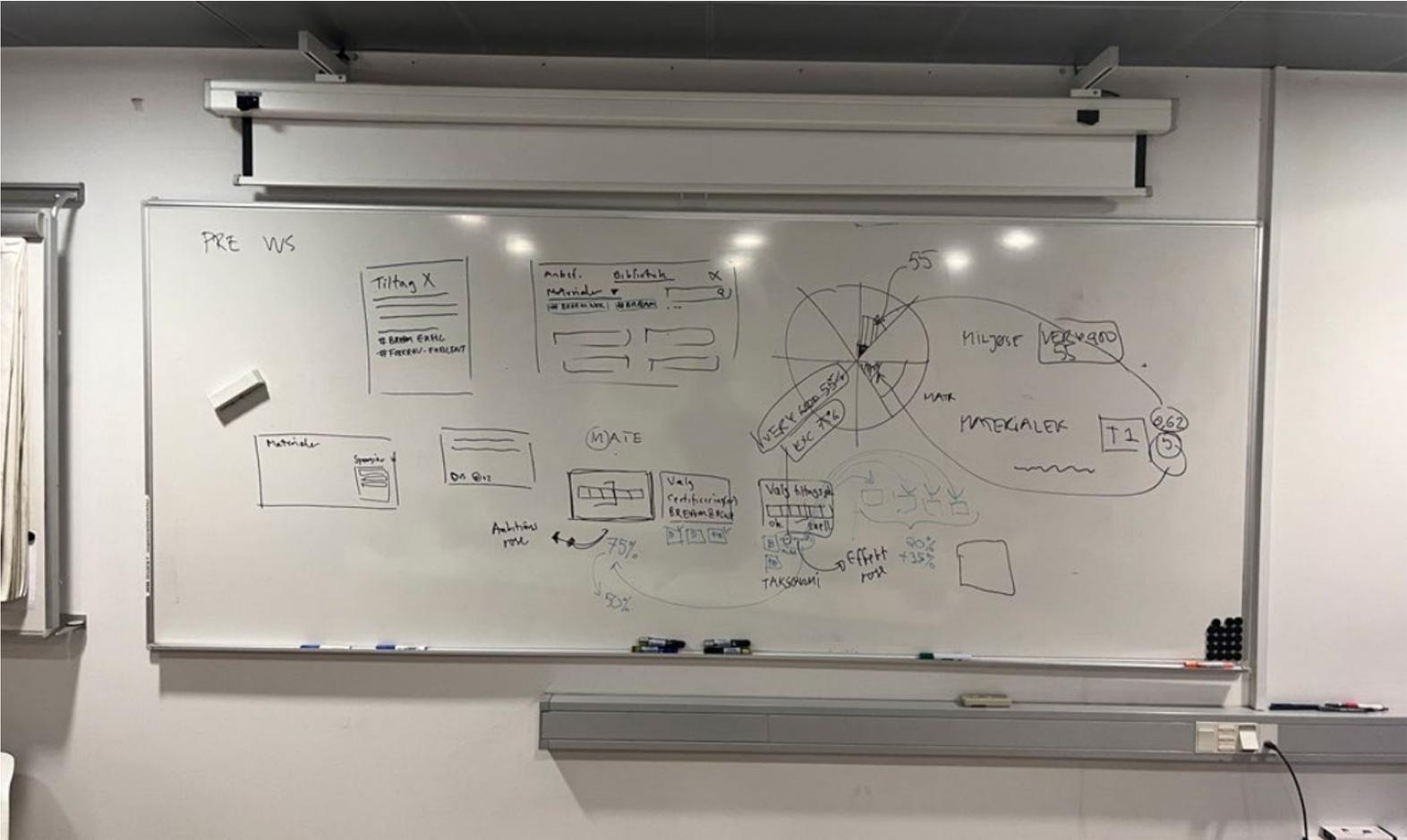
$$J(\theta) = \frac{1}{m} \sum_{i=1}^m [-y^{(i)} \log(h_{\theta}(x^{(i)})) - (1 - y^{(i)}) \log(1 - h_{\theta}(x^{(i)}))]$$







LINK Kompass®



LINK Kompass®

Platform Demo

LINK Kompass®

WORKSHOPØKT 1 Prosjektvisjon 2 3 4

Oversikt

PROSJEKTVISJON

- Vi ønsker å bygge fantastiske boliger på klodens premisser.
- Her må det være mulig å bo uten å eie bil.
- Prosjektet skal være CO2-nøytralt.
- Å bo her skal føles som å være i et feriehus.

