DISOBEDIENCE LAB

CREATIVITY FOR SUSTAINABILITY

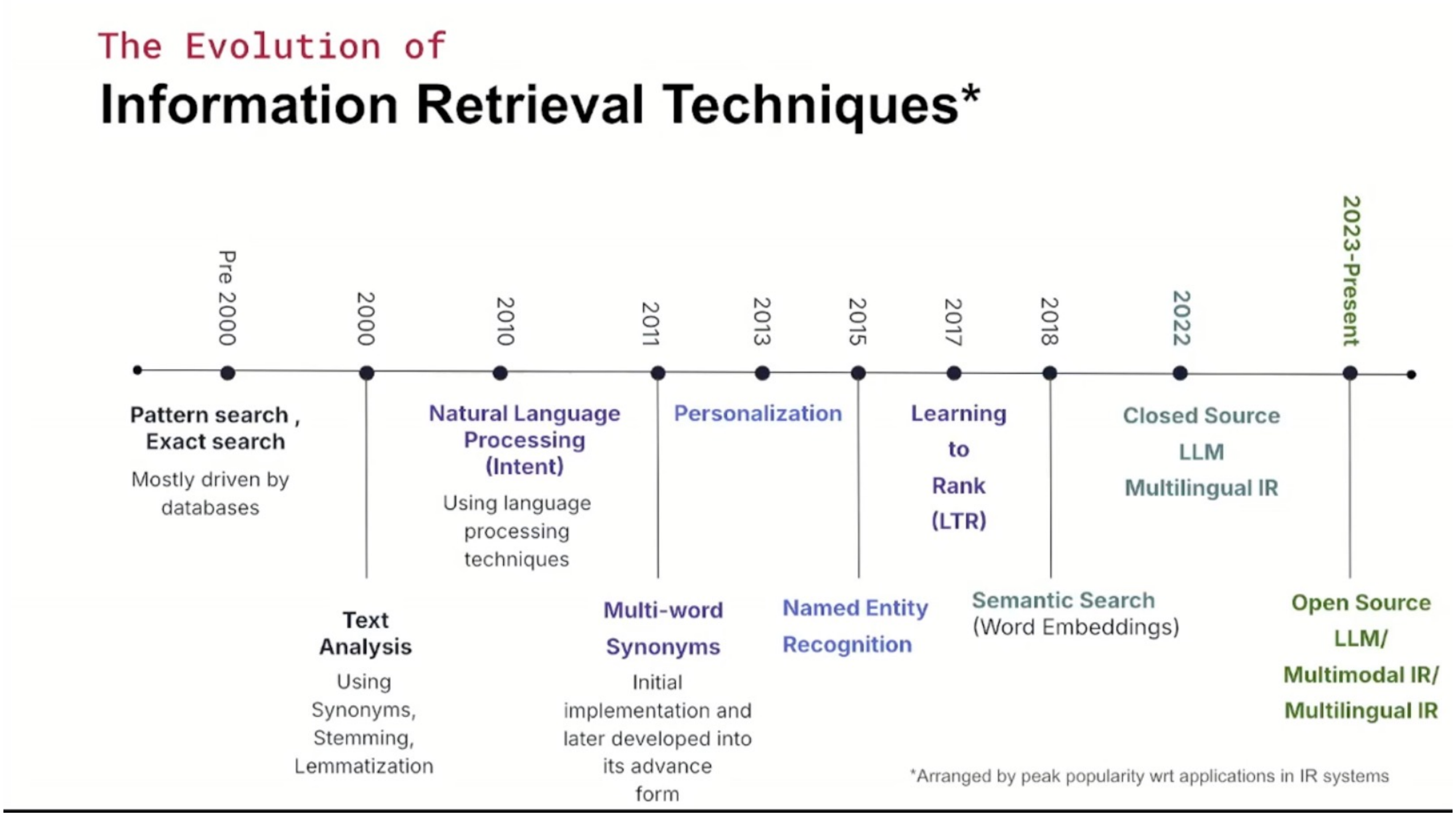
GEN AI and Sustainability

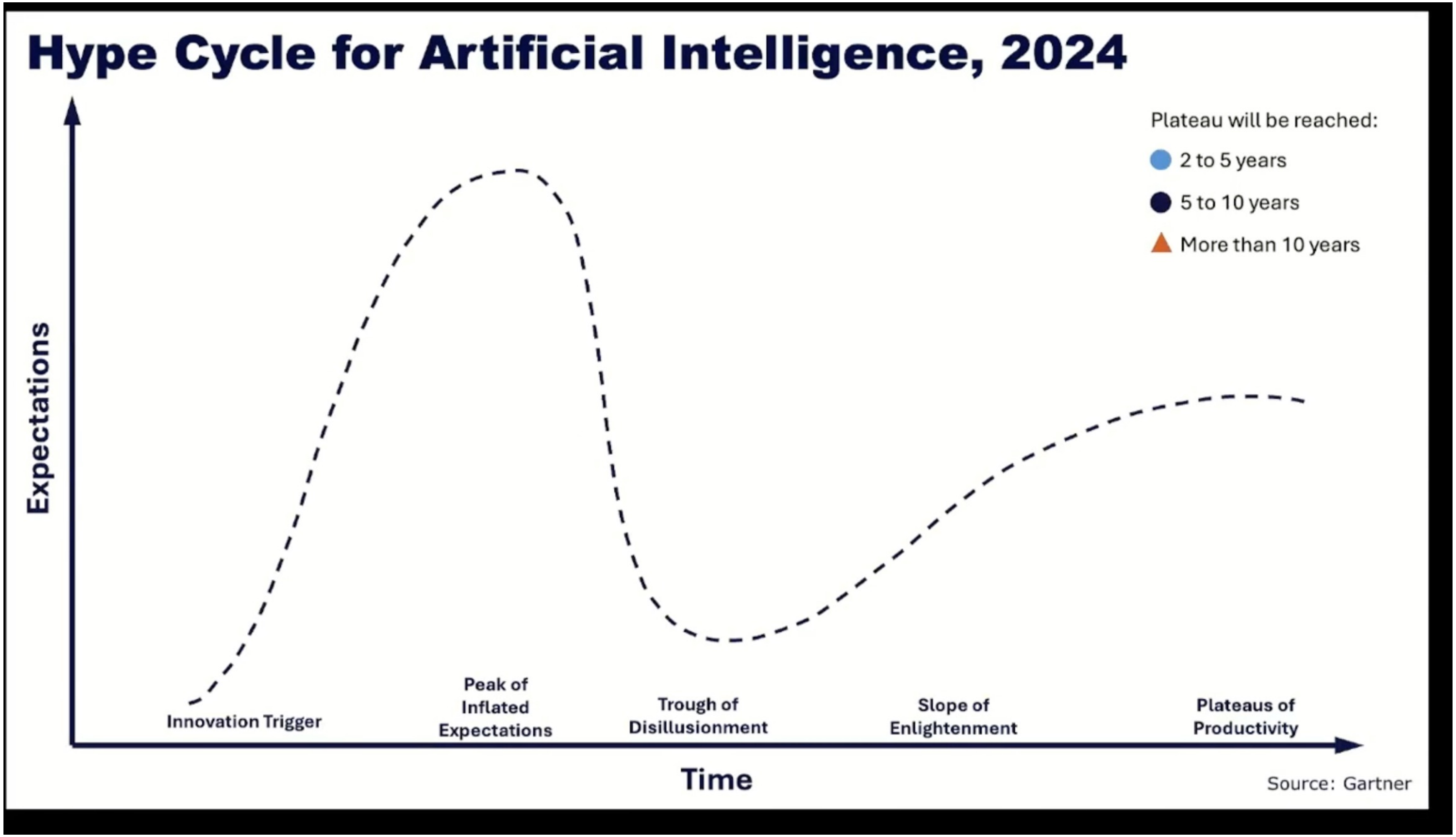
Learning Objectives

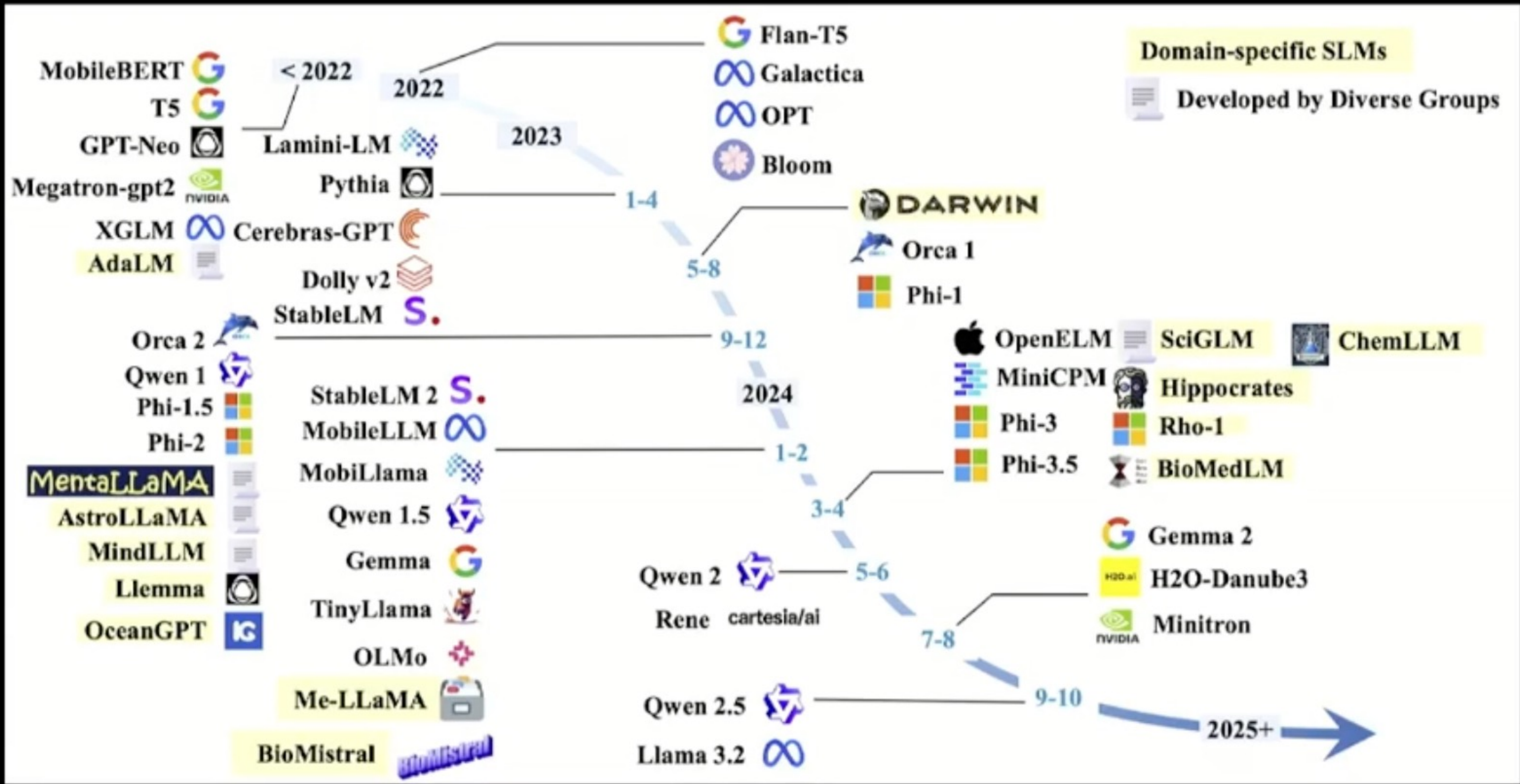
- ✓ Define AI Fundamentals and Generative Models.
- ✓ Deconstruct Large Language Models (LLMS)
- ✓ Analyze the behavior of LLMs and their understanding, completion, and prediction of text.
- ✓ Investigate the phenomenon of hallucinations in LLMs and its underlying causes.
- ✓ Identify the diverse applications of ML, DL, and Generative AI in business, retail, health, and technology sectors.
- ✓ Gain practical experience by working with ChatGPT for text generation and DALL-E for image generation.

