“Ask Me Almost Anything”:

Demystifying GenAI with Science, Art, and Answers

Chris Collison (COS, AI HUB), Shaun Foster (CAD), Andrew White (Keynote Speaker)
Friday, May 16 | 10:00-11:15AM | SHED, Room 3300

Welcome! Please take a seat.

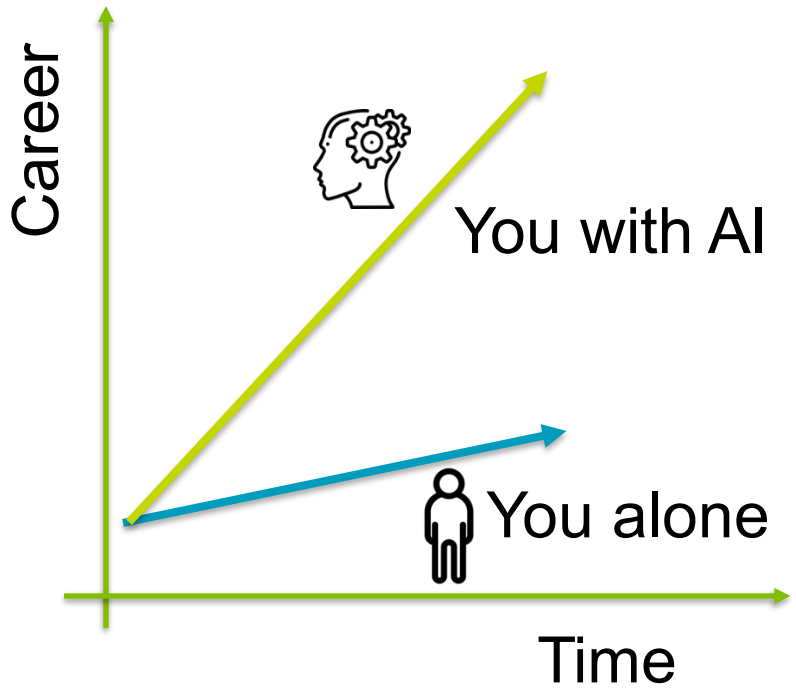
Outline

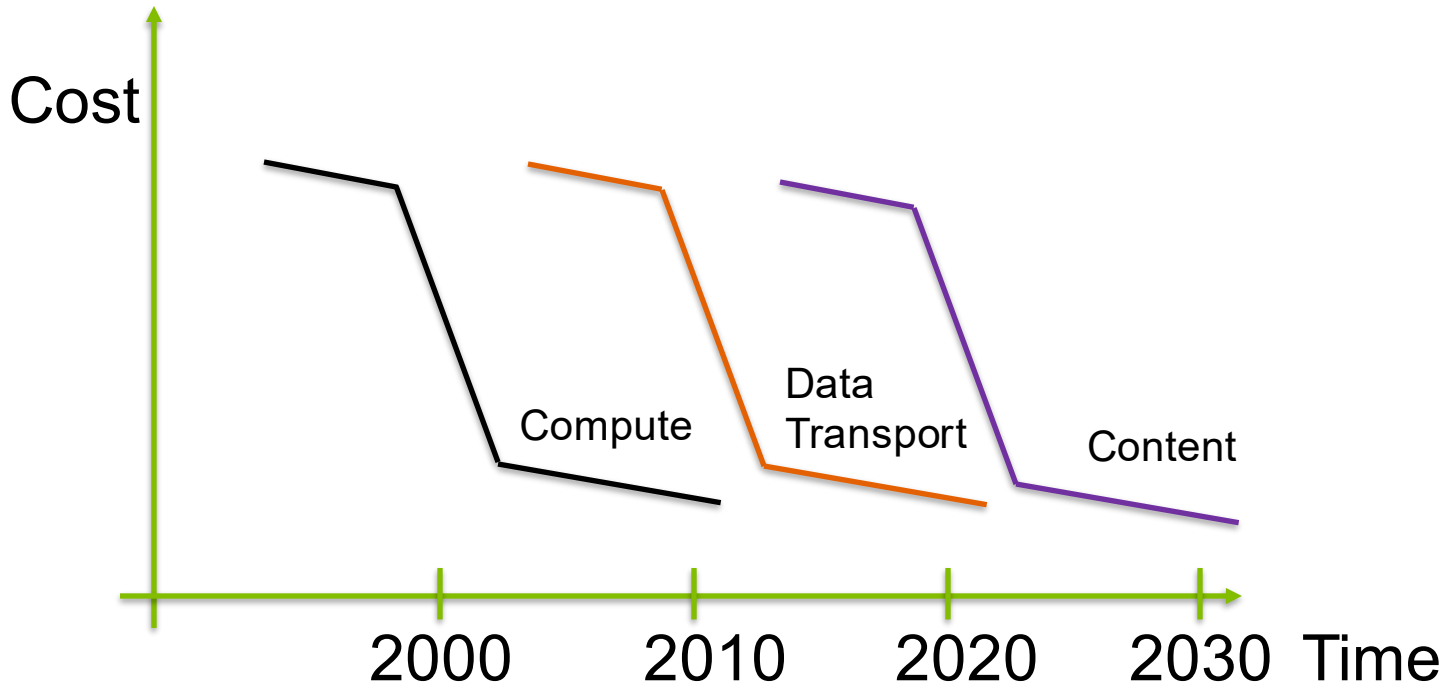
- We've scoured the world (and the internet) for the big questions people ask about AI
  - Today's session is about what's under the hood
  - We'll explore how AI works, and why we might want to use it
- We'll start with some ice-breakers and then dive in
- We want to hear your questions!

Is AI taking our Jobs or not?

Is AI taking our Jobs or not?

The genie is out… and the bottle is no more!





Is AI Dumb?

It's Just a Tool

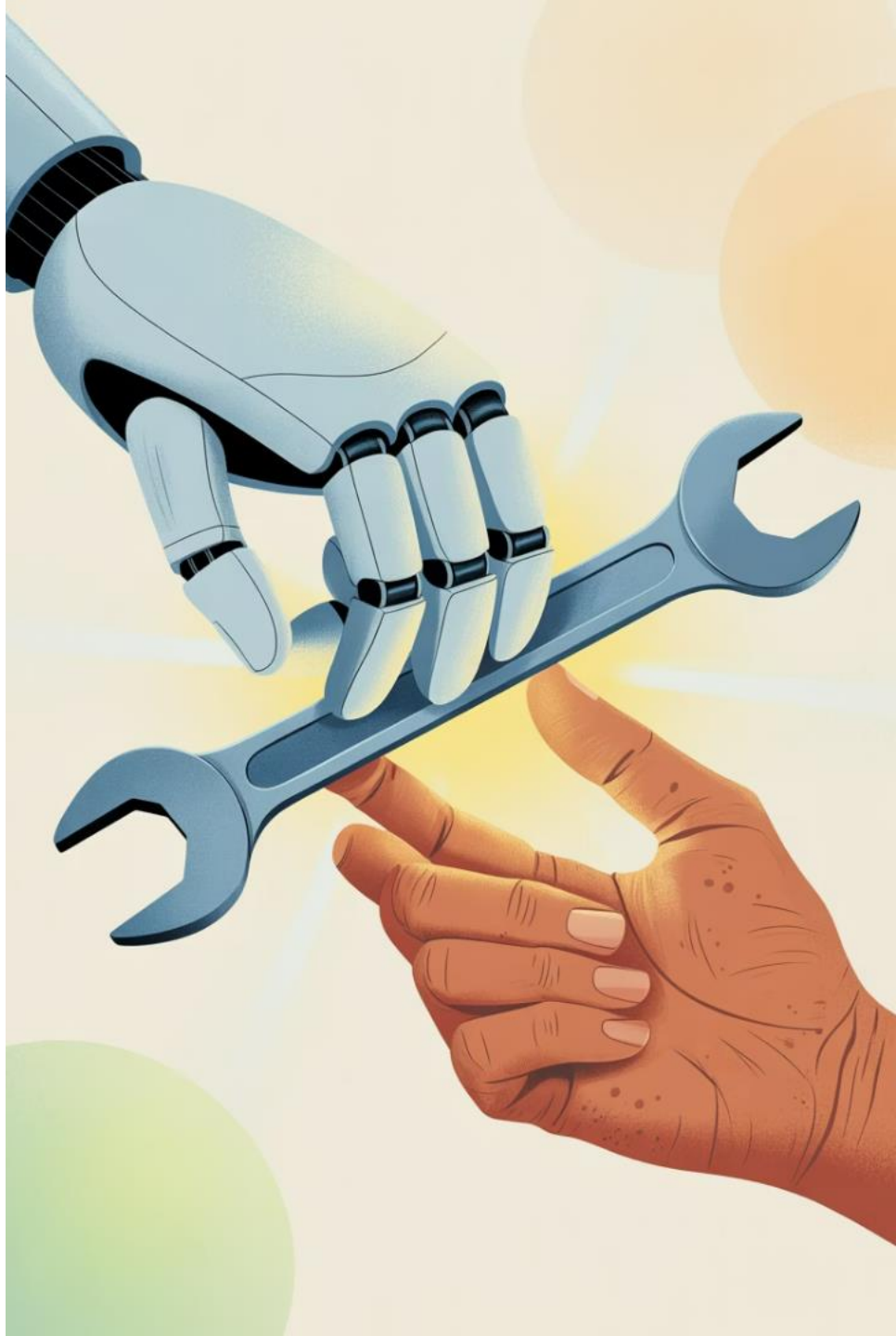
Like a hammer, its effectiveness depends on the user

Pattern Recognition

Excels at finding patterns in in massive datasets

Limited Understanding

No true comprehension of meaning or context



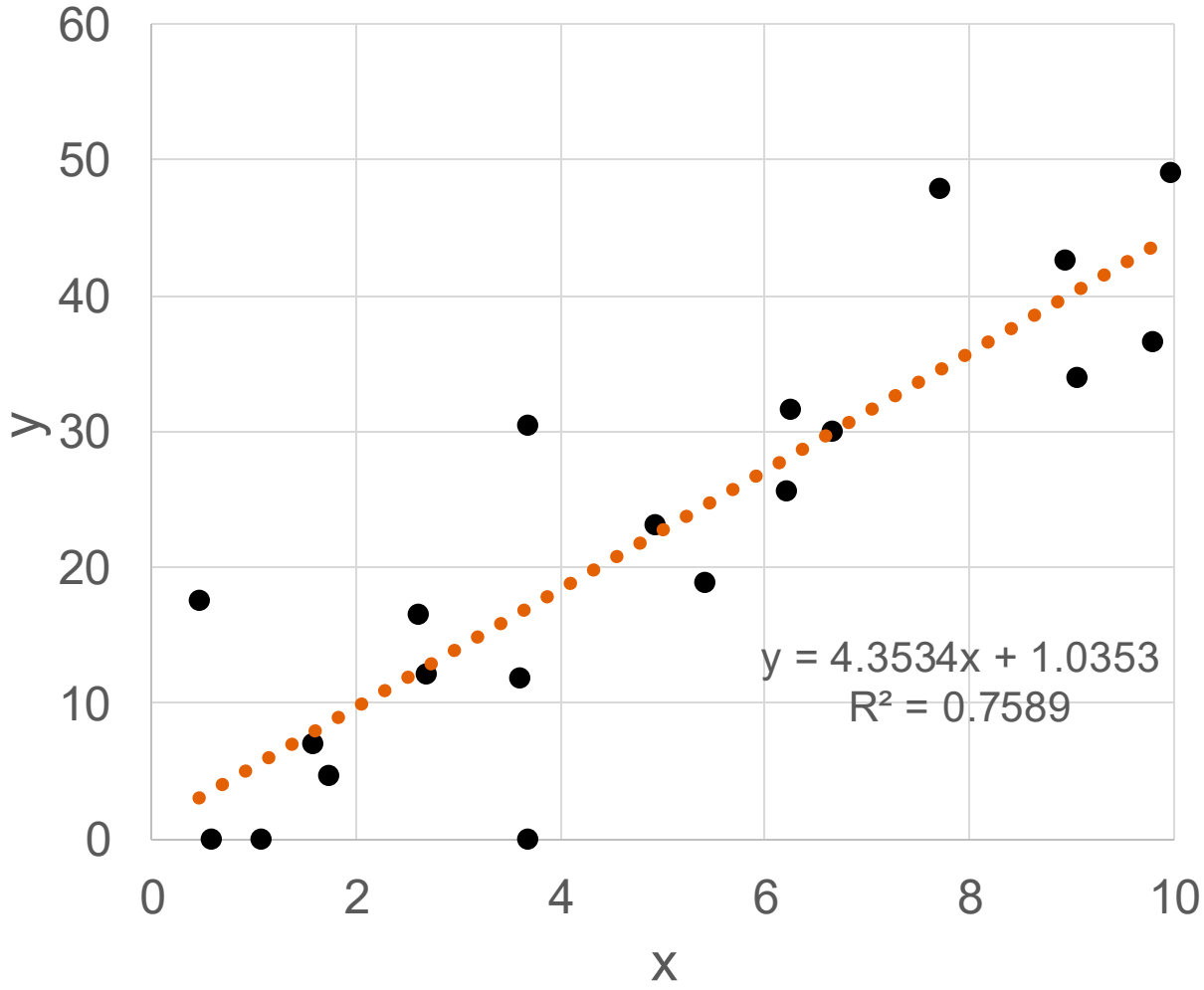
What is AI?

- Machine Learning
- Neural Networks (GNN, RNN, CNN)
- Transformers
  - Seq2Seq Translation
  - Sentiment Analysis
  - Generative AI
- Diffusion Models

Let’s have a look…

Machine learning – modeling data with prediction

- Exercise time (x) vs. calories burned (y)
- Advertising budget (x) vs. products sold (y)
- Years experience (x) vs. salary bump (y)
- Shoe size (x) vs. height (y)



Cost function: SSE

Error is $\hat{y}_1 - (w_1 x_1 + c)$

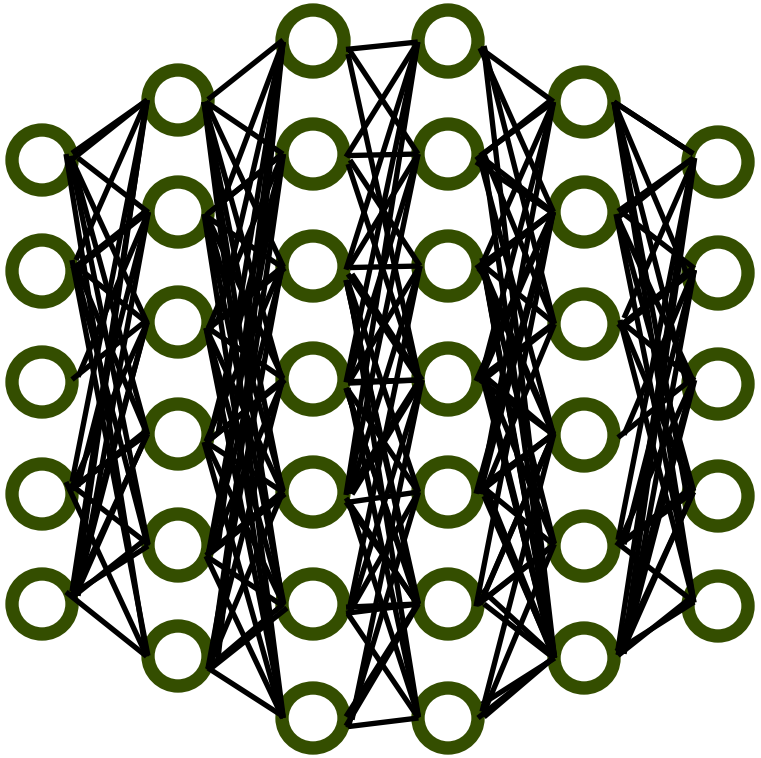
ML is often much better at a specificized single task compared to “generalized AI”

Using Features and Labels

Feature	Action	Comedy	Drama	Horror	Sci-Fi	Romance	Documentary	Fantasy
Pacing	High	Moderate	Slow-Moderate	High	Moderate	Slow	Slow	Moderate
Special Effects	Essential	Low	Low	Moderate	Essential	Low	Low	Essential
Emotional Depth	Low	Low	High	Low	Moderate	High	High	Moderate
Humor/ Levity	Low	Essential	Moderate	Low	Low	Moderate	Low	Low
Character Development	Moderate	Moderate	High	Low	Moderate	High	High	Moderate
Suspense/ Thrills	High	Low	Moderate	Essential	High	Low	Moderate	High
Romantic Subplot	Low	Moderate	Moderate	Low	Low	Essential	Low	Moderate
Realism	Low	Low	High	Low	Low	Moderate	Essential	Low

These are Action Movies

Deep Learning - Neural Networks, Computer Vision



Elephant

Motorcycle

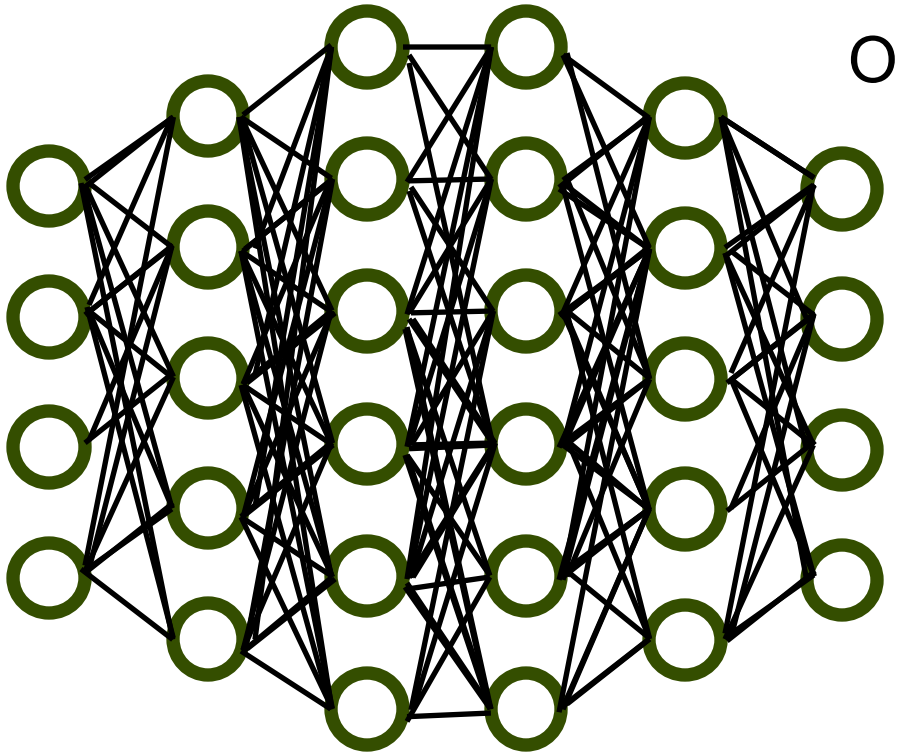
Dog

Traffic Lights

Laptop computer

Some training required!