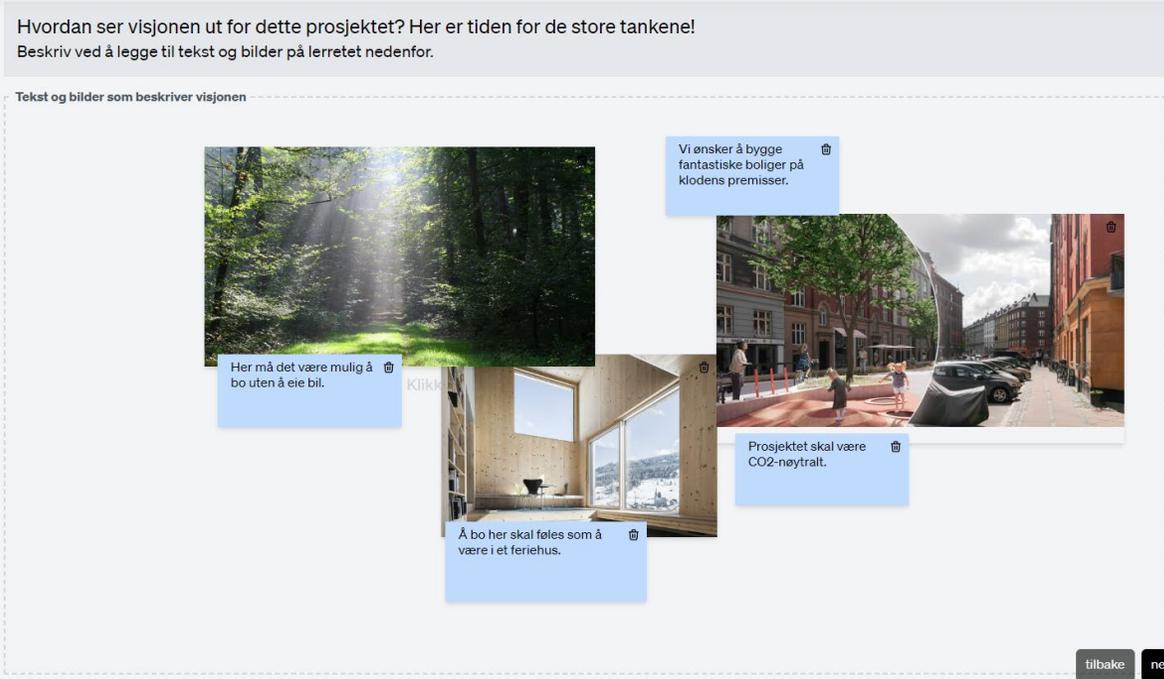


RAMMEVERK
LINK Kompass



Hvordan ser visjonen ut for dette prosjektet? Her er tiden for de store tankene!
Beskriv ved å legge til tekst og bilder på lerretet nedenfor.

Tekst og bilder som beskriver visjonen



Vi ønsker å bygge fantastiske boliger på klodens premisser.

Her må det være mulig å bo uten å eie bil.

Prosjektet skal være CO2-nøytralt.

Å bo her skal føles som å være i et feriehus.

tilbake neste

workshop (1).docx • Saved to this PC

Insert Draw Design Layout References Mailings Review View Help Acrobat

Font Paragraph Styles Editing Adobe Acrobat Voice Sensitivity Editor Reuse Files

LINK ARKITEKTUR

Notat

Bærekraftsscreening med LINK Kompass® 20.05.2023

Prosjekt Seminar
Kunde [navn på kunde]

1. Om notatet
Dette notatet inneholder resultater fra bærekraftsscreening av prosjektet Seminar gjennomført den 20.05.2023 22. Screeningen ble utført som en workshop der LINK Kompass ble brukt som dialogverktøy av prosjektets sentrale miljø- og bærekraftambisjoner. Formålet med workshopen var å:

[Sett inn formål med workshop]

LINK Arkitektur sin strategi for bærekraftig prosjektering innebærer å ta et overordnet grep for bærekraft fra første strek. LINK Kompass® ble brukt i workshopen som et dialogverktøy for å diskutere og identifisere de sentrale bærekraftkvaliteter, målsettinger og tiltak for prosjektet Seminar. Tidligfase arbeidsmetode-dikk med LINK Kompass® skal involvere og engasjere alle prosjektets interessenter for å få frem konkrete bærekraftige løsninger som skal være styrende for prosjektets design. Det er viktig å velge en omforent og diskutert med alle sentrale interessenter i prosjektet, og særlig at de samsvarer med de kvaliteter som byggherre ønsker at prosjektet skal kjennes for.