A brief history of Gen AI

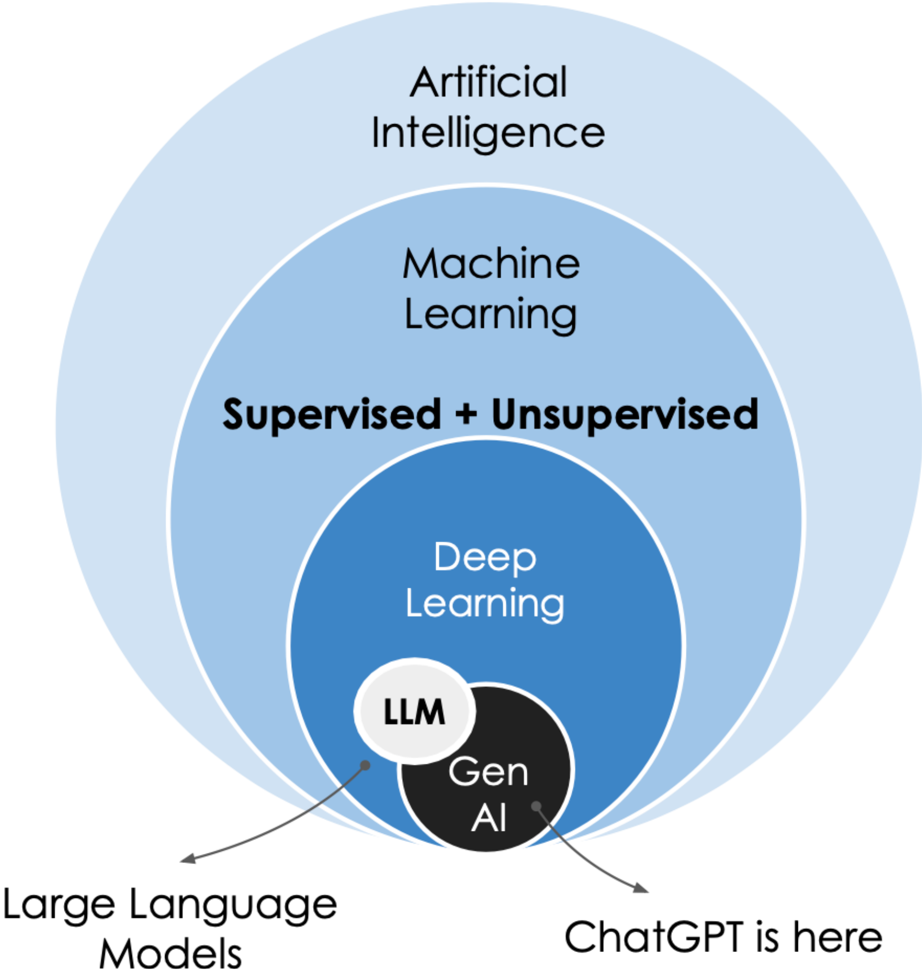
- **2014** Generated realistic images - GANs - Ian Goodfellow
- **2017** Breakthrough in NLP - Transformer Models
- **2018** Advent of large language models - BERT, GPT
- **2019** Advanced GANs - style GANs
- **2020** OpenAI released GPT-3 - this was state-of-the-art
- **2021** Vision Transformers - CLIP, DALL-E
- **2022** OpenAI released ChatGPT (GPT-3.5), DALL-E 2
- **2023** GPT-4, Stable Diffusion Model - market is flooded with GenAI apps