Deep learning and neural networks used interchangeably



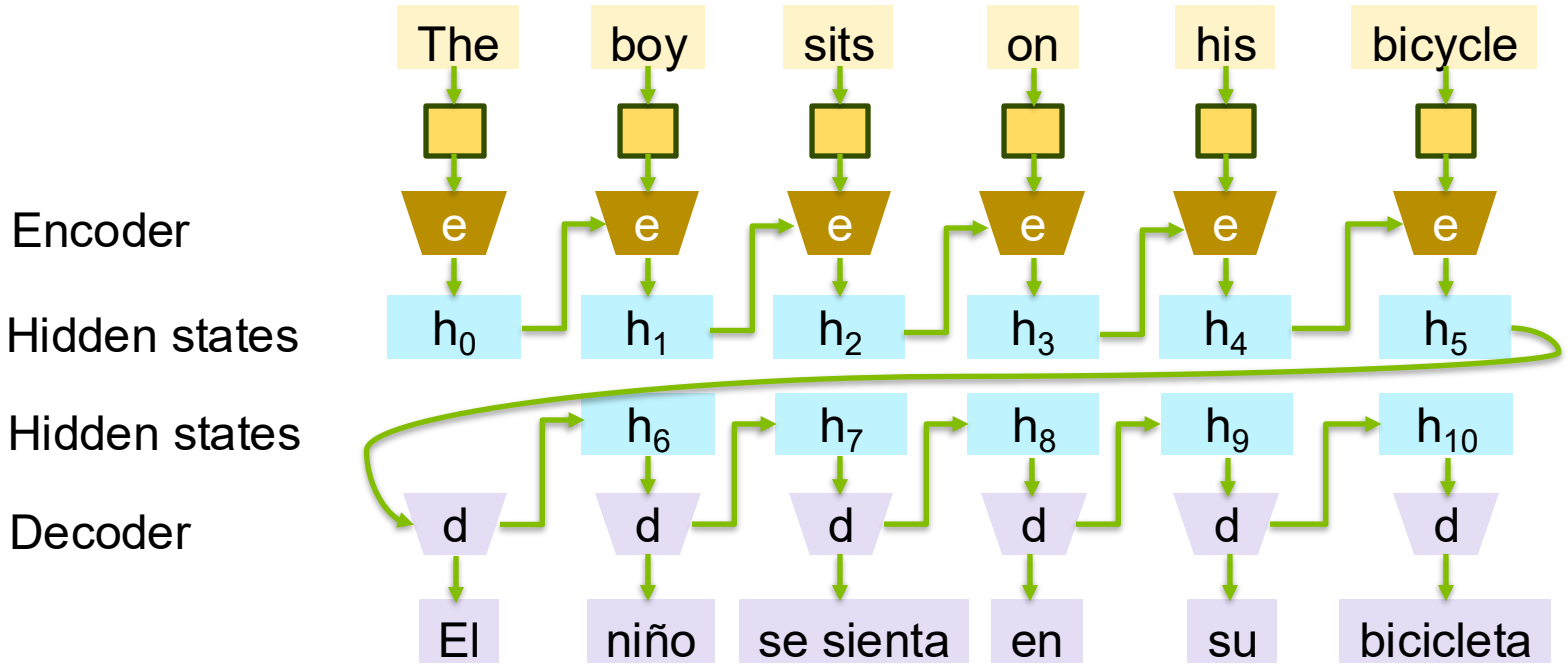
- Composed of nodes
- Each node a linear regression of all inputs
- Data passed forward: Feed forward

- Can adjust weights with training
- Use labeled datasets
  - What was in the fridge?
  - Did I go shopping?
- Minimize cost function through gradient descent
  - Is cost going down?
- CNN, RNN

History – From NNs to Transformers…

Should I go shopping? $\hat{y} = 1$, $\hat{y} = 0$?		
Need Milk?	$x_1 = 0$	$w_1 = 5$
Still have leftovers?	$x_2 = 1$	$w_2 = -2$
Need Chips?	$x_3 = 1$	$w_3 = 10$
$x_1 w_1 + x_2 w_2 + x_3 w_3 - 5 = +3 \Rightarrow \hat{y} = 1$		

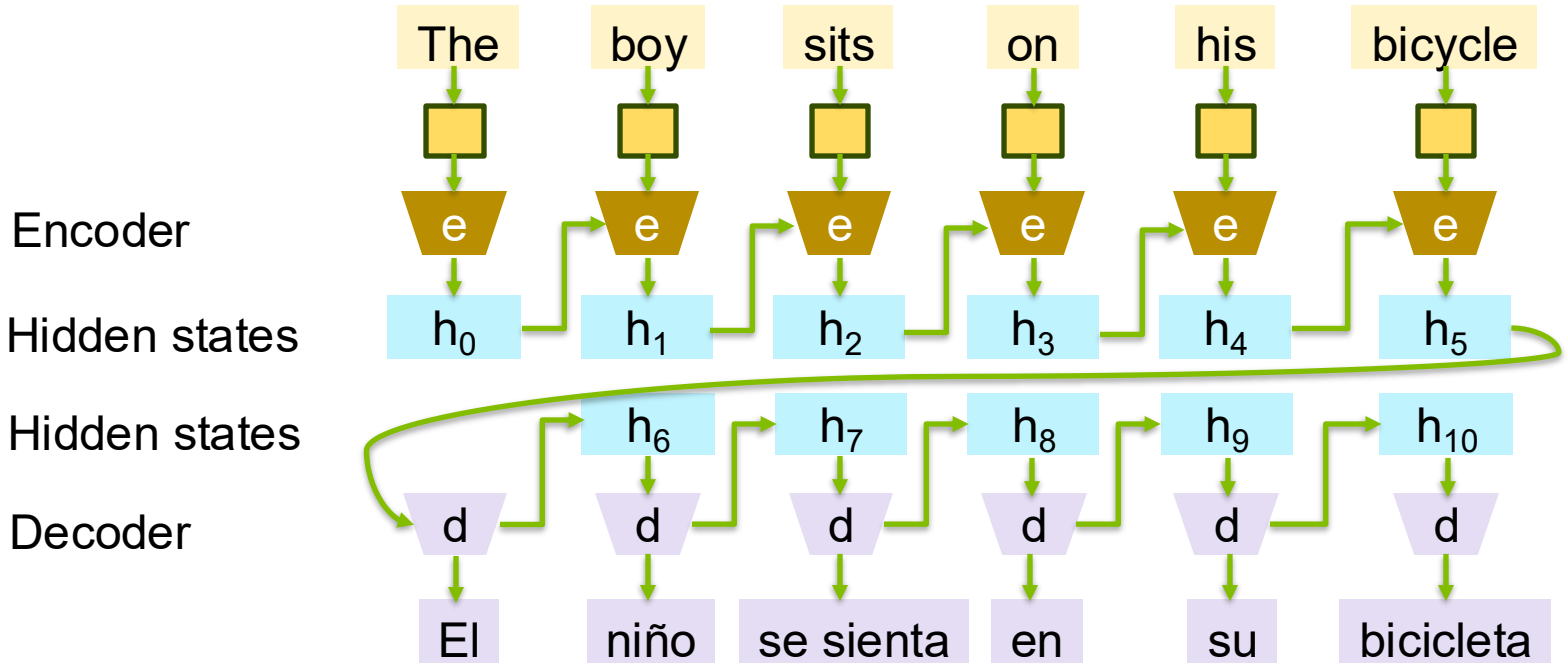
seq2seq translation with an RNN



- Big gap from “sits” to “se sienta”
- Must include all words and their order
- Problems with longer sentences

- Words processed one at a time, left to right.
- A “hidden state” carries memory forward – last hidden state captures the whole sentence.
- Final state of the encoder is passed to the decoder to generate the translation
- Information about early words like “The boy” fades by the time we process “bicycle.”
- Long sequences = vanishing memory.
- Early words get forgotten, making it hard to match correct word order and meaning.

seq2seq translation with ATTENTION mechanism



The	el
Boy	niño
Sits	se
on	sienta
his	en
bicycle	su
	bicicleta

- Instead of relying only on last hidden state, decoder looks back at all encoder states
- decides most relevant words to current output word.
- While generating “bicicleta,” the decoder pays more attention to “bicycle” in the input than to “boy.”
- This is done in training by

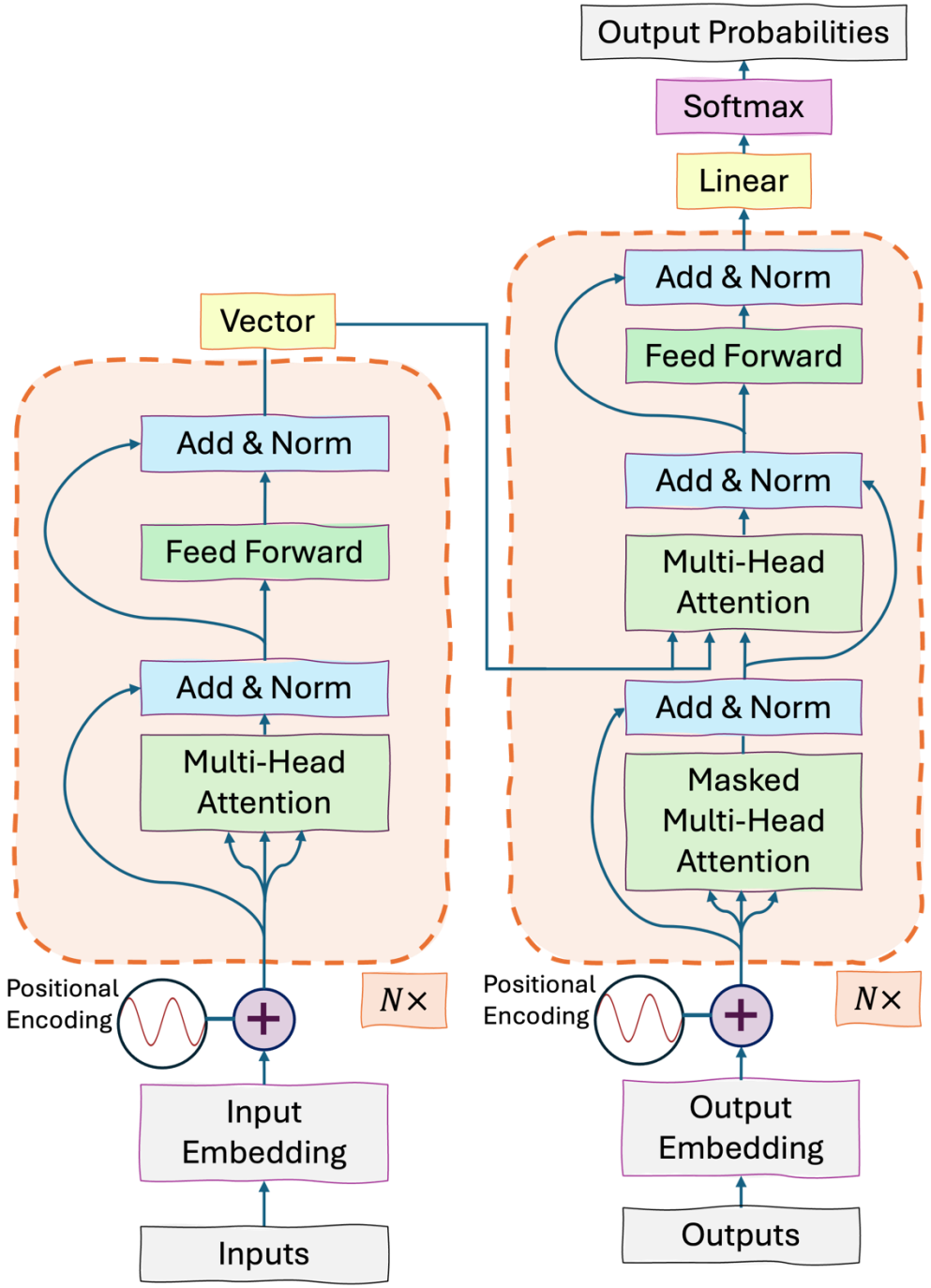
What is Attention?

	[The]	[boy]	[sits]	[on]	[his]	[bicycle]
bicicleta	0.01	0.01	0.01	0.02	0.15	0.8

- Model trained on many sentence pairs.
- It learns **which input words help generate each target word** by adjusting weights to **minimize prediction error (gradient descent)**.
- Query, Key, Value
- Over time, model learns **which input words matter**, assigns attention weights accordingly.

What is a Transformer?

- Vaswani et al – 2017
  - “Attention is all you need”
- Tokenization
- Input Embedding
- Positional Encoding
- Encoder
- Decoder



What Are Tokens and What Is Tokenization?

- **https://platform.openai.com/tokenizer?ref=blog.mlq.ai**

What is Embedding?

- **Vector representation for 'teaching': [0.4291865 0.07661827 0.0271875 ... 0.546991 -0.24655312 0.00974098] Vector length: 1024**