Kompasset inneholder 14 bærekrafttema med en rekke undertema. Samlet dekker de 14 tema en screening av prosjektets økonomiske, sosiale- og miljømessige bærekraftprofil. Alle temaer ble diskutert under workshopen. Notat oppsummerer først forslag til overordnede målsettinger og samlet resultat fra bærekraftsscreeningen som kan/ bør være førende for byggets bærekraftambisjoner. Deretter gis det forslag til mål og tiltak for hvert enkelt tema som ble diskutert. **Temaer som ble diskutert er vist i f. u. g.** under.

Miljøstyring	Klimastyring	CO ₂	Ervervsnett	Materialer	Utslepp	Bløttemater
Struktur og kvalitet	Transport og miljø	Trykkløst og trykkløst	Parallell og trykkløst	Nett og trykkløst	Nett	Nett

73 words Nonwegian (Bokmål) Accessibility: Investigate Focus

LINK ARKITEKTUR

2. Prosjektorganisering, ledelse og oppfølging av miljøprogram

Prosjektinformasjon
[Fyll ut relevant prosjektinformasjon]

Prosjekt navn:
Byggherre:
Prosjektnummer: null
Bygg type:
Areal:

Organisering av miljøarbeidet i prosjektet
[Fyll ut informasjon om organisering av miljøarbeidet i prosjektet dersom aktuelt]

Ledelse og oppfølging
Det skal utarbeides en samlet miljøkvalitets og oppfølging plan for prosjektet, kort betegnet «MOP». Miljøoppfølgingsplan skal utarbeides kompletteres og oppdateres gjennom hele byggeprosessen. Prosjektet skal i hele prosjekterings- og utførelsesfasen, utpeke en miljøansvarlig for prosjektet. Miljøansvar kan være tas av forskjellige personer i prosjektets faser, men det skal tilknyttet ansvarlig miljøansvarlig i alle faser.

Miljøansvarlig skal være ansvarlig for at prosjektet miljømål løpende følges. Status for miljøarbeidet rapporteres som del av den ordinære rapporteringen og miljø tas opp som tema på alle møter, herunder:

- Byggherremøter
- Prosjektmøter
- Byggemøter

Konsekvenser av eventuelle avvik fra Miljøprogrammet underveis i prosjektet skal forelegges og godkjennes av byggherre sin miljøansvarlig.

3. Forslag til plan og videre prosess.

Dette notatet sendes til deltakere i workshop for kommentering for supplering, og rettelser av eventuelle misforståelser. Forslag til videre prosess er som følger.

[Følgende forslag til tekst slettes eller suppleres ved behov]

Trinn 1 (uke xx)
Etter kommentering utarbeides det overordnede målsettinger og konkrete tiltak for byggets bærekraftkvaliteter.

Trinn 2 (uke xx)
Det utarbeides forslag til overordnede og delmålsettinger for prosjektet. Forslag fremsendes til byggherre, og går gjennom og diskuteres på et følgende møte uke 47. Overordnede- og delmålsettinger, tiltak og krav besluttet.

Trinn 3 (uke xx)
Overordnede og delmålsettinger legges inn Miljøoppfølgings- og kvalitetsplan – MOP. Krav til dokumentasjon av kvaliteter legges inn i MOP. For dokumentasjon av kvaliteter brukes det BREEAM-NQR kriterier der dette er mulig. MOP fremsendes til byggherre for kommentering.

Trinn 4 (uke xx)
Møte med gjennomgang av siste versjon av MOP.