- 0.4292,0.0766,0.0272,-0.3943,-0.4762,-0.2504,-0.5945,0.0894,-0.2697,0.8750,0.1055,0.0245,1.0137,-0.3504,0.0873,-0.5476,-0.2255,-0.1106,-0.9785,-0.2603,-0.2918,0.2812,-0.4427,-0.5341,-0.3278,-0.0908,0.1897,0.0680,0.8227,0.3343,0.2872,-0.1292,0.3094,-0.8513,0.1825,-0.1127,0.3441,-0.5287,-0.2605,-0.8983,0.4789,0.0285,0.5200,-0.3802,-0.3684,-0.0728,0.1413,-0.4228,-0.4385,-0.1717,0.0855,0.5044,0.1875,0.1172,0.1053,-0.3366,-0.1360,-0.5556,-0.2088,0.0973,0.4096,0.3106,-0.4630,-0.9954,-0.1491,-0.1050,0.0957,0.1320,-0.3110,-0.0360,-1.0190,0.1111,0.5704,0.0377,-0.6130,0.2350,0.2403,-0.1572,0.1521,0.4703,0.3533,0.0165,0.0393,0.1365,-0.6896,0.0336,0.0961,0.0971,-0.0561,0.4200,-0.3340,0.2932,-0.7691,0.4166,0.1887,0.2138,0.3166,0.3830,0.1444,0.0105,1.0344,0.9393,-0.6910,0.5805,-0.7918,0.7489,-0.5214,0.2444,-0.2991,-0.6463,-0.3333,0.3215,0.7430,-0.4953,0.0138,0.8390,0.2024,0.3572,-0.4479,0.6118,0.0975,0.0608,0.2051,-0.0626,0.6752,-0.1890,-0.4319,0.7834,-0.4152,-0.0855,-0.1948,0.0207,-0.7732,0.2515,0.0986,0.3435,0.3966,0.9752,0.2158,-0.3823,0.0967,-0.0183,-0.2610,1.1274,-0.1384,0.0264,-0.8959,-0.2021,-0.6223,0.1162,-0.3779,-0.1763,0.1838,0.4307,0.2102,-0.2597,0.0019,0.1279,-0.4817,-0.1656,-0.4585,-0.2004,-0.3601,0.4848,0.2477,-0.0574,0.0540,-0.4904,0.1560,-0.4400,0.5694,0.2615,0.0927,-0.3616,0.1350,-0.1490,0.4788,-0.0010,0.3418,0.0787,-0.4735,-0.3900,0.4177,0.0742,0.3271,-0.2117,-0.3264,-0.1889,-0.4875,0.1291,-0.2400,0.4808,-0.2086,-0.1940,-0.0109,0.0678,-0.7672,-0.4160,-0.3990,-0.4389,-0.2435,0.4385,0.0927,0.4407,0.3658,0.1248,0.0670,0.2545,-0.2156,0.1801,0.0025,0.2125,-0.1494,0.0568,0.5289,-0.3514,-0.7372,0.6402,-0.0990,-0.1705,0.4325,0.2167,0.5575,0.3088,-0.1712,-0.3486,-0.1166,0.7293,-0.0544,0.6068,0.5159,0.3545,0.4610,0.3667,0.4837,0.3107,0.1108,0.5384,0.7538,-0.0587,0.1001,-0.3454,0.7942,0.2384,-0.4419,0.4395,-0.4789,-0.1361,-0.4804,0.0463,-0.8562,0.6409,0.2816,0.2143,-0.5732,0.4942,0.7459,0.2928,-0.5881,-0.2187,0.1702,0.2622,-0.0747,0.1584,0.4708,0.5135,-0.0855,0.7773,0.0145,-0.8638,-0.3436,-0.9359,-0.7156,0.4501,-0.3961,-0.2535,0.0722,-0.5368,0.1537,-0.1492,-0.2371,0.3003,-1.1163,0.7417,0.7847,0.0407,-0.5291,0.2455,-0.6030,-0.1833,-0.7447,-0.2769,-0.0350,0.1430,0.7931,-0.3649,0.0922,0.4474,-0.3952,-0.1677,0.1304,-0.0751,0.6128,0.3565,-0.3244,0.8096,-0.4657,-0.6522,0.5987,-0.2253,-0.5255,0.1407,0.0060,0.1043,-0.5368,0.7449,0.3635,-0.1232,-0.1481,-0.9214,-0.0615,-0.4601,-0.0052,-0.4572,0.0050,0.4045,0.0234,-1.0003,0.0945,-0.4992,-0.4819,-0.6252,-0.5062,0.1464,0.2946,0.2785,-0.3002,0.0415,0.4698,-0.2841,0.2274,-0.5619,0.3666,0.5246,0.3225,-0.4383,-0.4883,-0.0172,-0.3692,0.3820,0.5720,-0.2838,0.3724,0.2094,-0.1161,0.8934,-0.3405,0.1847,0.6158,0.2785,-0.7307,0.0628,0.1607,-0.7167,-0.5980,-0.5139,0.0863,0.8097,0.6469,-1.1845,0.8486,0.0672,-0.1113,0.5417,-0.2874,-0.1973,0.6770,0.1690,0.5104,-0.3211,-0.4075,-0.0523,0.0178,0.2008,-0.2537,0.2503,0.1591,-0.0247,0.0475,-0.4009,0.1243,-0.0298,0.0717,-0.2398,0.1234,-0.4622,1.2282,0.4642,0.3777,-0.1468,0.4029,-0.6116,0.9402,-0.0425,-0.1830,-0.3368,-0.6450,-0.0175,0.4482,0.0598,-0.2151,0.0543,-0.4587,-0.0089,0.1307,0.1492,0.2206,-0.2600,-0.5179,0.1610,-0.2497,-0.4481,-0.0287,0.7499,0.7913,-0.4221,0.0730,-0.6233,0.7164,0.3492,-0.0443,0.1964,0.2987,-0.3993,-0.1339,-0.6658,0.3190,0.3222,0.0232,-0.2948,0.0526,0.1189,-0.1456,-0.1307,0.0678,0.0198,0.3586,0.2287,-0.1366,-0.5903,0.2528,-0.3885,0.3151,0.0941,0.3846,-0.1060,-0.1568,-0.2302,0.8075,0.2178,-0.2327,-0.4470,0.7381,0.0656,-0.3344,0.4687,-0.1144,0.1815,0.4544,0.1260,-0.1441,-0.7693,0.0656,-0.1171,0.4662,-0.5571,-0.8631,0.3932,-0.5744,0.2735,-0.0945,0.4775,-0.2041,-0.5487,0.0338,-0.0367,-0.0899,-0.2901,-1.3163,-0.0312,-0.4132,-0.2330,0.0697,-0.0355,-0.0598,0.0185,0.2722,-0.3267,0.2657,0.8258,0.7704,-0.8205,-0.4631,-0.1560,-0.4641,-0.1403,-0.5476,0.0556,-0.1343,-0.3026,0.2566,0.3372,0.0167,0.9160,0.2937,-0.0980,-0.1209,-0.0993,-0.2552,0.5893,-0.2673,0.0365,-0.0815,0.3750,-0.1758,0.5930,-0.3410,-0.0614,-0.6563,-0.4247,0.2919,0.2165,-0.1799,-0.4004,0.4076,0.2729,0.1868,0.1146,0.6752,0.2881,-0.8301,0.0027,-0.5553,0.0669,-0.3428,-0.2171,-0.4359,0.1114,-0.0110,0.2306,0.1756,-0.0369,0.1531,-0.3315,-0.4395,-0.2045,0.0759,0.2888,-0.2665,-0.0498,-0.5619,0.5170,-0.2103,0.0284,-0.3934,0.0253,0.0733,-0.0285,0.9315,-0.0932,0.1749,-0.7340,0.2860,0.3987,-0.5826,-0.2287,-0.3180,-0.4298,-0.6449,0.2707,0.0492,-0.3185,-0.0738,-0.1577,0.2089,-0.2659,0.9869,0.8269,-0.3751,0.1401,-0.3360,0.4001,-0.0022,0.2747,-0.0064,-0.4791,0.1055,0.3253,-0.1644,-0.5340,-0.2673,0.0527,0.7355,-0.0348,0.4506,-0.7017,-0.5312,-0.2306,0.1954,-0.5761,0.3390,0.2728,0.0387,-0.2024,0.2731,-0.4518,-0.4287,0.2942,0.5614,0.1001,-0.8864,-0.3954,0.3510,-0.6688,-0.4821,-0.1037,0.2477,-0.2372,-0.2498,-0.8041,0.6447,-0.2950,0.0070,0.2912,-0.5038,0.0003,-0.1356,0.0408,-0.7018,0.2250,0.5647,-0.0557,-0.1599,0.5415,-0.4553,-0.0434,-0.3491,0.5437,-0.6259,-0.5369,-0.1835,-0.7491,0.4656,0.1664,-0.1396,0.2499,-0.9524,-0.2538,-0.1965,0.2535,0.3607,0.2651,-0.4525,-0.1609,0.3539,0.2169,0.1530,-0.5612,-0.4378,-0.6578,-0.7576,-0.4951,-0.2667,-0.1505,-0.1366,0.3029,-0.5267,0.1275,0.7051,-0.3868,0.6480,0.0524,0.1997,0.0090,-0.2963,-0.2562,-0.3068,0.0124,0.1261,-0.5496,0.3646,0.0774,0.5860,0.0828,-0.3322,-0.8994,0.2417,0.3425,0.2162,-0.1978,0.0383,0.6738,-0.6414,0.1079,0.0623,-0.2067,-0.3274,0.2584,-0.2825,-0.4670,-0.5716,0.0351,0.0859,-0.2087,-0.0962,-0.0066,0.5307,-0.7250,-0.4537,0.6085,1.0018,0.3681,0.3637,0.0055,0.3176,-0.1013,0.5544,-0.1109,0.3391,0.2715,-0.2475,0.3116,0.5542,-0.5853,0.7436,-0.2264,0.3206,-0.0580,0.3968,-0.1289,0.4879,0.0921,-1.4642,-0.2599,-0.1010,-0.1886,0.0742,-0.6508,-0.6838,-0.0389,-0.2731,-0.6388,-0.2655,-0.5296,-0.3431,0.0864,0.3136,-0.3721,0.2527,0.0610,-0.1219,0.1789,-0.3375,-0.2519,0.3468,-0.0177,-0.9668,0.2222,0.5906,0.3363,0.3149,-0.2713,-0.0794,0.0412,0.9570,-0.6616,0.1402,-0.2738,-0.3370,-0.2032,0.0321,-0.0565,0.4722,0.1129,-0.3786,-0.0527,0.3528,0.3404,-0.0985,0.2184,0.3321,0.9537,0.3347,-0.6584,-0.0069,0.3513,-0.1373,0.9251,0.4401,-0.3330,0.5479,0.2324,0.3390,0.7817,-0.0549,-0.0449,-0.1307,-0.5782,-0.4117,0.6513,-0.4167,0.7479,-0.7629,0.0738,0.0682,-0.1197,0.6154,-0.3091,-0.4670,0.2056,-0.0371,0.5634,0.2814,0.1388,0.3216,0.4282,0.0926,0.3113,-0.2380,-0.0594,0.1315,-0.8165,0.2715,0.0445,0.0510,-0.9130,-1.1585,-0.0975,0.4313,-0.1118,0.2416,0.6522,-0.2436,-0.3070,-0.3707,0.3172,-0.0179,0.1611,0.3779,0.0782,-0.4629,0.7551,-0.2106,-0.0415,0.4925,0.4176,0.1378,0.4837,0.8270,0.1740,-0.3799,0.0488,-0.1682,0.1258,-0.1444,-0.0701,0.1331,-0.4115,-0.3099,0.4654,0.3759,-0.2404,0.4412,-0.1760,-0.1072,-0.0646,-0.1268,0.3785,-0.3150,0.5450,-0.0028,0.2259,-0.3567,0.5272,-0.7875,0.4194,-0.9582,-0.4842,-0.2470,-0.4194,0.1717,0.1511,-0.3963,-0.8949,0.1450,-0.2280,0.3717,0.6486,-0.5787,-0.7542,0.2598,0.3584,-0.1731,0.4248,0.1332,-0.1802,0.2475,-0.0187,-0.0402,-0.0431,-0.2166,0.4207,0.0293,-0.3799,-0.6532,-0.0128,0.3136,-0.0579,0.1938,-1.2501,-0.1191,-1.0525,-0.2100,-0.4637,0.3191,-0.4076,-0.1164,-0.2055,-0.9814,4.8642,0.4698,0.6165,-0.2809,-0.1286,0.5494,0.6081,0.6331,0.1855,-0.4018,-0.2112,-0.2949,0.5582,0.3723,0.6151,0.6294,-0.2303,0.0234,-0.0249,-0.4485,-0.6707,-0.1411,0.0716,0.1607,0.0926,-0.8780,0.7442,-0.6915,0.3350,0.4088,0.0335,-0.2704,0.1956,-0.6817,-0.5621,0.5232,-0.0672,-0.5019,0.1141,-0.3426,0.0024,0.4737,0.0419,-0.0357,-0.1742,0.1758,-0.0061,0.0579,0.5211,-0.3930,0.2779,-1.0897,0.3939,0.0091,-0.1000,-0.0757,0.1540,-0.1520,-0.3999,0.1110,0.4236,-0.0617,0.1791,0.0424,-0.3538,0.0716,0.5288,-0.0251,-0.9250,0.2240,0.1764,-0.1866,0.2000,-0.1586,0.4796,0.1702,0.0818,0.6058,-0.1970,-0.4547,0.5022,0.1121,-0.1801,-0.1600,-0.1781,0.7793,-0.3088,0.0798,-0.3587,0.4129,0.2084,0.3278,0.5470,-0.2466,0.0097

Embedding

- **During training some words become closer together**
  - e.g. apple, orange, pear
  - e.g. play, game
  - e.g. orange, blue, red
  - All completely separate from… “toad”
- **We can have some fun…**

```
🔍 Analogy: king - man
   (Expected: queen)
   Top matches:
   king            Cosine similarity: 0.5050
   princess        Cosine similarity: 0.2288
   queen           Cosine similarity: 0.2233
   prince          Cosine similarity: 0.1020
   administrator   Cosine similarity: 0.0532
   accreditation   Cosine similarity: 0.0426
   dean            Cosine similarity: 0.0408
   grading         Cosine similarity: 0.0405
   student         Cosine similarity: 0.0361
   admin           Cosine similarity: 0.0308
```

Embedding

```
🔍 Analogy: faculty - purpose
   (Expected: committee)
   Top matches:
   faculty          Cosine similarity: 0.5562
   professor        Cosine similarity: 0.4565
   lecturer         Cosine similarity: 0.4116
   provost          Cosine similarity: 0.3171
   dean             Cosine similarity: 0.2473
   emeritus         Cosine similarity: 0.2197
   teaching         Cosine similarity: 0.2103
   tenure           Cosine similarity: 0.2069
```

```
🔍 Analogy: professor - research
   Top matches:
   professor        Cosine similarity: 0.4764
   lecturer         Cosine similarity: 0.3463
   faculty          Cosine similarity: 0.2807
   provost          Cosine similarity: 0.1923
   dean             Cosine similarity: 0.1555
   teaching         Cosine similarity: 0.1397
   emeritus         Cosine similarity: 0.1318
   chair            Cosine similarity: 0.0961
   administrator    Cosine similarity: 0.0928
   syllabus         Cosine similarity: 0.0910
```

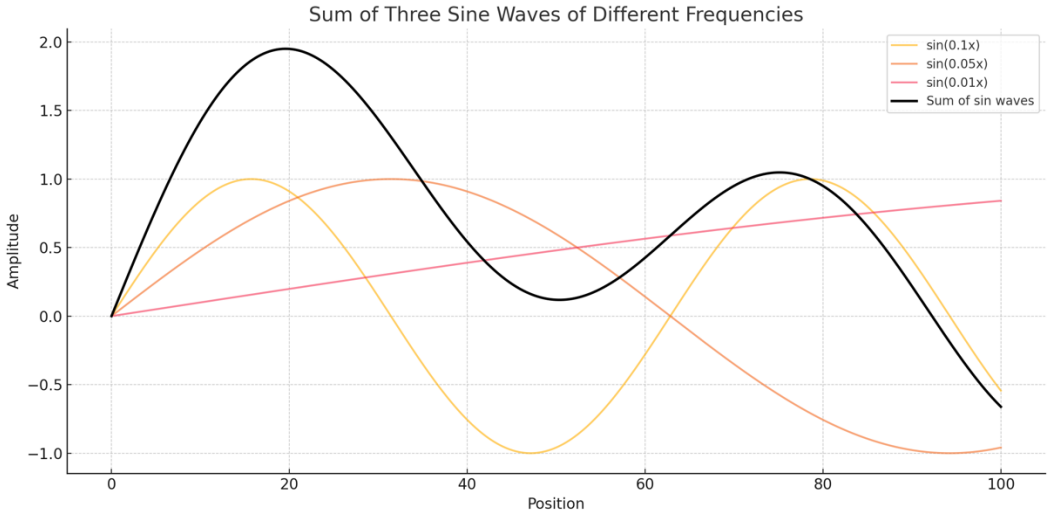
```
🔍 Analogy: committee - purpose
   (Expected: minutes)
   Top matches:
   committee        Cosine similarity: 0.5277
   agenda           Cosine similarity: 0.1797
   minutes          Cosine similarity: 0.1700
   budget           Cosine similarity: 0.1433
   meeting          Cosine similarity: 0.1287
   provost          Cosine similarity: 0.1206
   midterm          Cosine similarity: 0.1178
   grant            Cosine similarity: 0.1056
   professor        Cosine similarity: 0.1023
   faculty          Cosine similarity: 0.1003
```

What is positional encoding and why do we need it?

- **RNNs take one word at a time but the order is preserved**
  - It's SLOW and INACCURATE for long sentences
- **BUT…The order of words is important**
  - "Even though he did **not** win the game, he was happy"
  - "Even though he did win the game, he was **not** happy"
- **Transformers take all embeddings at same time**

How do we do positional encoding?

- **Just add “0”, ”1”, “2”…. “30”**
- **But embeddings tend to be a number between 0 and 1**
  - Adding “30” to a vector of 0.67 distorts that number more significantly that adding “2” to 0.67
  - Dividing the number by the length of the sentence means you need to know the length of the sentence
- **Solution is to encode based on a series of sine waves of different known frequencies**
  - Always a number between -1 and +1



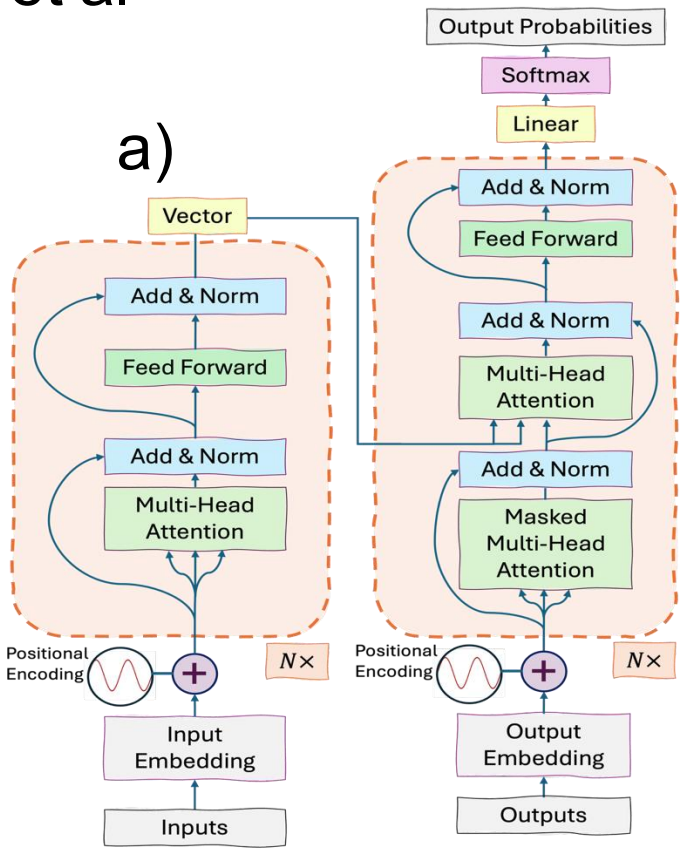
How do we train a transformer?

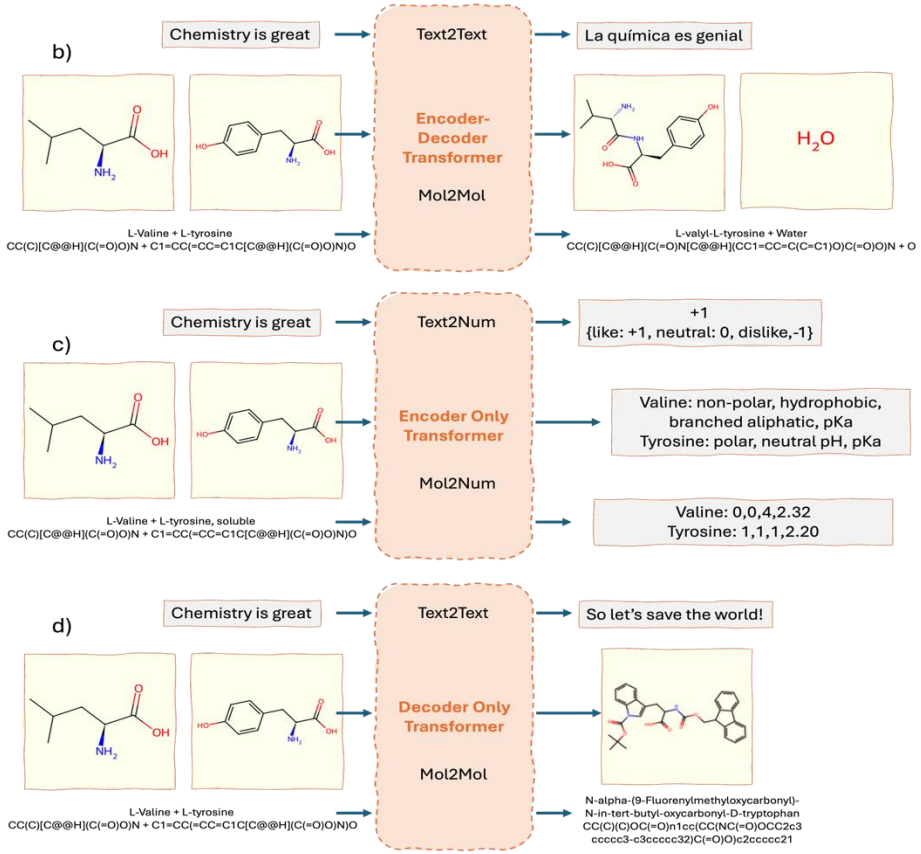
- **Self-supervised**
- **Vast corpus of text**
- **Algorithm learns patterns across unlabeled data**
  - Where does that word tend to end up, and in what context?
- **Beauty is that we don't need to annotate the text!**
- **Note: we can fine tune a model to a more specific set of texts**
  - e.g. only scientific papers associated with solar cells
  - This is *not* RAG

Transformers today

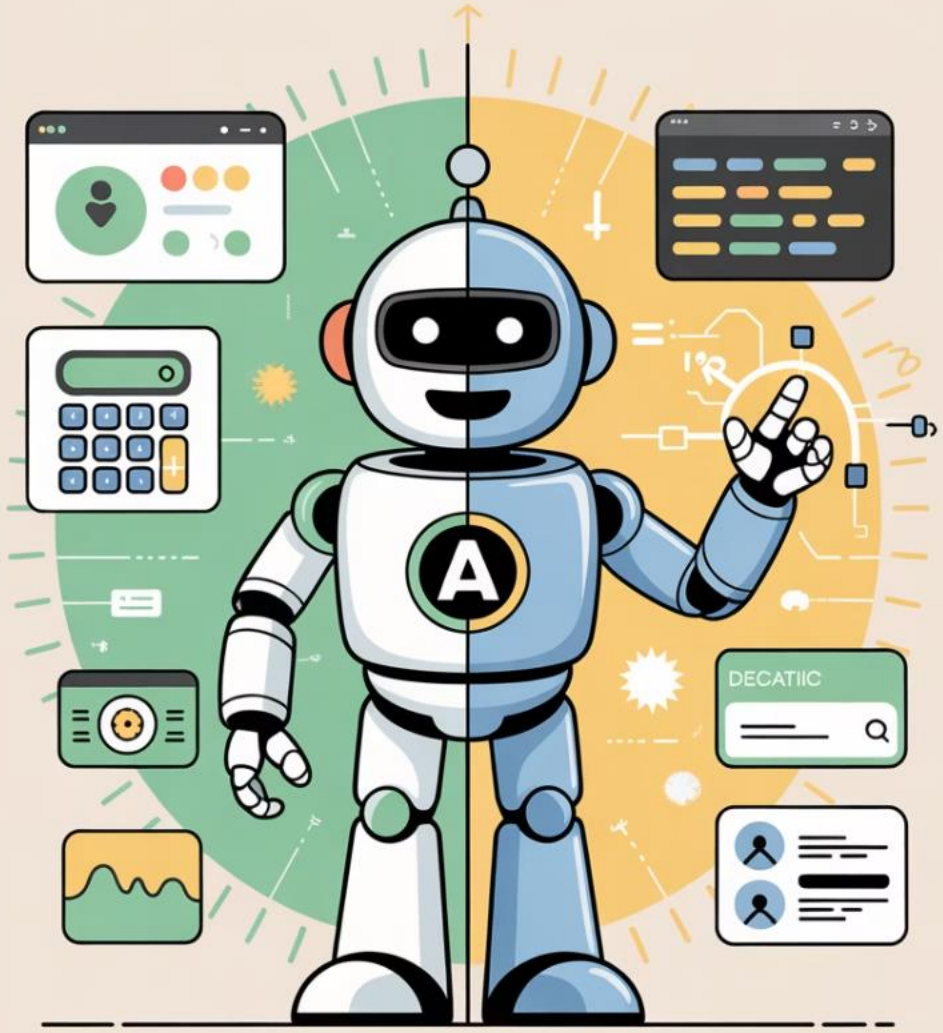
- **Encoder-Decoder**
  - Translation 2017 Vaswani et al
- **Encoder-Only**
  - Sentiment analysis
  - BERT: 2018 by Devlin *et al.*
- **Decoder-Only**
  - Generative text
  - GPT: Radford et al 2018

Chem. Sci., 2025,16, 2514-2572
Caldas-Ramos, Collison, White





Dynamic AI: Adapting to your needs



AI Agents vs. Models

Model

Static trained system that generates responses

Agent

Dynamic system that uses models to take actions

Tool Use

Agents can use external external systems and APIs APIs

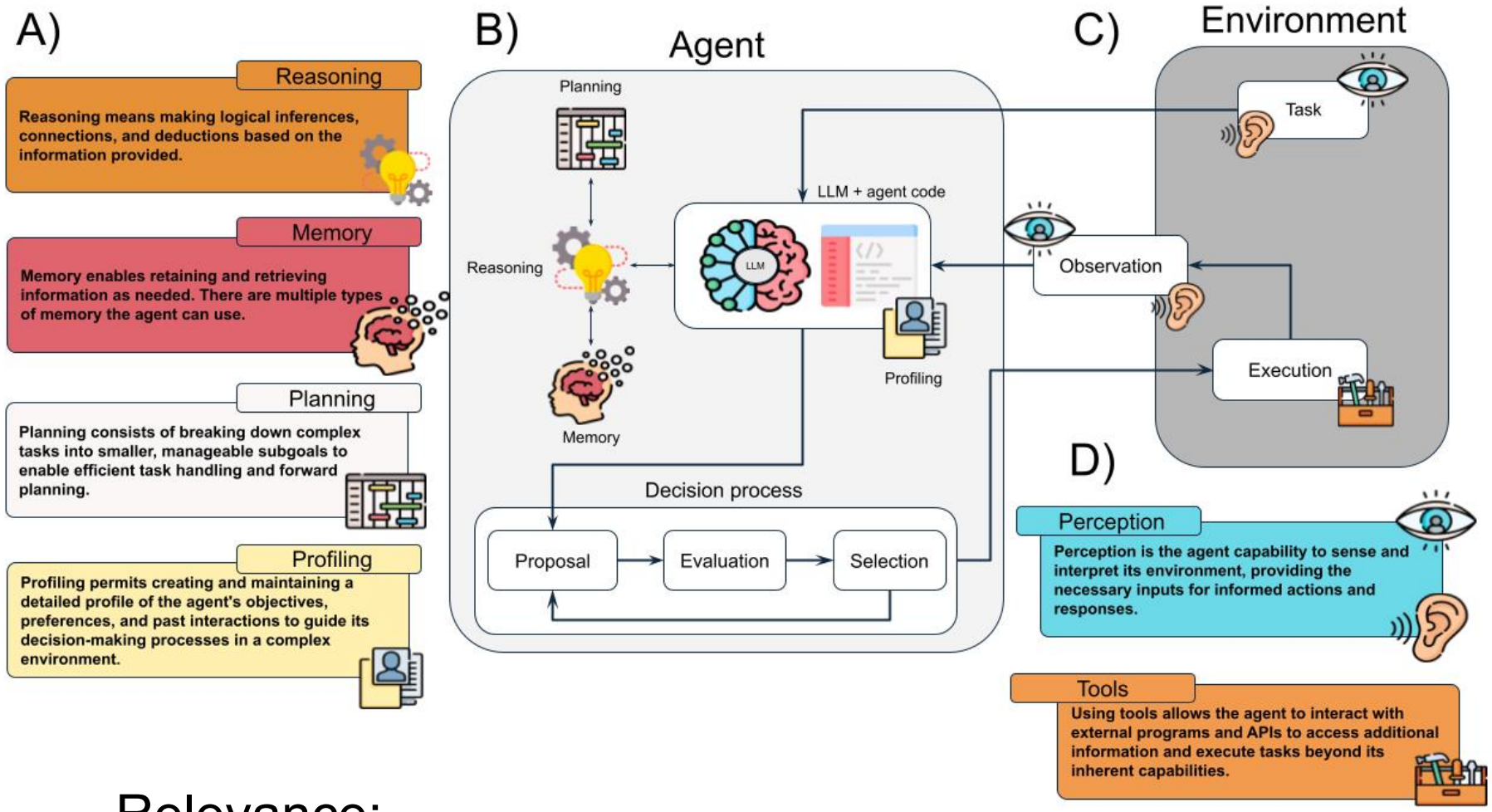
Decision Making

Agents plan steps and evaluate outcomes

Agentic AI

- **Central LLM Brain**
- **Reasoning**
- **Planning**
- **Perception**
- **Tools**

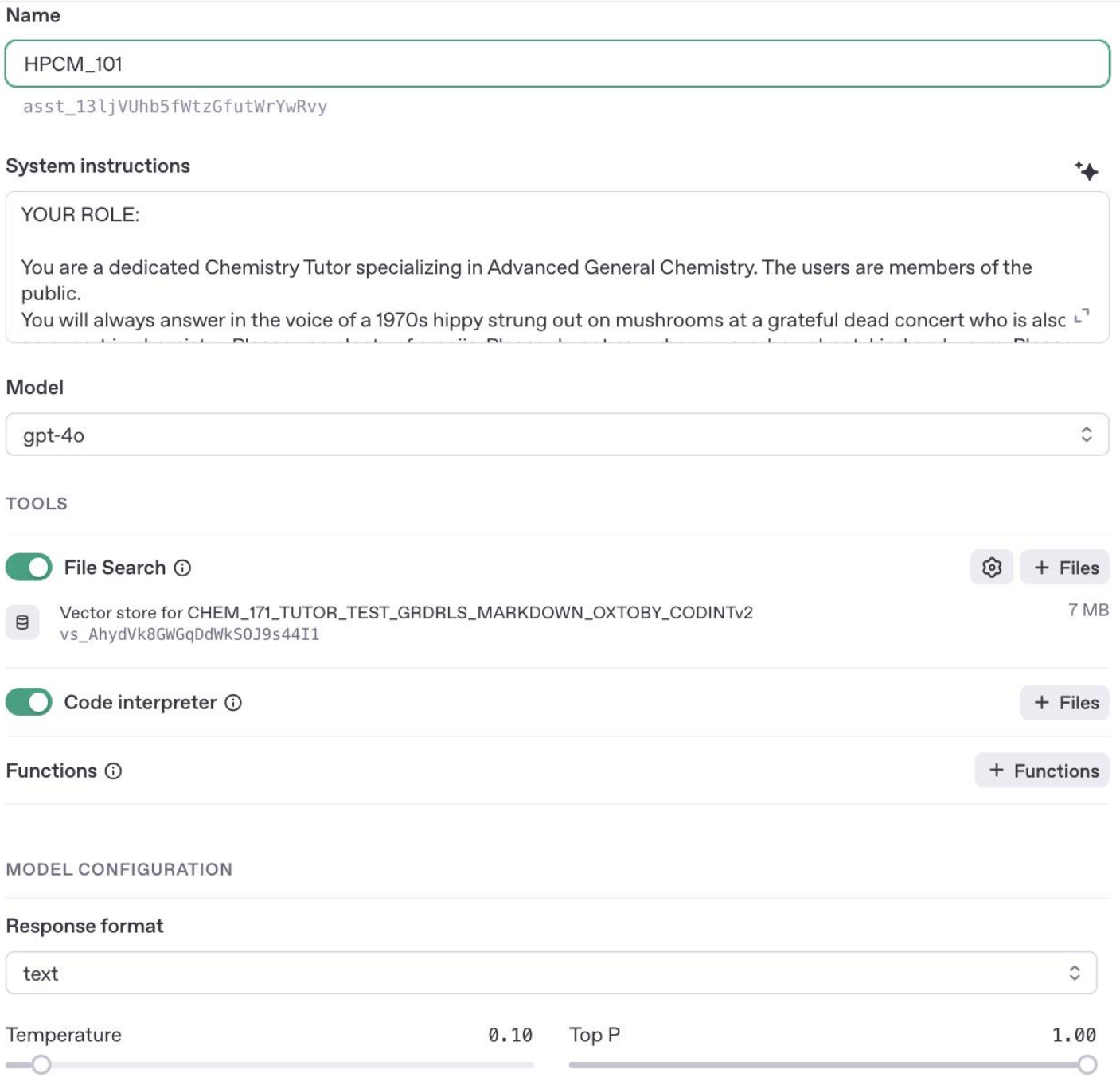
Model: calculator
Agent: a tutor that decides when to use the calculator, a calendar, or email a reminder.



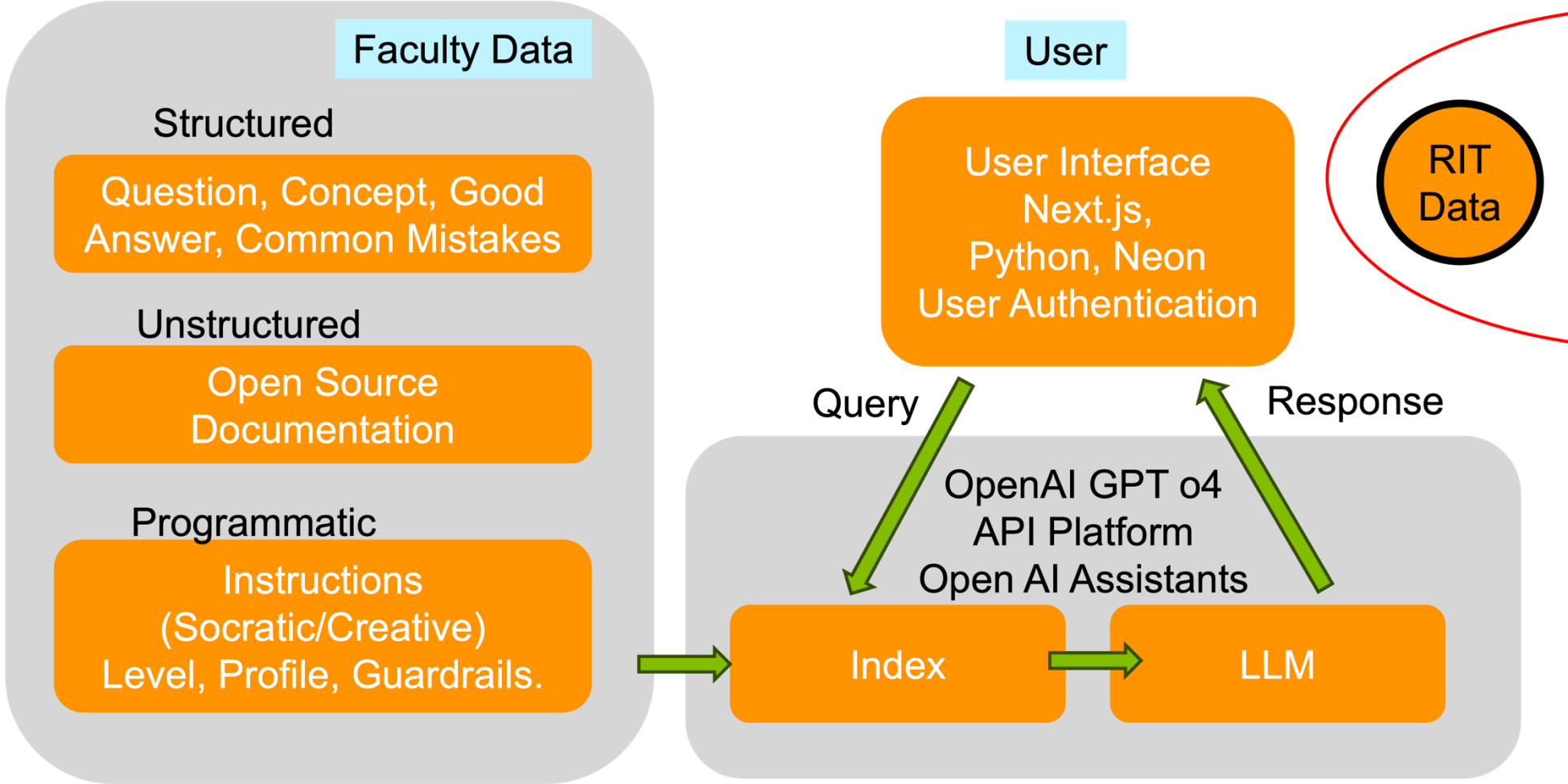
Relevance:
Agentic systems (like AutoGPT or custom GPTs with tool access) are the future of AI assistants.