PROSJEKTVISJONER

- **Rona torv skal sette mennesket i sentrum.** Hva skal Rona torv assosieres med? Rona torv skal sette mennesket i sentrum for det som jobbes med. Det skal være et godt sted å bo, oppholde seg, være. Det vil også være kommersielt smart. Det skal oppleves godt å være menneske. Bydelssenter har en samfunnsmessig betydning. Vi må tenke helhet for området - hva kan prosjektet tilby Kr.sand og området rundt? Vi må gjøre området kommersielt mer attraktivt for leietakere. Tomta egner seg godt for kontor/ næring, hvordan henger disse sammen med planlagt utbygging rundt eks. Strømsøya - hvordan kan vi utkonkurrere kontorene her? Funksjoner og mennesker Rona torv 50 år frem i tid: Flere eldre som skal bruke området frem i tid. Møteplassene må kunne brukes av både barn ungdom og eldre. Mye drabantby som skal bygges rundt torget . Rona torg må derfor bli et sted med identitet og sjel, se til eks Grunerløkka. For å få til dette må det komme nok folk til torget. Tetthet er suksessfaktor. Bra kommersielt, men også for prosjektet og stedets fremtid. Dekke behovene i dag og om 50 år. Demografi - i dag og frem i tid - en bydel for alle I dag ung befolkning. Rekkehus og spredt bebyggelse. Fremover i tid: flere eldre. Hvem skal bruke torget og hvem bor i nærheten? Alternative boformer/ store nok leiligheter/ skiftes behovet for hvordan vi bor?
- **Rona torv skal tilby sambruk av arealer og ulike funksjoner til alle:** Sirkulære strategier og fleksibilitet: Hvordan er Rona om 20-30 år? Hvordan skal vi bygge for fremtiden? Hvordan legger vi til rette for fleksible planer/ funksjoner så bygget kan bli noe annet om 50 år? Kommunen har løftet frem sambruk av arealer og ulike funksjoner til ulike typer grupper, dette vil være viktig å jobbe med i prosjektet. Dette har et økonomisk perspektiv og er en viktig visjon for prosjektet.
- **Rona torv skal bli et bydelssenter med trygge og gode uterom:** Regional plan for bolig og mobilitet: Rona torv skal bli et bydelssenter. Planen definerer hva et bydelssenter skal være. Området er et stort knutepunkt. Hvordan organiseres trafikken? Hvordan unngå "kaos"? Å strukturere trafikken vil bli viktig å arbeide med. Kommunikasjonsknutepunktet skal dekke et stort område. Rona torg er starten på et større boligområde Drangsvann som skal bygges ut. Kommunen er klar for å satse på tomta og er tydelige på hvilke funksjoner - en rekke servicefunksjoner for barn og voksne. Prosjektet er med på å gjøre området mer tilgjengelig, men hvordan legger vi til rette for grønn mobilitet i fremtiden? Skal vi definere en overordnet visjon for hele området og/ eller visjoner for deler av prosjektene? Hvordan reiser folk i 2050? Færre privatbiler, legge til rette for sykkel og gange, flere tettsteder. Hvordan kan p-plassene brukes dersom det ikke lenger er behov for samme antall p-plasser?

Vision generated manually, ca. 3 hours

Jan Buthke fredag 11:51

Teksten diskuterer den fremtidige visjonen for "Rona Torg," et utviklingsprosjekt.

Menneske-sentrisk tilnærming: Rona Torg skal være et sted som setter mennesket i sentrum, et godt og attraktivt sted å bo og oppholde seg, samtidig som det er kommersielt smart.

Fleksibilitet og sirkulære strategier: Det legges vekt på fremtidsrettet bygging, med fleksible planer som kan tilpasse byggets funksjoner over tid, også til noe annet om 50 år.

Regional bolig- og mobilitetsplan: Rona Torg skal bli et bydelssenter, og det diskuteres hvordan trafikken kan organiseres for å unngå kaos. Grønn mobilitet er også et fokusområde.

Større boligområde og kommunale planer: Rona Torg er starten på et større boligområde, Drangsvann, som skal bygges ut. Kommunen planlegger å tilby forskjellige tjenester for barn og voksne.

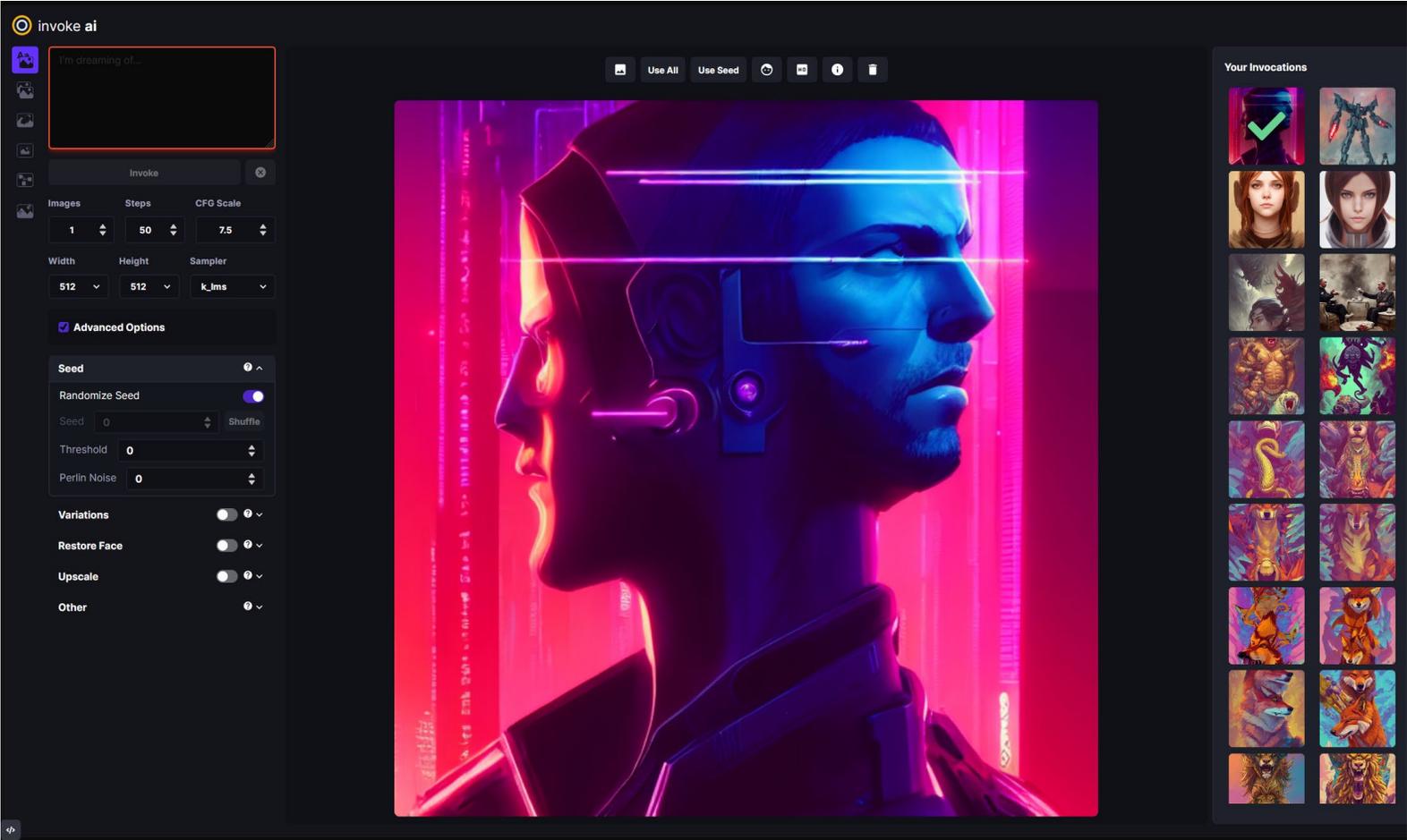
Demografi og boformer: Det diskuteres hvordan området skal tilpasse seg forventede demografiske endringer, med flere eldre som vil bruke området. Diskusjonen inkluderer også spørsmål om alternative boformer og hvordan behovene for bolig kan endre seg.

Samfunnsmessig betydning: Rona Torg skal ikke bare være et bydelssenter, men også ha en bredere betydning for samfunnet rundt. Det er også et mål å gjøre området mer attraktivt for leietakere, med en god blanding av kontor/næring i forhold til den planlagte utbyggingen rundt.