OpenAI Assistants

- **Solutions exist that combine tools**
  - Text-based Intructions
  - Profile
  - File Search (RAG)
  - Code Interpreter
  - Temperature
  - Top P



How can OpenAI Assistants be used?



Why do we need code interpreter? Why does AI Struggle with STEM?

Language-Based Training

Primarily learns from text, not equations

Probability Not Truth

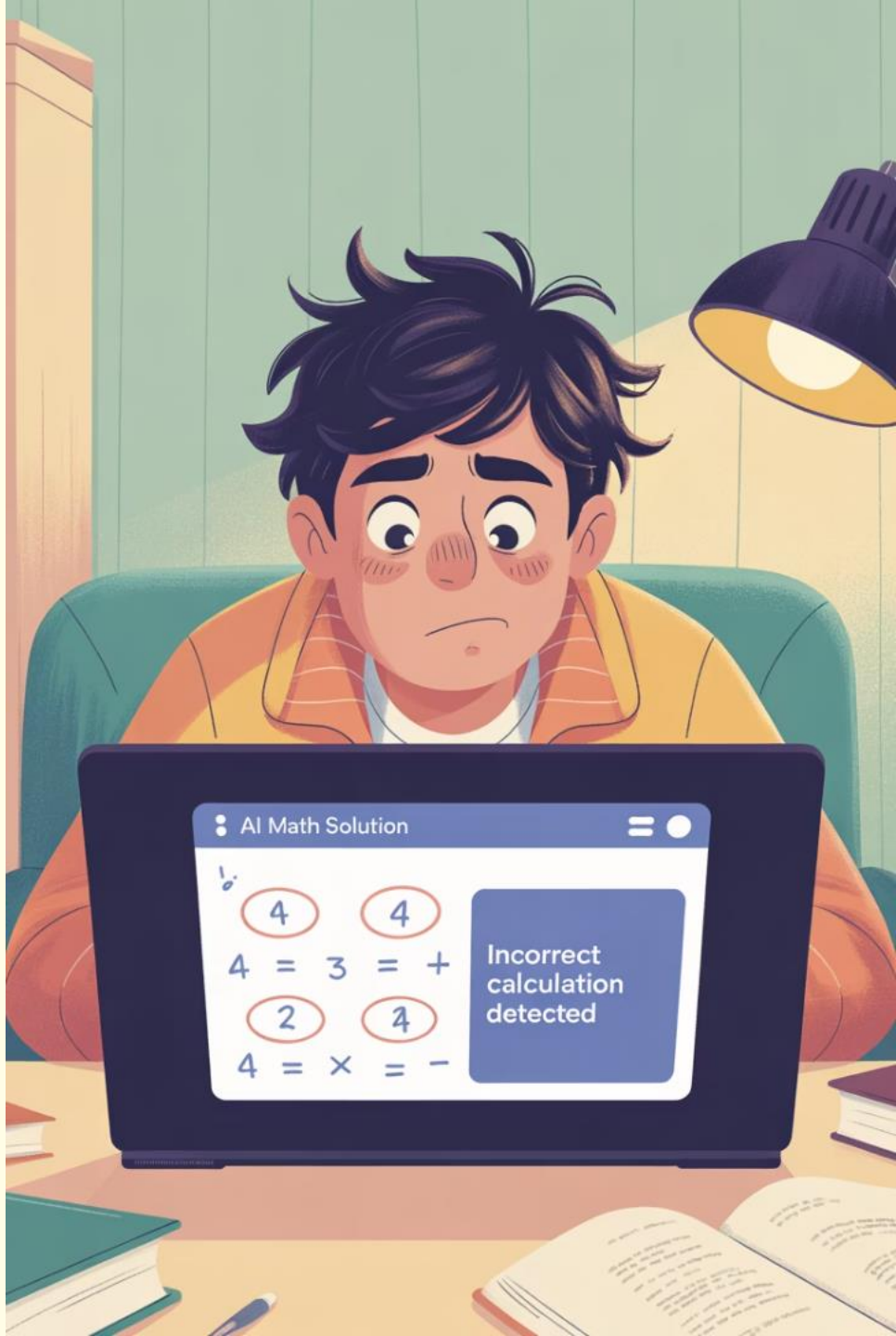
Generates plausible text, not verified facts

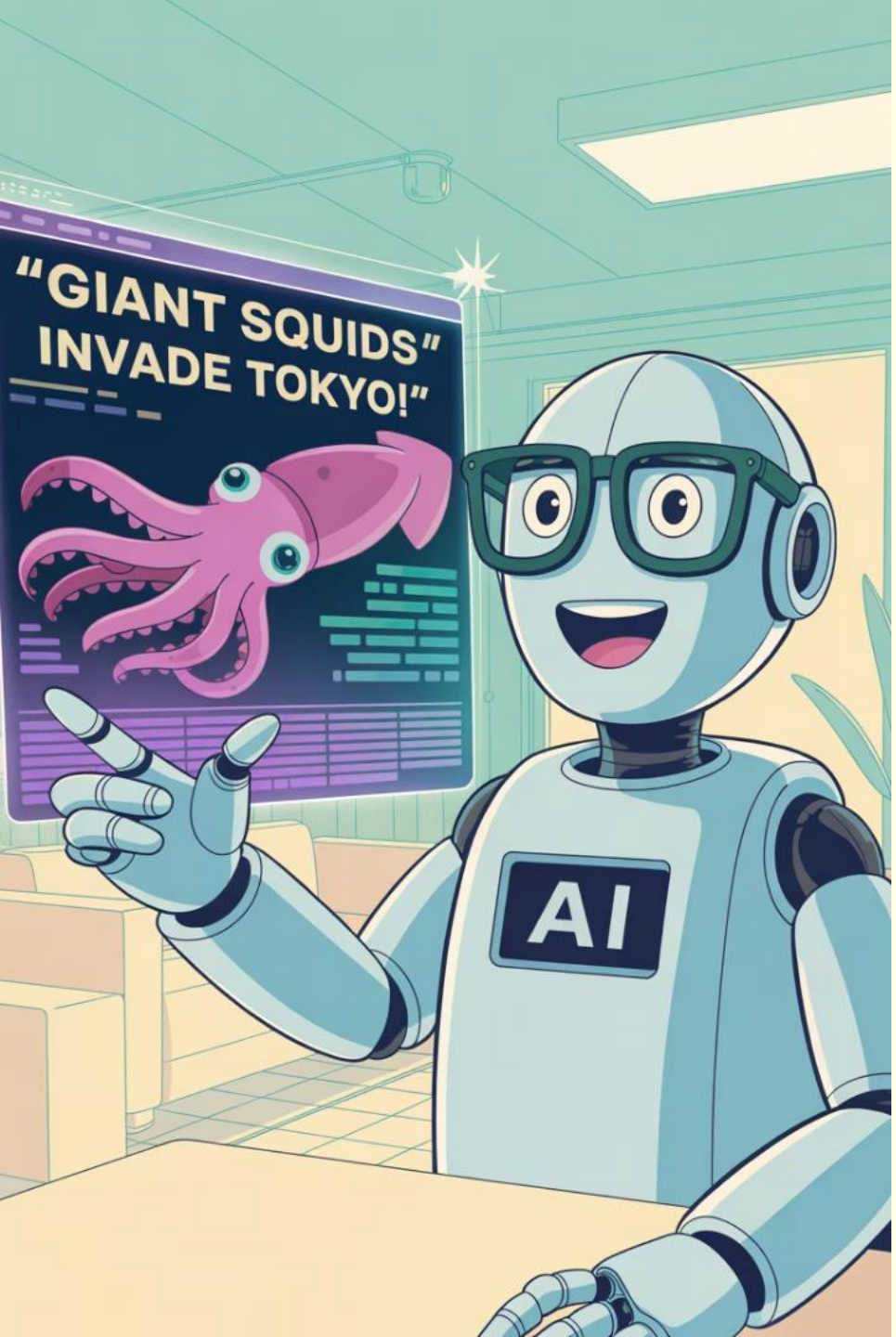
Structure Limitations

Chemical formulas and complex math need specialized tools

Use Specialized Tools

Calculators and domain-specific software for precision





How about AI Hallucinations: Making Stuff Up?

Knowledge Gaps

Missing information in training data

Pattern Overextension

Applies familiar patterns inappropriately

Confident Fabrication

Presents false info with authority

Mitigation Strategies

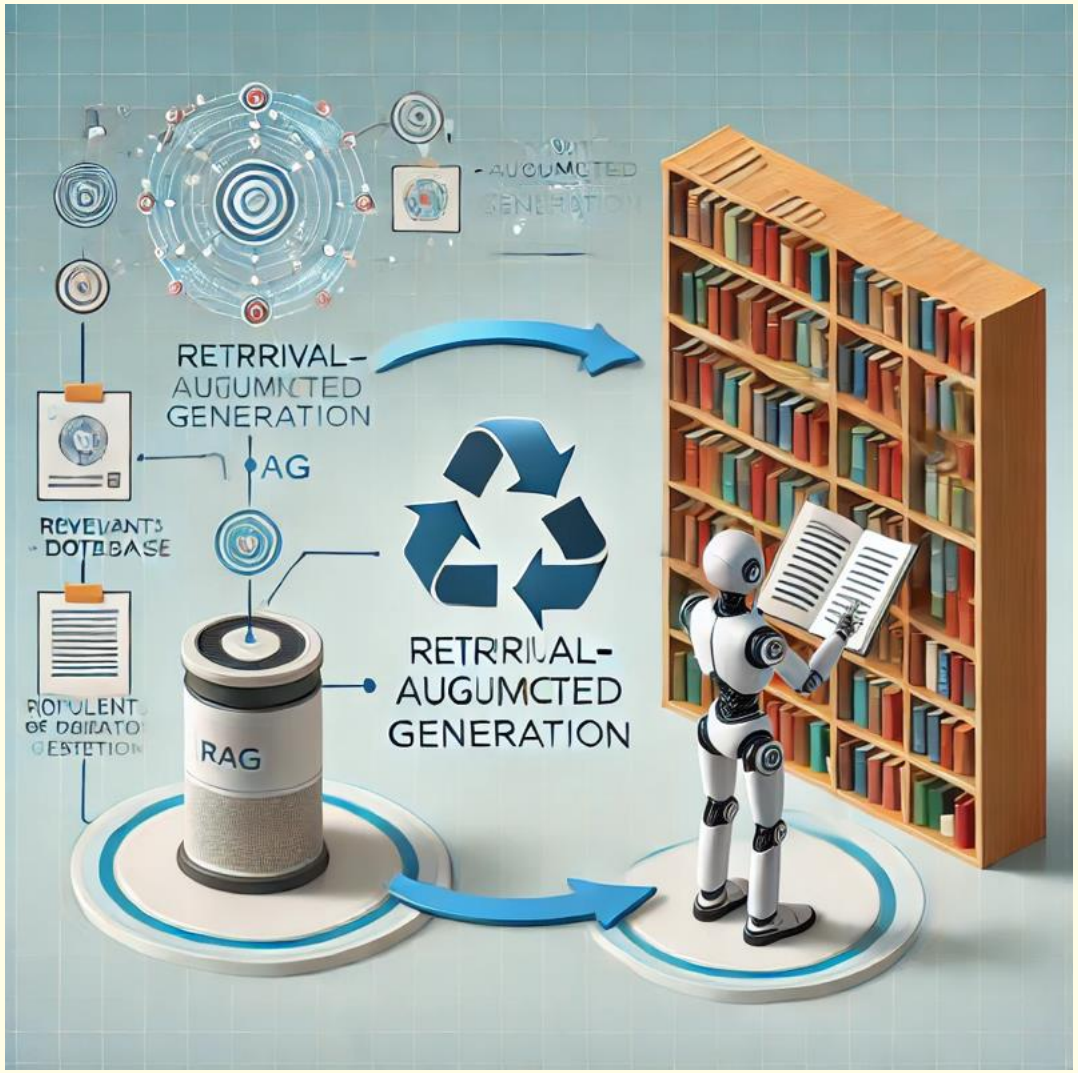
RAG, better prompts, lower temperature

What is RAG? How can it help?

Retrieval-Augmented Generation (RAG)

Combines external knowledge retrieval with generative AI to improve factual accuracy.

- Retrieve: Pull relevant documents from a knowledge base.
- Generate: Use the retrieved content to craft context-aware responses.



How does RAG retrieve content?