Vision generated by ChatGPT

Open Source and an outlook to the future

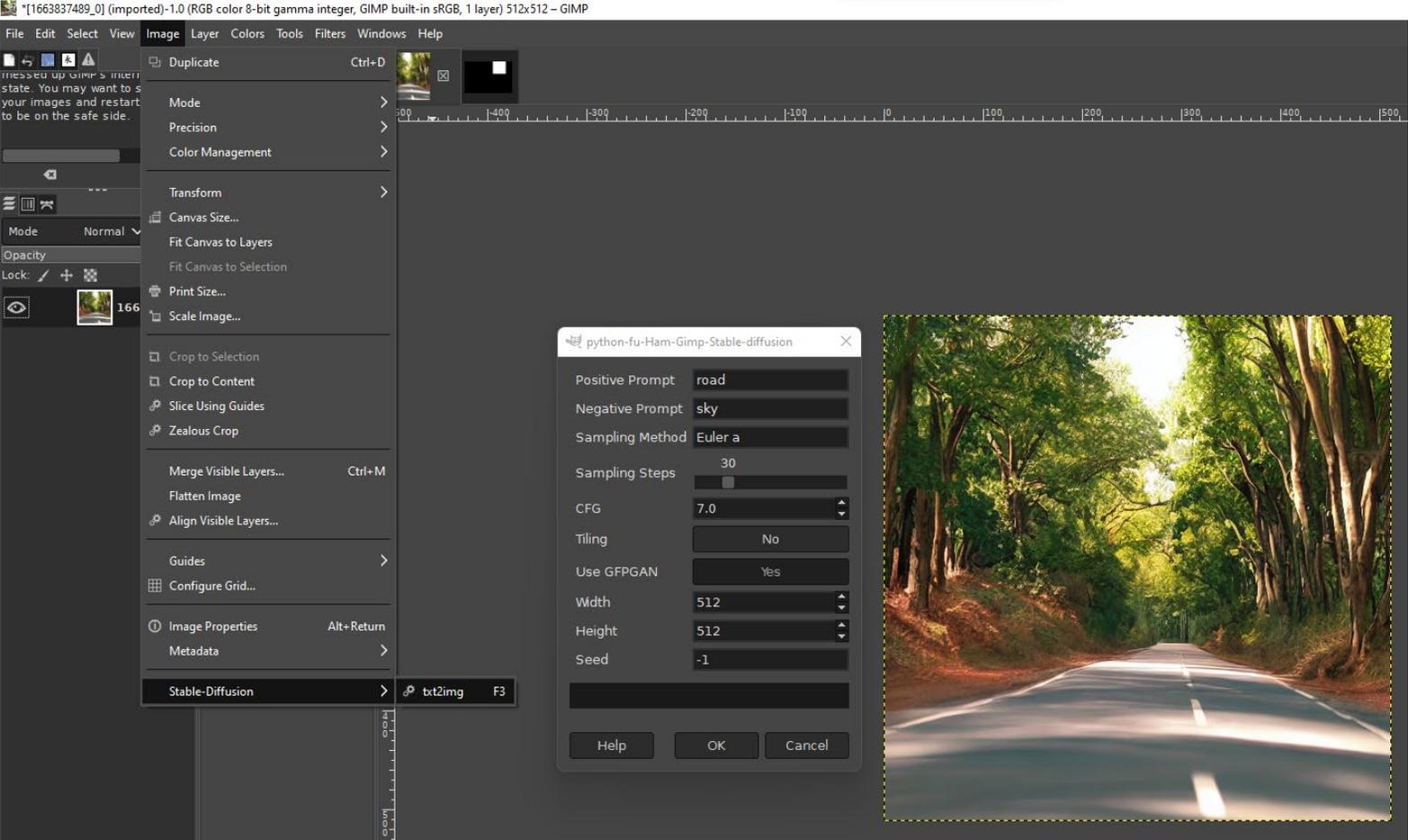
Invoke AI



Invoke AI an open-source front end for Stable Diffusion
<https://github.com/invoke-ai>



Gimp – Stable Diffusion Integration



Open Source LLM's

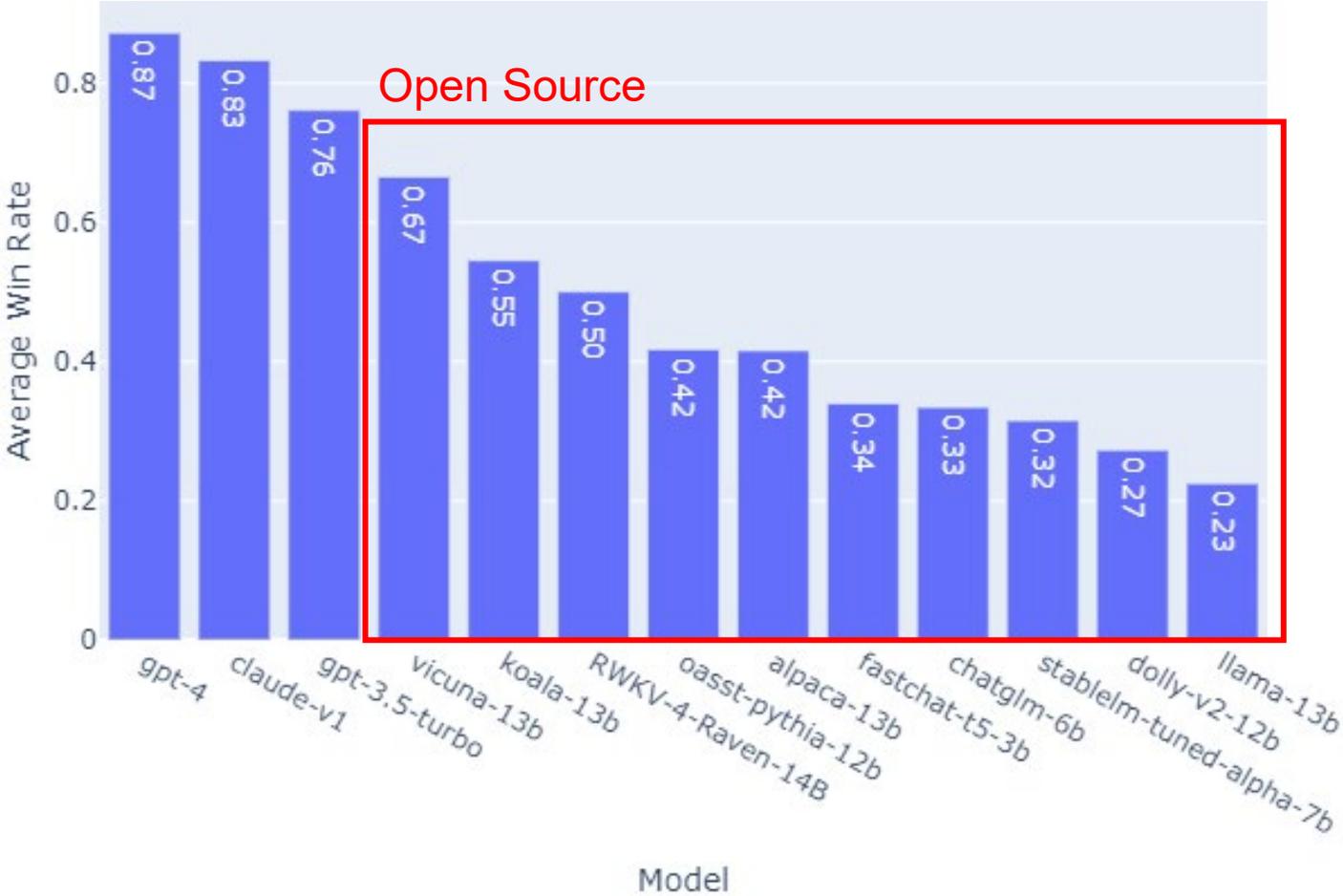
https://github.com/eugeneyan/open-llms

Language Model	Release Date	Checkpoints	Paper/Blog	Params (B)	Context Length	Licence
T5	2019/10	T5 & Flan-T5, Flan-T5-xxl (HF)	Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer	0.06 - 11	512	Apache 2.0
UL2	2022/10	UL2 & Flan-UL2, Flan-UL2 (HF)	UL2 20B: An Open Source Unified Language Learner	20	512, 2048	Apache 2.0
Cerebras-GPT	2023/03	Cerebras-GPT	Cerebras-GPT: A Family of Open, Compute-efficient, Large Language Models (Paper)	0.111 - 13	2048	Apache 2.0
Open Assistant (Pythia family)	2023/03	OA-Pythia-12B-SFT-8, OA-Pythia-12B-SFT-4, OA-Pythia-12B-SFT-1	Democratizing Large Language Model Alignment	12	2048	Apache 2.0
Pythia	2023/04	pythia 70M - 12B	Pythia: A Suite for Analyzing Large Language Models Across Training and Scaling	0.07 - 12	2048	Apache 2.0
Dolly	2023/04	dolly-v2-12b	Free Dolly: Introducing the World's First Truly Open Instruction-Tuned LLM	3, 7, 12	2048	MIT
DLite	2023/05	dlite-v2-1_5b	Announcing DLite V2: Lightweight, Open LLMs That Can Run Anywhere	0.124 - 1.5	1024	Apache 2.0
RWKV	2021/08	RWKV, ChatRWKV	The RWKV Language Model (and my LM tricks)	0.1 - 14	infinity (RNN)	Apache 2.0
GPT-J-6B	2023/06	GPT-J-6B, GPT4All-J	GPT-J-6B: 6B JAX-Based Transformer	6	2048	Apache 2.0
GPT-NeoX-20B	2022/04	GPT-NEOX-20B	GPT-NeoX-20B: An Open-Source Autoregressive Language Model	20	2048	Apache 2.0
Bloom	2022/11	Bloom	BLOOM: A 176B-Parameter Open-Access Multilingual Language Model	176	2048	OpenRAIL-M v1
StableLM-Alpha	2023/04	StableLM-Alpha	Stability AI Launches the First of its StableLM Suite of Language Models	3 - 65	4096	CC BY-SA-4.0
FastChat-T5	2023/04	fastchat-t5-3b-v1.0	We are excited to release FastChat-T5: our compact and commercial-friendly chatbot!	3	512	Apache 2.0
h2oGPT	2023/05	h2oGPT	Building the World's Best Open-Source Large Language Model: H2O.ai's Journey	12 - 20	256 - 2048	Apache 2.0
MPT-7B	2023/05	MPT-7B, MPT-7B-Instruct	Introducing MPT-7B: A New Standard for Open-Source, Commercially Usable LLMs	7	84k (ALiBi)	Apache 2.0, CC BY-SA-3.0
RedPajama-INCITE	2023/05	RedPajama-INCITE	Releasing 3B and 7B RedPajama-INCITE family of models including base, instruction-tuned & chat models	3 - 7	2048	Apache 2.0
OpenLLaMA	2023/05	OpenLLaMA-7b-preview-300bt	OpenLLaMA: An Open Reproduction of LLaMA	7	2048	Apache 2.0