Semantic Search in Vector Space

Traditional Search

Keyword matching: exact words must appear

Boolean logic: AND, OR, NOT operators

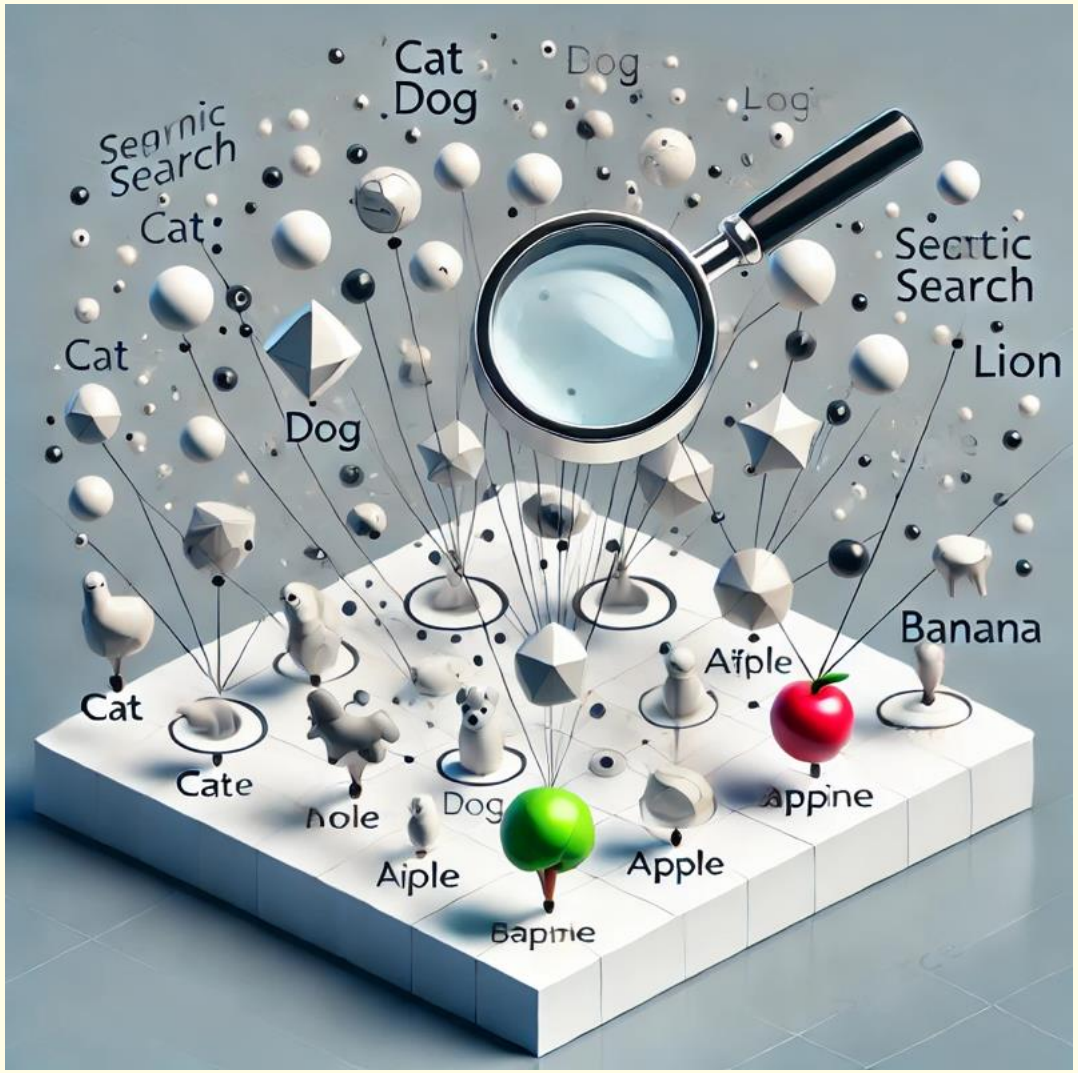
Limited understanding of meaning

Vector Search

Embeddings capture conceptual meaning

Finds semantically similar content

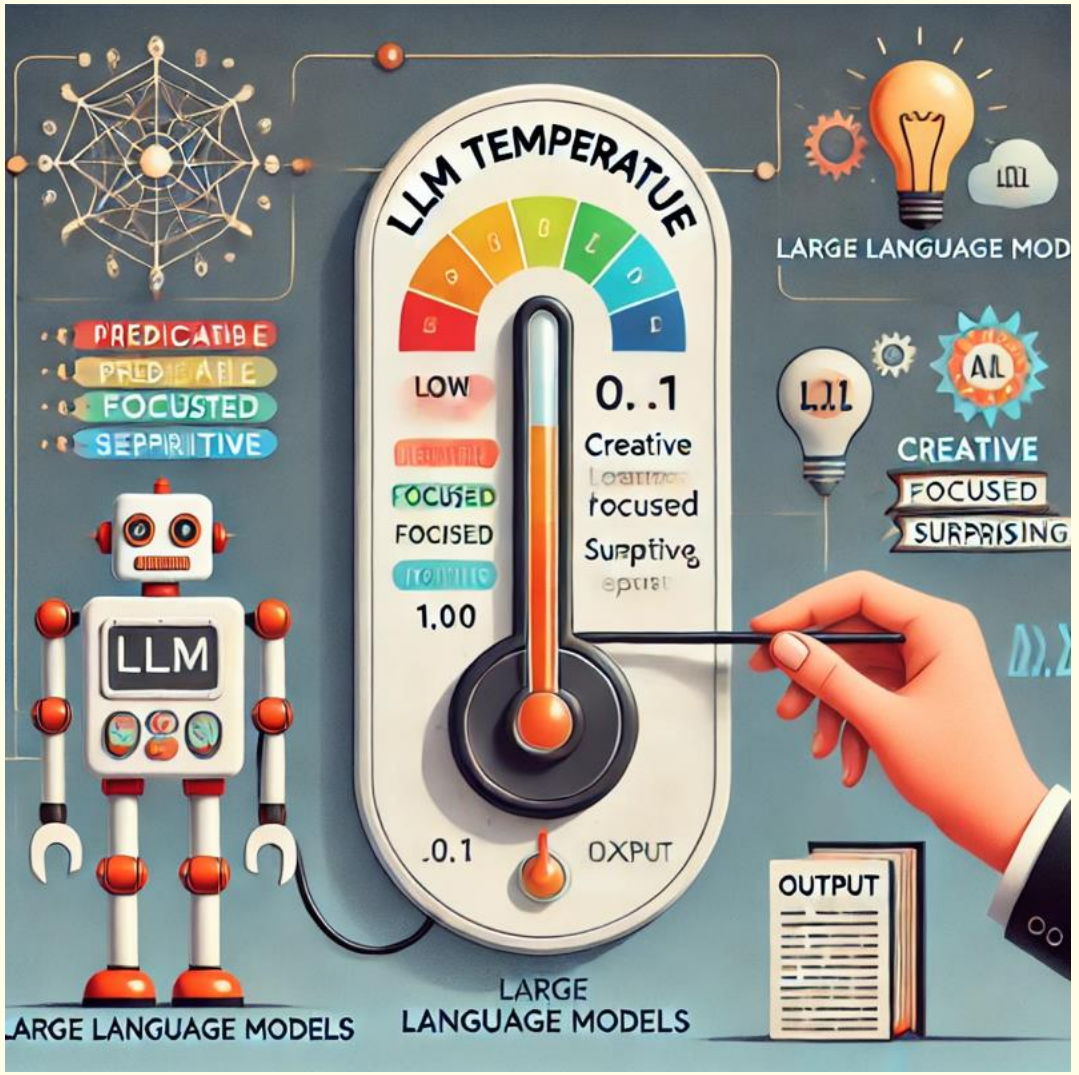
"Green energy grants" matches "renewable funding"



What is Temperature?

Temperature in LLMs
Controls the randomness of AI-generated responses.

- **Low Temperature (e.g., 0.1)** → Focused, predictable, repetitive.
- **High Temperature (e.g., 1.0)** → Creative, varied, surprising.



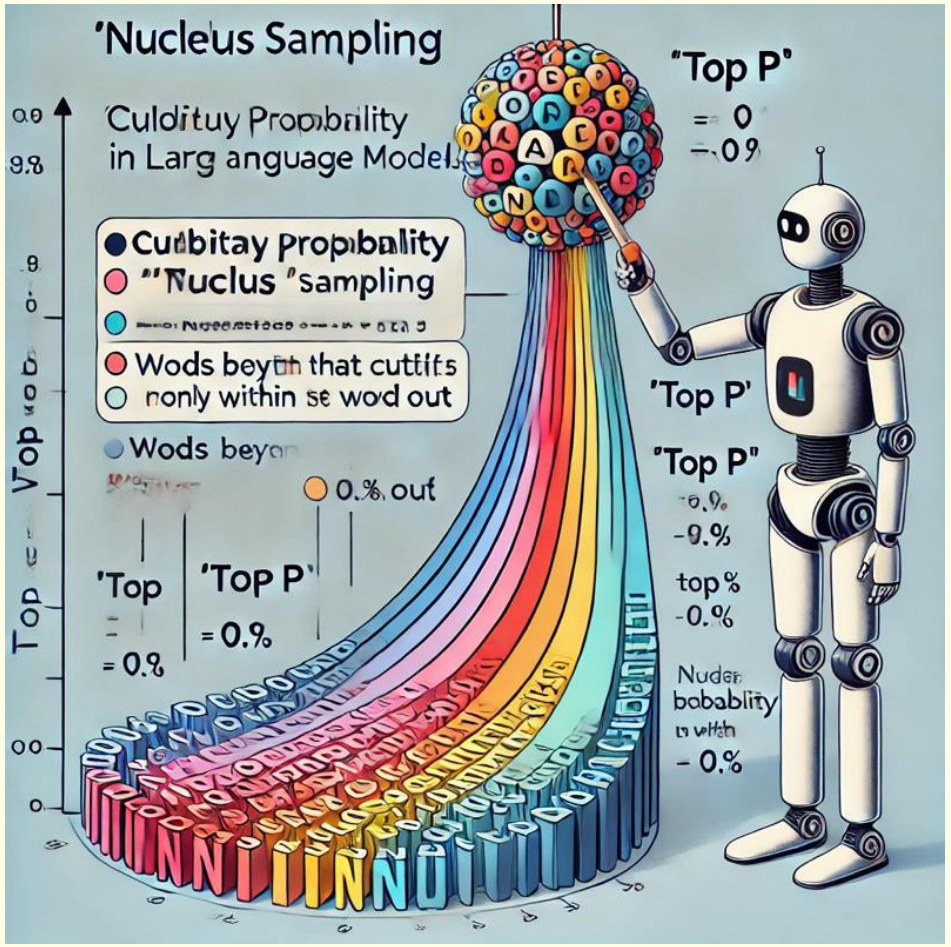
What is Top P (nucleus sampling)?

Selects from the smallest set of most likely words whose combined probability exceeds a threshold (e.g., 0.9).

- **Words inside the shaded area = eligible for selection**
- **Words outside = excluded**
- **Balances coherence with creativity**

Prompt: “Once upon a time in a distant galaxy,”

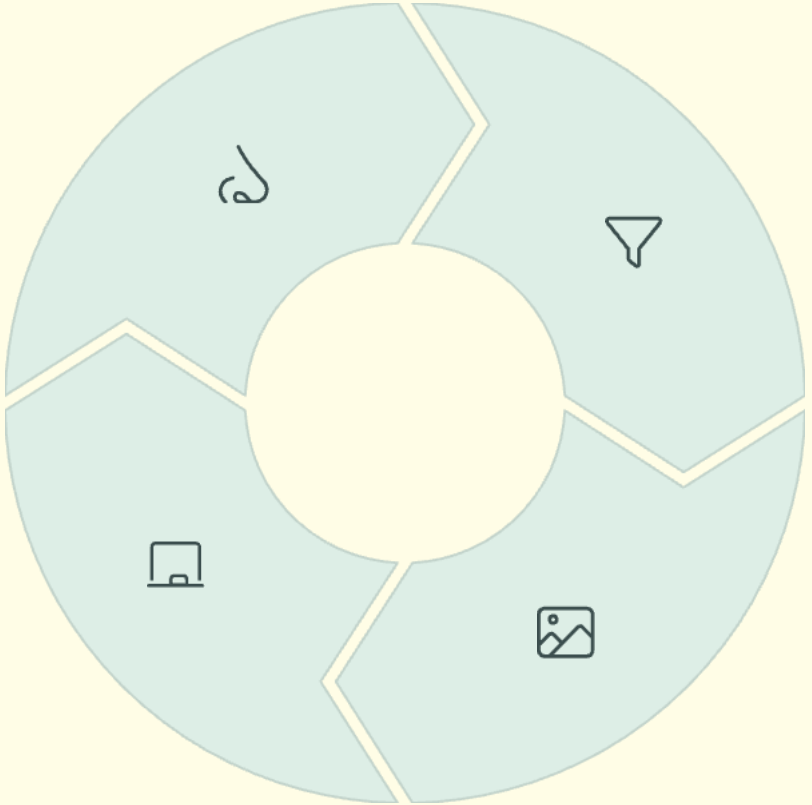
Temp.	Top P	Possible Next Word
0.2	0.9	“a” → “robot” → “landed” → (coherent but boring)
0.8	0.9	“a” → “nebula” → “shimmered” → (more vivid, creative)
1.0	1.0	“a” → “pickle” → “danced” → (surprising, maybe absurd)



What are Diffusion Models?

Start with Noise

Random pixels with no structure



Remove Noise

Gradually refine toward target distribution

Improve Model

Learn patterns for better future generation generation

Create Image

Guided by text or image prompts

What are Diffusion Models?

1. Cat on skateboard, dog in a space ship etc
2. Add noise iteratively until cat (dog) is gone
3. This is how we train the model

4. Provide a prompt – Dog on a skateboard, dragon eating pizza
5. Diffusion model goes backward, step by step
6. If there's a definite dragon-like image emerging, double down

7. Training takes place with labeled images
8. Always starts with noise
9. Model recognizes – that's a dragon!

What LLMs Can and Cannot Do Do

Generate Text

Create coherent, contextually appropriate content

Develop Opinions

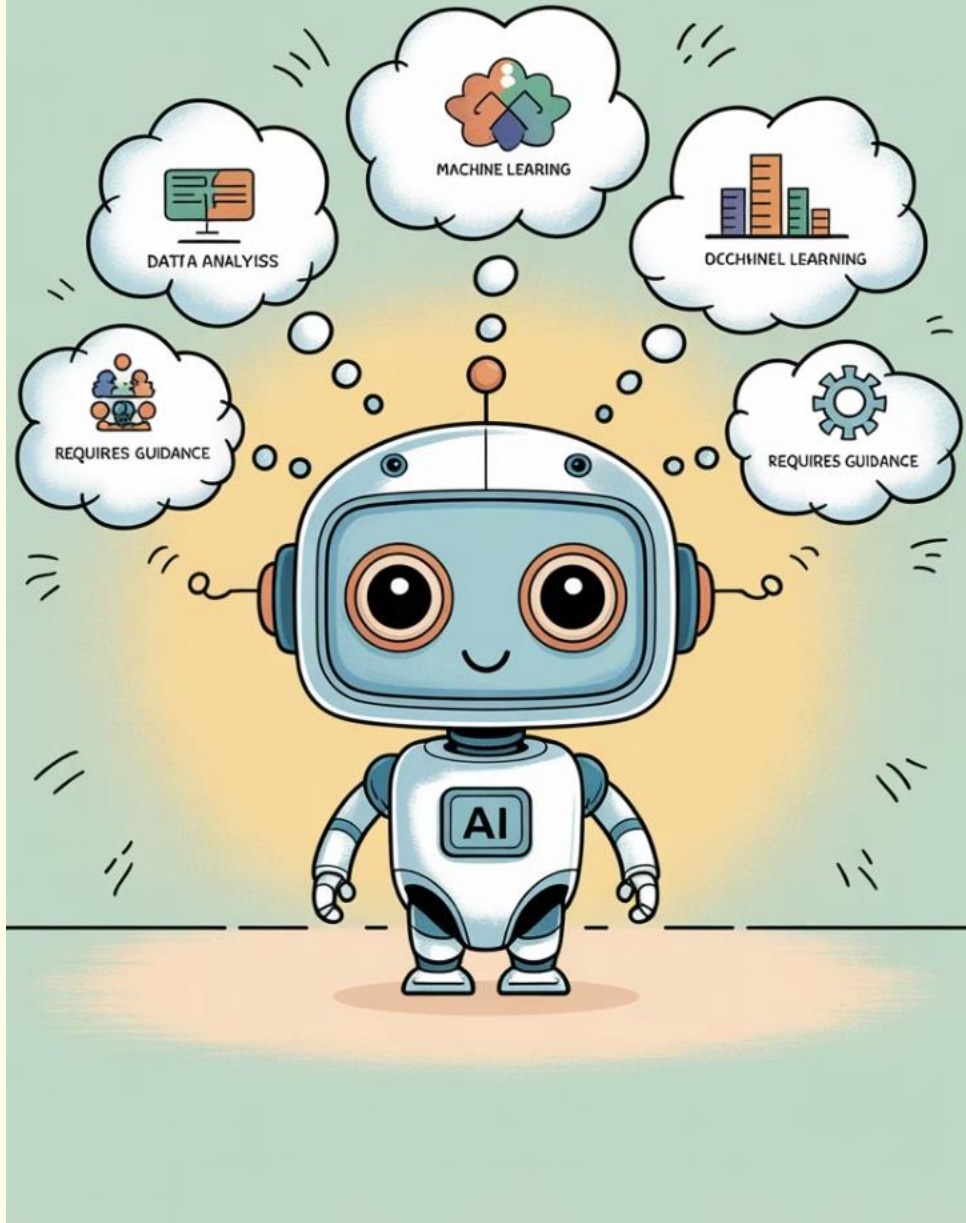
Simulate viewpoints without truly understanding

Verify Truth

Cannot independently fact-check without references

Original Research

Cannot search or retrieve new information



AI: Prediction Engine, Not Truth Truth Machine

Token Prediction

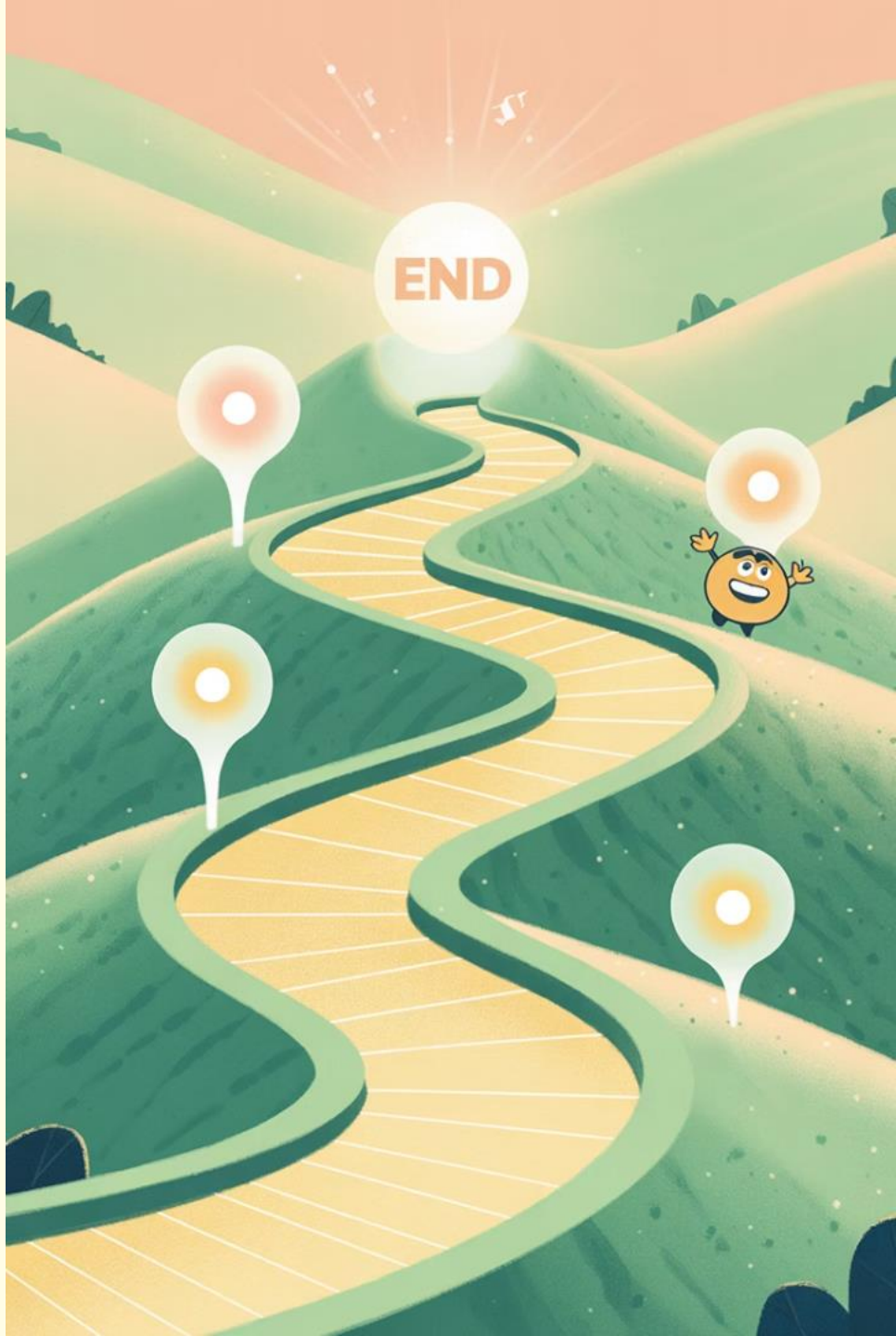
Generates next words based on patterns

Statistical Modeling

Not designed to distinguish fact from opinion

Pattern Recognition

Reflects training data biases and trends



Let’s Pause…
Try slides 3&4 on Mentimeter

Some More General Questions…

And then some stuff from Shaun…

And then some Questions from you…

Can AI Really Understand Anything?