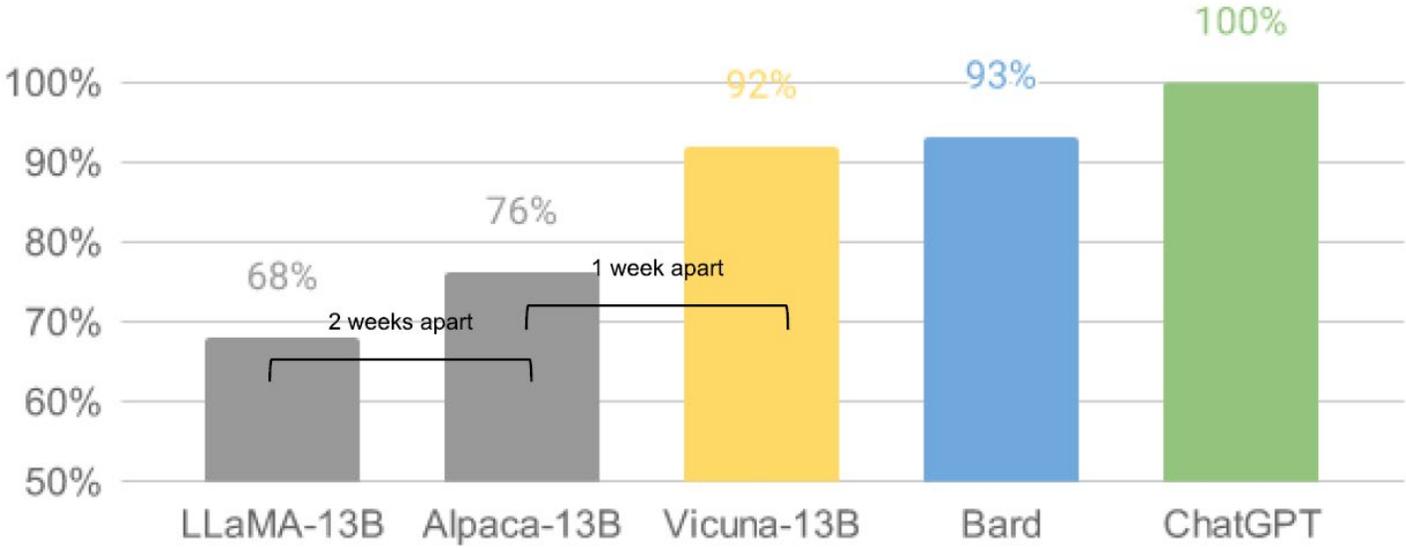


Benchmarking of LLM's

Leaderboard on lmsys.org



Google "We Have No Moat, And Neither Does Open AI"



*GPT-4 grades LLM outputs. Source: <https://vicuna.lmsys.org/>

Leaked Internal Google Document Claims Open Source AI Will Outcompete Google and Open AI



Open ended questions:

- How can we start using the data we have as architects.

Maybe we need a library/tool that let us programmatically query our models and extract structured data from them. While forgoing heavy solutions such as Revit.

- Is data really ours, if we require a license to read it.

Driftig by Intention Online, 24.05.2023

Povl Filip Sonne-Frederiksen
PhD Fellow
+45 61 93 61 44
pfs@linkarkitektur.dk

LINK Arkitektur
linkarkitektur.com