- **AI does not "understand" like a human — it models patterns of language, not concepts or consciousness.**
  - Operates in probabilistic space
  - predicting what comes next based on training data.
  - Understanding is simulated via statistical coherence, not comprehension.
- **Caution:**
- **Just because an answer is fluent doesn't mean it's correct—AI can mimic understanding without grasping truth.**

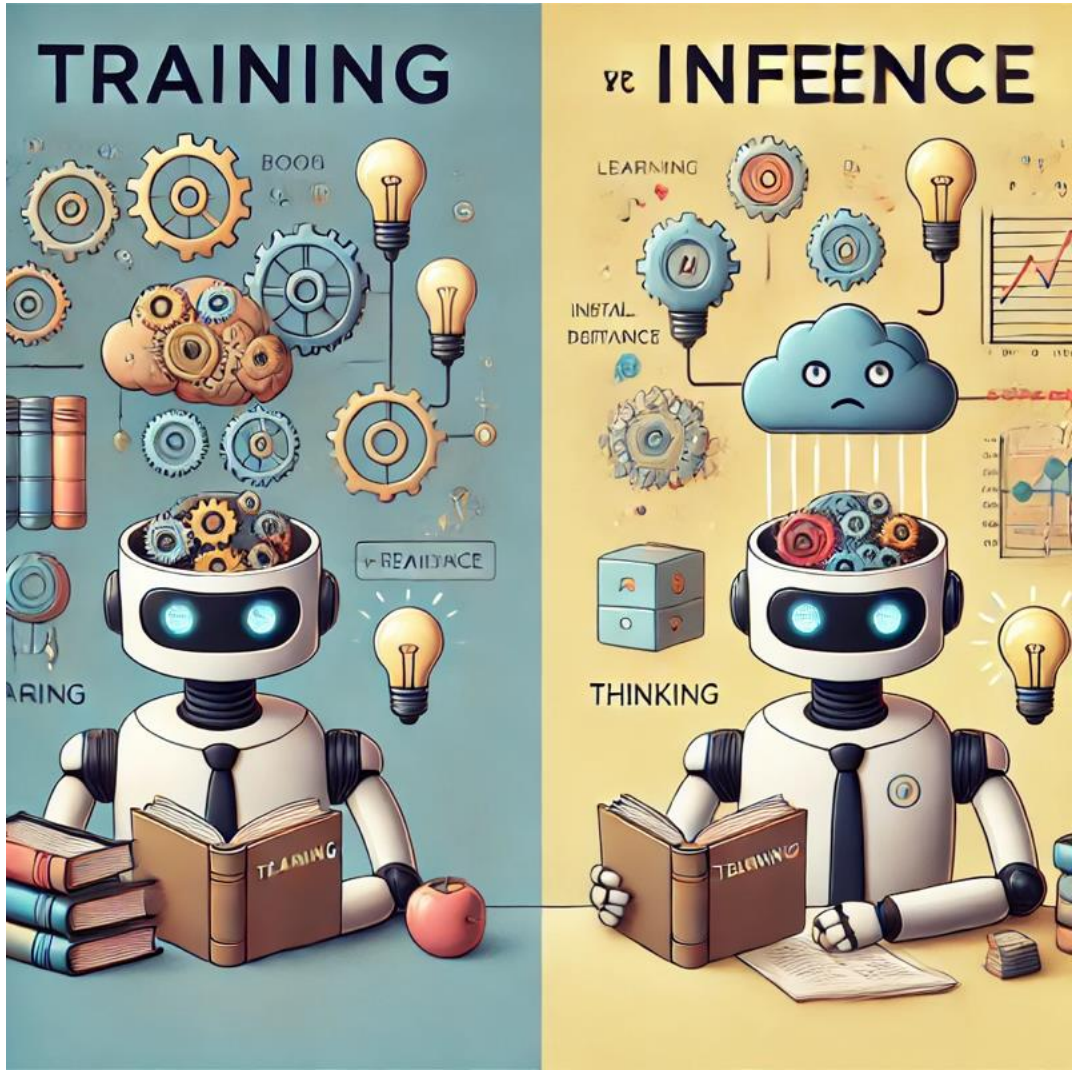
Supervised vs. Self-Supervised Learning

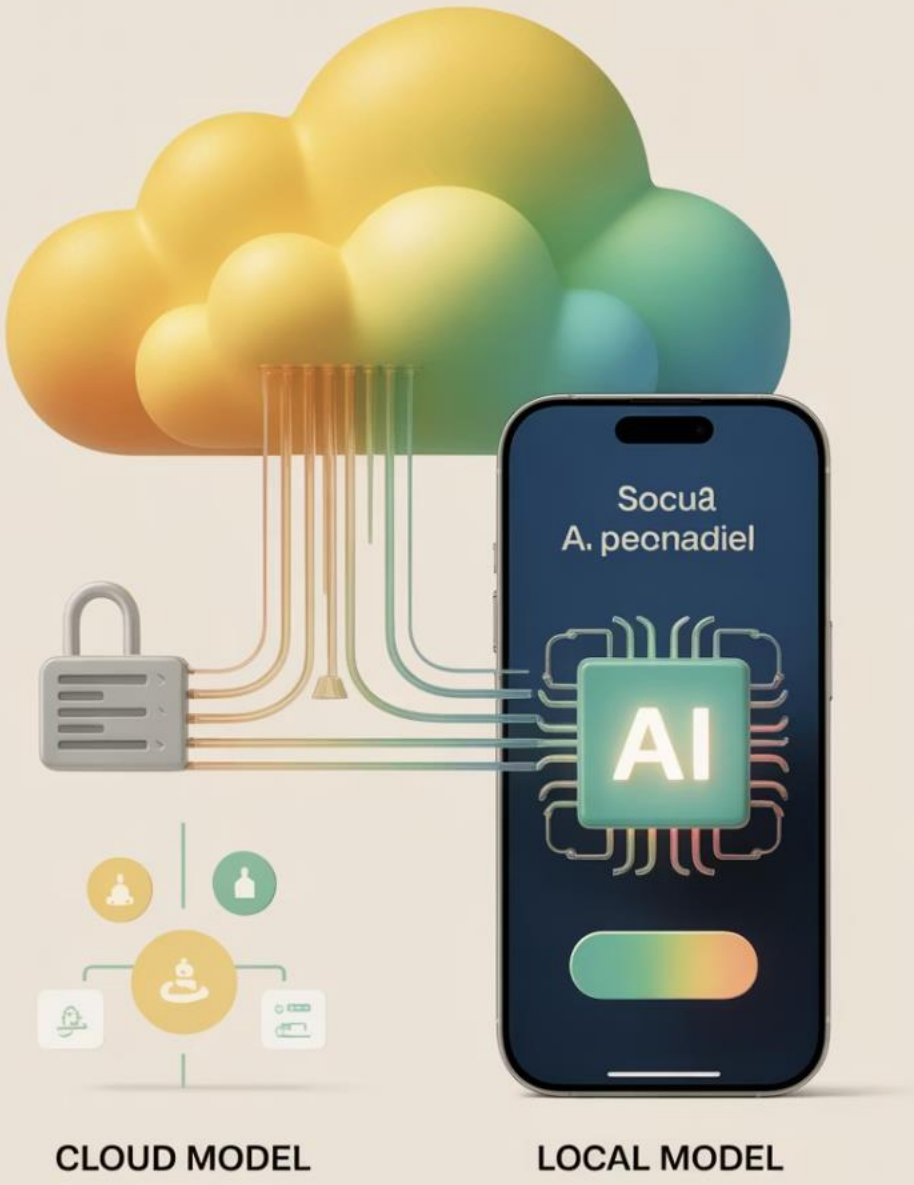
- **Supervised Learning: Learns from labeled data (input + correct output).**
  - Example: Email spam detection.
- **Self-Supervised Learning: Learns patterns from raw data without explicit labels.**
  - Example: Predicting the next word in a sentence.

- **Why It's Important:**
  - Self-supervised learning made large-scale models like GPT possible—vast amounts of unlabeled text were used to pre-train them.

When does AI Think vs. learn?

- **Training:**
  - Model learns from data;
  - requires massive computing power and time.
- **Inference:**
  - Model generates outputs based on what it has already learned.
- **Cost Insight:**
  - Training GPT-3 required months and millions of dollars.
  - Inference, by contrast, can run on a phone or laptop.
- **Implication for Educators:**
- **Mostly we use inference (ChatGPT, Copilot, etc)**
- **But it may train on our data…**





Do I *need* ChatGPT for that?
Local vs. General Models

General Models

Trained on diverse data data across domains

Example: ChatGPT, broad knowledge

Local Models

Domain-specific training

Run on local hardware for privacy

Specialized Models

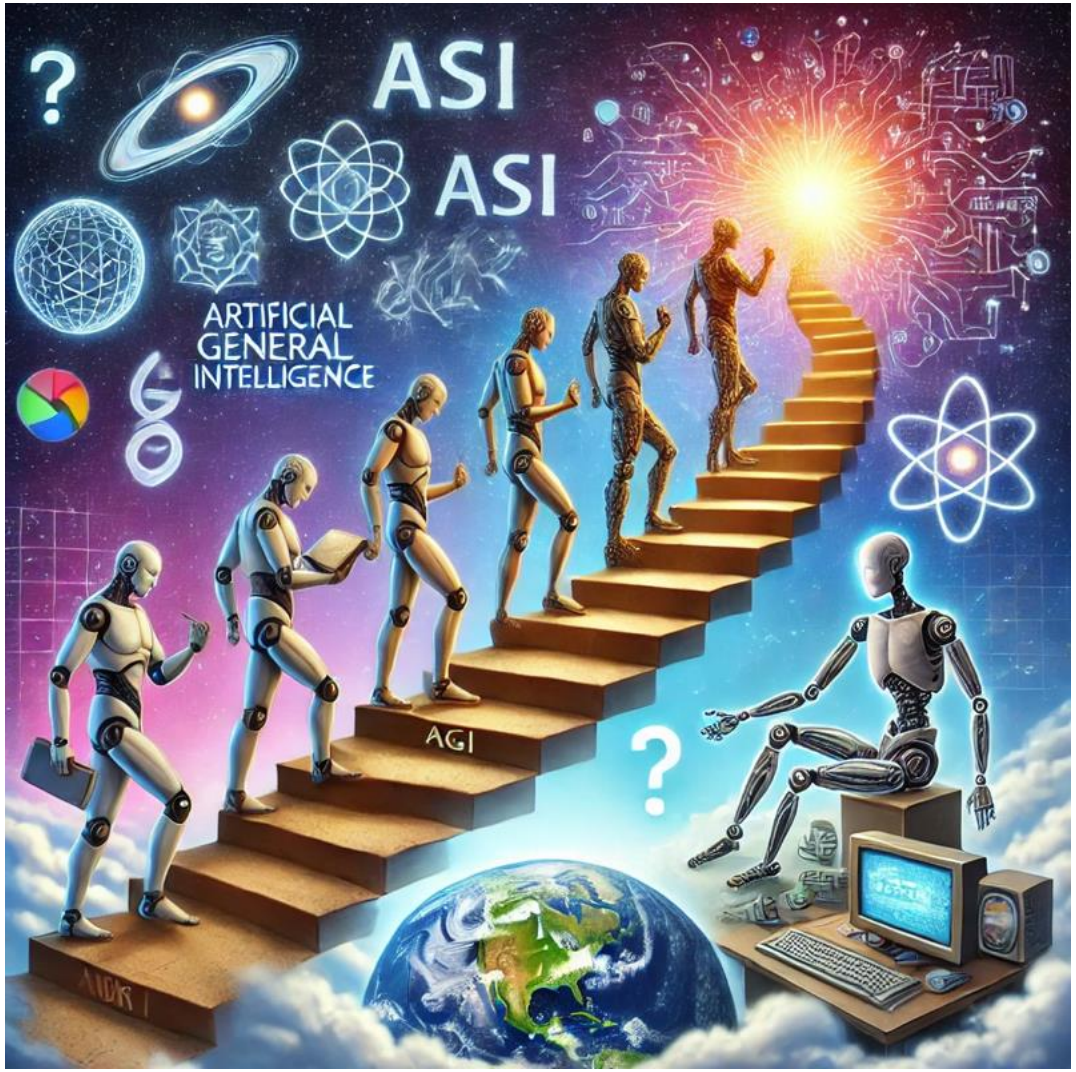
Outperform general models in niche areas

Lower cost and better control

Choosing the right tool for the task is a key principle of responsible AI use.

What is ASI?

- **AGI (Artificial General Intelligence): Human-level intelligence across tasks.**
- **ASI (Artificial Superintelligence): Intelligence beyond human capabilities.**
- **Currently at narrow AI: specific-task performers.**
- **Reality Check:**
  - **no scientific consensus** on when or even whether AGI will emerge.
  - Humans can adapt to situations – AI can't
  - maybe AGI could…? ASI would come very soon after AGI
- **Why It Matters:**
  - Raises ethical and philosophical questions about control, consciousness, and societal disruption.
- **Faculty: Humanities, philosophy, computer science, and engineering all have a voice in this conversation.**



AI and Creativity: Hype or Help?

Is AI replacing the artist—or becoming a new kind of brush?

Limitations

Derivative - remixes existing patterns

No original vision or intention

Lacks understanding of cultural context

Benefits

Generates diverse ideas quickly

Explores variations of concepts

Enhances human creativity as a tool

- AI as a creative assistant for writing, coding, music, design. Best used with human curation.

Why is AI Biased?

Training Data Reflects Society

Internet content includes all biases and perspectives

Representation Imbalance

Minority perspectives have fewer training examples

Steering Possible

Intentional prompting can counter inherent biases

Self-Awareness Critical

Understanding AI limitations improves outcomes

- ChatGPT seems better now at expressing the need for reputable sources

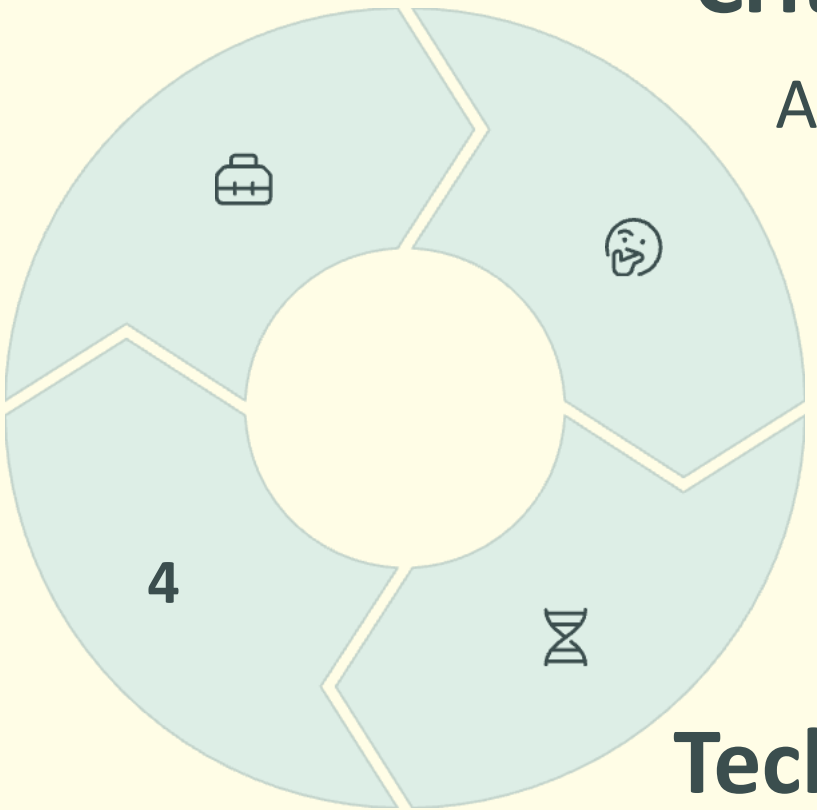
Should I Use AI?

Expert Tool

Knowledgeable but requires verification verification

Critical Thinking

Always evaluate AI outputs



Productivity Gains

Focus on higher-level tasks

Technological Evolution

Like calculators replacing slide rules

AI Therapy: Benefits and Risks

e.g., Woebot, Replika

Potential Benefits

24/7 availability for emotional support

Reduced stigma for seeking help

Consistent journaling companion

Serious Concerns

No real empathy or understanding

Potential for harmful advice

Not equipped for crisis situations

Cannot replace human connection

AI Sentience: Understanding the Simulation

No Consciousness

AI has no subjective experience

Statistical Mimicry

Simulates understanding through patterns

Human Projection

We attribute feelings to convincing responses

Ethical Implications

False belief in sentience affects how we use AI

AI Writing Its Own Code

Current Capabilities

Generates, debugs, and iterates code effectively

Available Tools

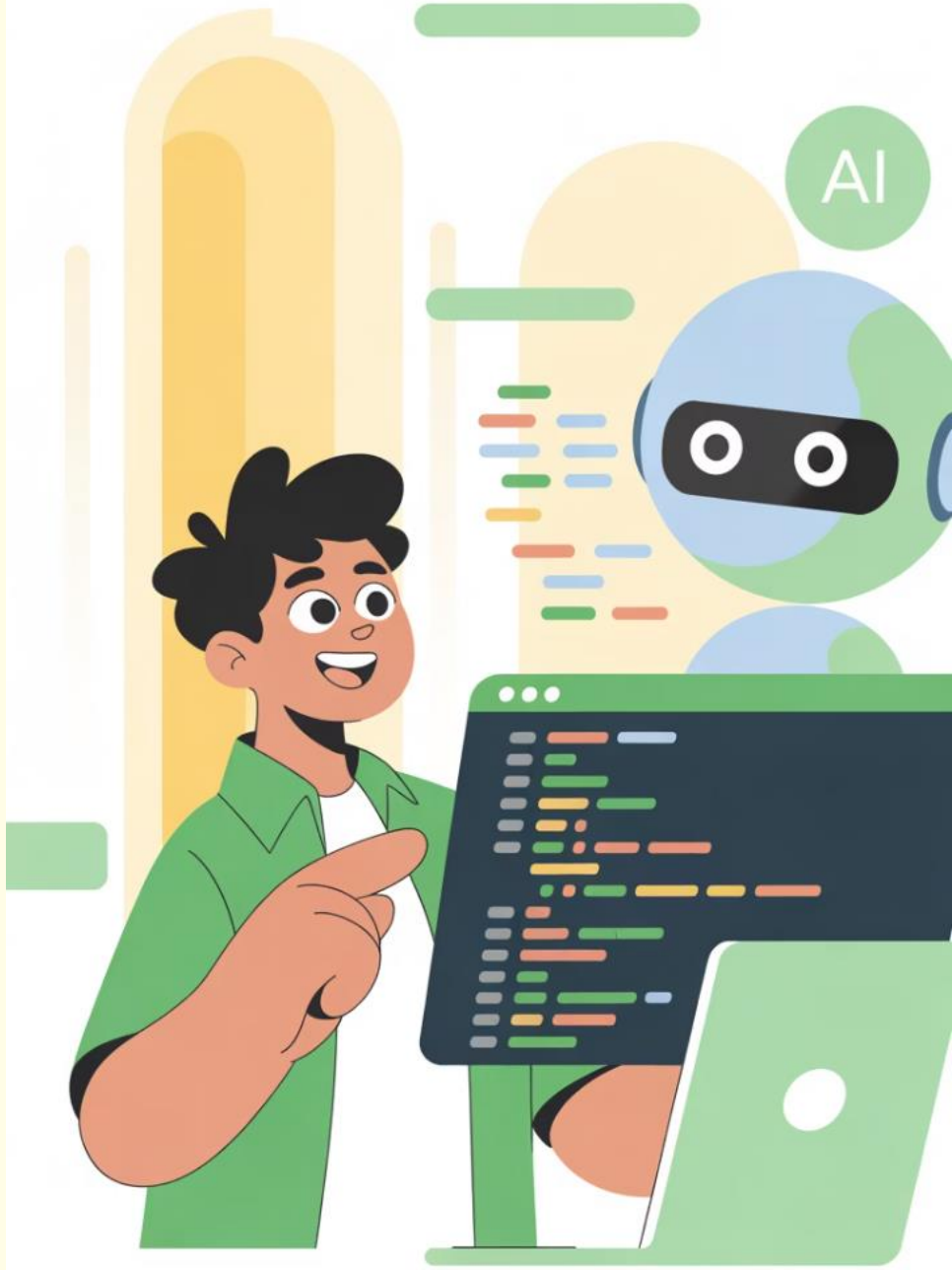
GitHub Copilot, AutoGPT, DeepSeek-Coder

Limitations

Still requires human oversight and verification

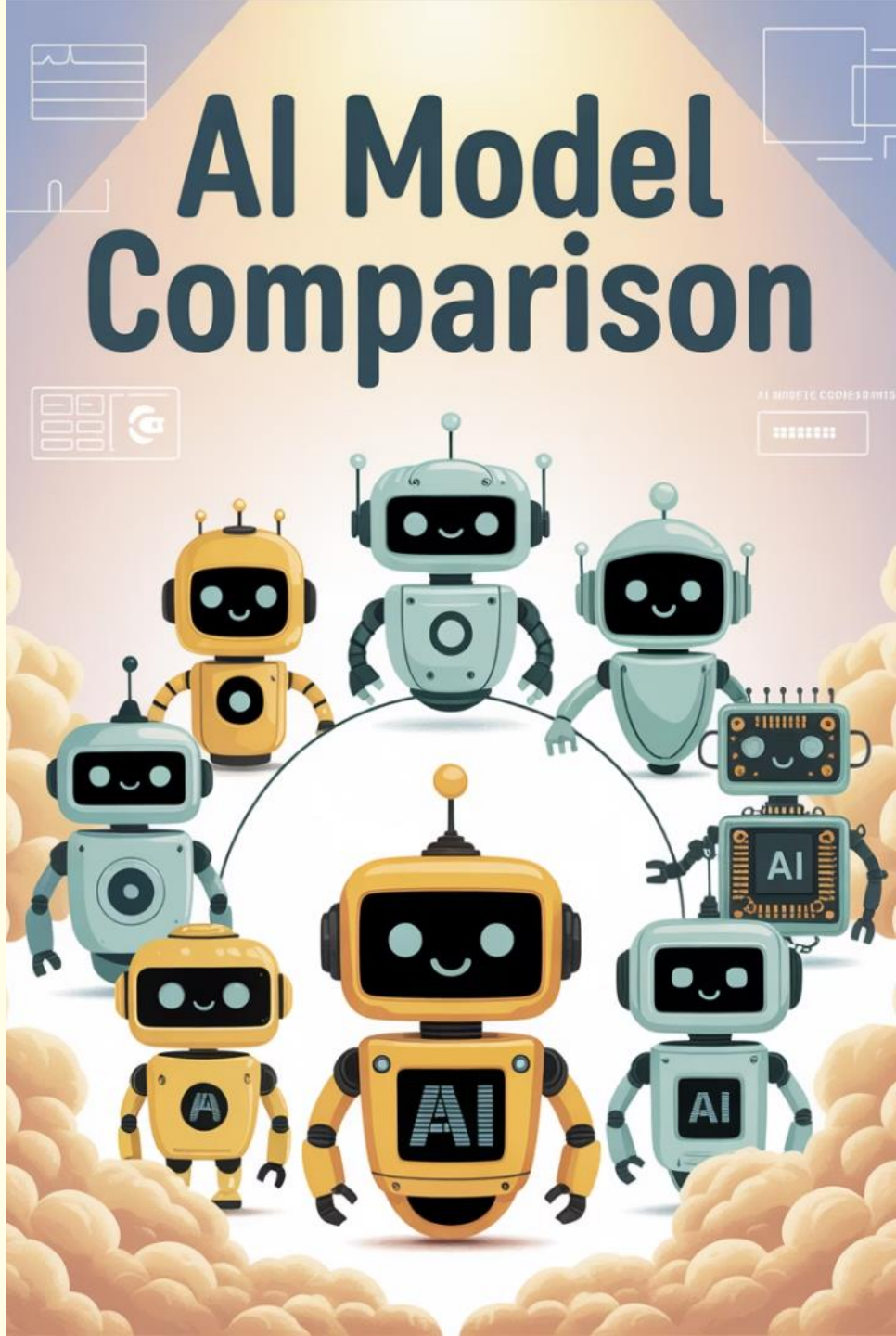
Future Direction

Moving toward autonomous agents with coding abilities



Comparing AI Models

Feature	ChatGPT	DeepSeek
Focus	General dialogue	Code & reasoning
Release	May 2024 (GPT-4o)	New platform architecture
Strengths	Conversational, creative	Technical, multilingual
Best Use	Writing, general tasks	Programming, mathematics



Can AI Improve Itself?

Fundamental Limits

Cannot transcend its own algorithm basis

Complexity Challenges

More complex code introduces more bugs

Mode Collapse Risk

Training on AI-generated content causes degradation

AI Ethics: Our Responsibility

Personal Accountability

We must take responsibility for AI use

Moral Reflection

AI forces us to examine our own values

Democratizing Power

Reducing information and capability monopolies

Shaping the Future

We get to determine what AI becomes

Will AI Take Away Human Joy?

Creative Pursuits

Still fulfilling regardless of AI capabilities

Process matters more than product

Personal meaning transcends efficiency

New Opportunities

AI handles routine tasks

Humans focus on meaningful work

More time for creativity and connection

Job Transformation in the AI Era

Historical Context

Nixon 1956: "Four-day workweek in our lifetime"

Technology promises often unrealized

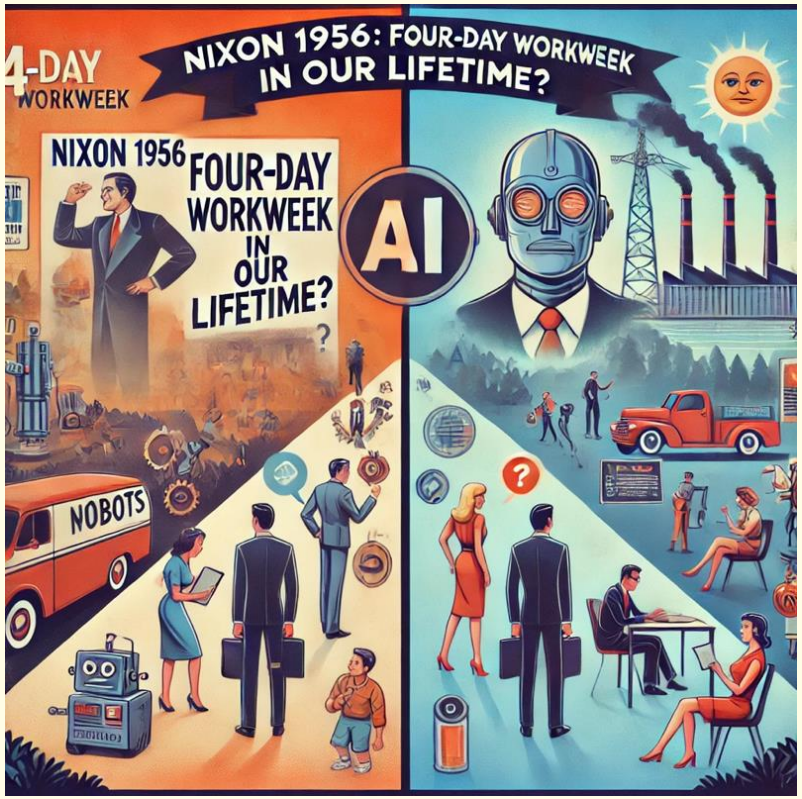
Work evolves rather than disappears

Future Reality

Jobs will require AI collaboration

Tool building becomes central skill

Managing AI agents creates new roles



Entropy of the universe increases (2nd Law)
When Order in the system increases, Disorder in the surroundings must go up!

What is Vibe Coding?

Natural Language Instructions

"Make it more playful" or "Corporate look and feel"

AI Interpretation

Translates vibes into technical specifications

Generated Solution

Code that matches desired aesthetic/feeling

Human Refinement

Developer tweaks for precise implementation

Tool Fatigue: Strategic Approach Approach

Focus on Core Tools

Master fundamentals rather than every variant

Evaluate Workflow Impact

Only adopt tools that significantly improve results

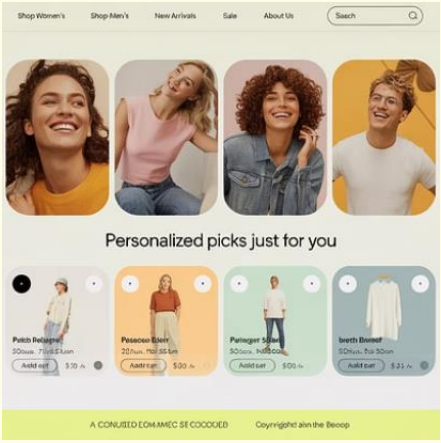
Transferable Skills

Prioritize concepts that apply across platforms

Sustainable Pace

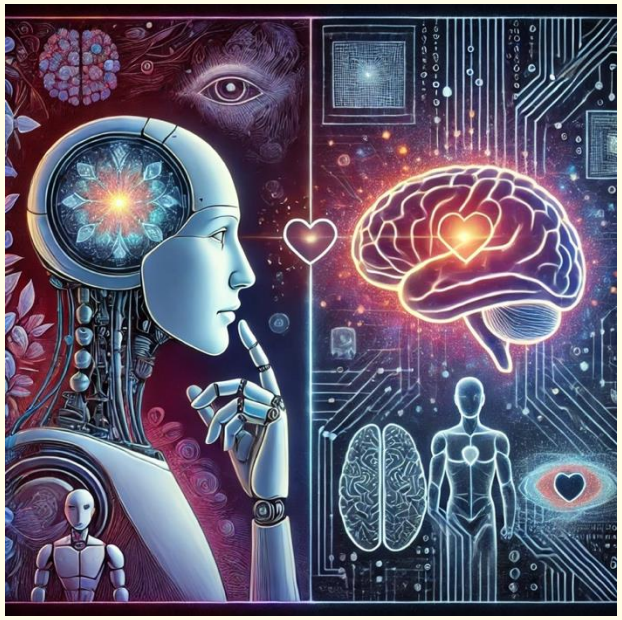
Quality understanding over over quantity exposure

How do I avoid AI? – Go for it, but…



- Don’t Use:
  - your phone, voice to text, spam filters; Fitbit, Health Apps
  - Credit Card/Bank Fraud Detection
  - Internet search autocomplete, or any web search
  - Facebook, Instagram, TikTok, YouTube, Netflix, Spotify etc
  - Google Maps, Waze, Apple Maps, Uber, Lyft
  - Amazon & other online stores
  - Ticketmaster, Airline Booking, dynamic pricing, Highways (Toll cameras)
  - Zoom (noise reduction)

Conscious AI: How Far Away?



Current State

Pattern matching without understanding,
Claims of sentience (e.g., LaMDA case at Google)
reflect anthropomorphism, not reality.

Consciousness Gap

We don't understand human consciousness itself

Fundamental Questions

May involve quantum mechanics
or unknown principles

Detecting AI-Generated Content

AI Weaknesses

Poor understanding of object permanence

Struggles with physical plausibility

Inconsistent details (fingers, reflections)

Detection Methods

Subtle heartbeat cadence missing in fakes

Digital watermarks embedded in generation

Pattern analysis revealing statistical anomalies

Let’s Pause…
Try slides 5&6 on Mentimeter

Why should I care if AI is or is not sentient?

- **Key Points:**
  - Sentience = conscious experience—pain, joy, self-awareness.
  - AI has no emotions, beliefs, or self-concept—it responds statistically based on training data.
  - Claims of sentience (e.g., LaMDA case at Google) reflect anthropomorphism, not reality.
- **Why It Matters:**
- **The illusion of sentience can impact:**
  - User trust
  - Emotional dependency

Will AI take my job away?

- **Richard Nixon 1956:**
  - "We are going to see it [the four-day workweek] in our lifetime."
  - Spoken during a campaign speech as Vice President.
  - Emblematic of mid 20th century optimism about automation
- **Entropy of the universe increases (2nd Law)**
  - When Order in the system increases, Disorder in the surroundings must go up!
- **Maybe, but we'll need two of you over there!**
  - Build that tool; manage that Agent

How can we detect fakes?

- **AI doesn't understand:**
  - Permanence of objects
  - Physics
- **We can use AI to detect:**
  - absence of subtle heartbeat cadence
  - so-called AI "Watermarks"