The Evolution of
Information Retrieval Techniques*
Pre 2000
Pattern search , Exact search
Mostly driven by databases
2000
Text Analysis
Using Synonyms, Stemming, Lemmatization
2010
Natural Language Processing (Intent)
Using language processing techniques
2011
Multi-word Synonyms
Initial implementation and later developed into its advance form
2013
Personalization
2015
Named Entity Recognition
2017
Learning to Rank (LTR)
2018
Semantic Search
(Word Embeddings)
2022
Closed Source LLM
Multilingual IR
2023-Present
Open Source LLM/
Multimodal IR/
Multilingual IR
*Arranged by peak popularity wrt applications in IR systems


Hype Cycle for Artificial Intelligence, 2024
Plateau will be reached:
2 to 5 years
5 to 10 years
More than 10 years
Expectations
Innovation Trigger
Peak of Inflated Expectations
Trough of Disillusionment
Slope of Enlightenment
Plateaus of Productivity
Time
Source: Gartner



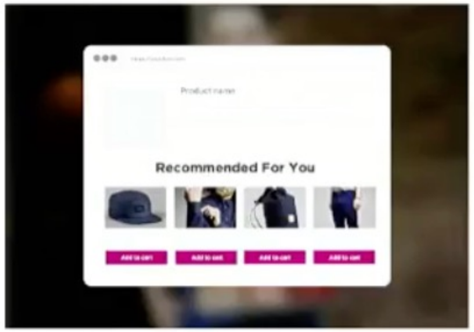
Timeline of various Small Language Models

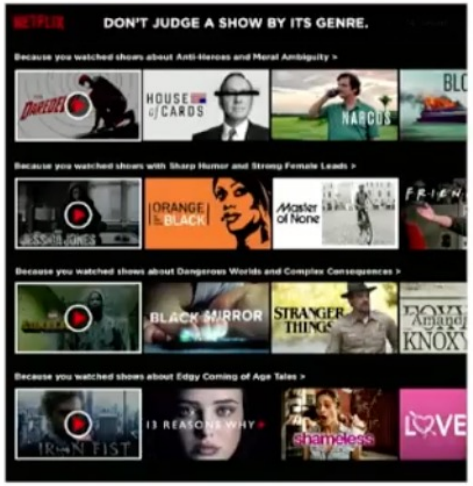


Getting the Definitions Right

Structured vs Unstructured Data use-cases

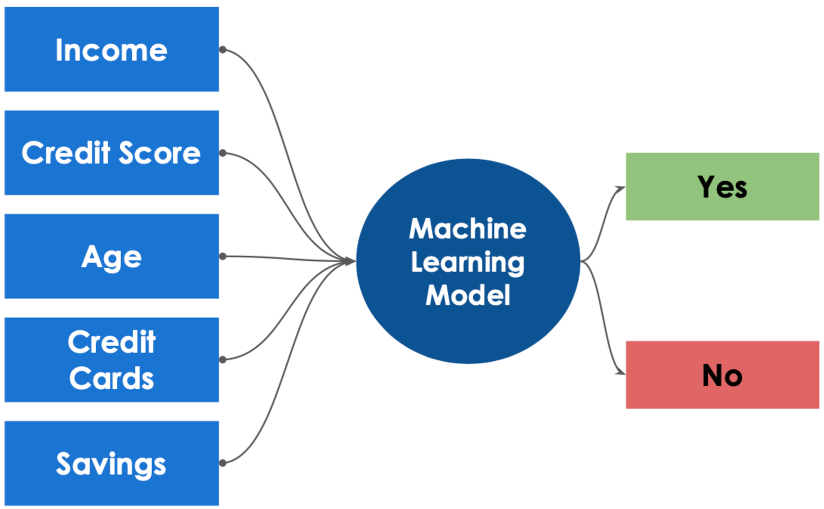
Name	FName	City	Age	Salary
Smith	John	3	35	$280
Doe	Jane	1	28	$325
Brown	Scott	3	41	$265
Howard	Shemp	4	48	$359
Taylor	Tom	2	22	$250

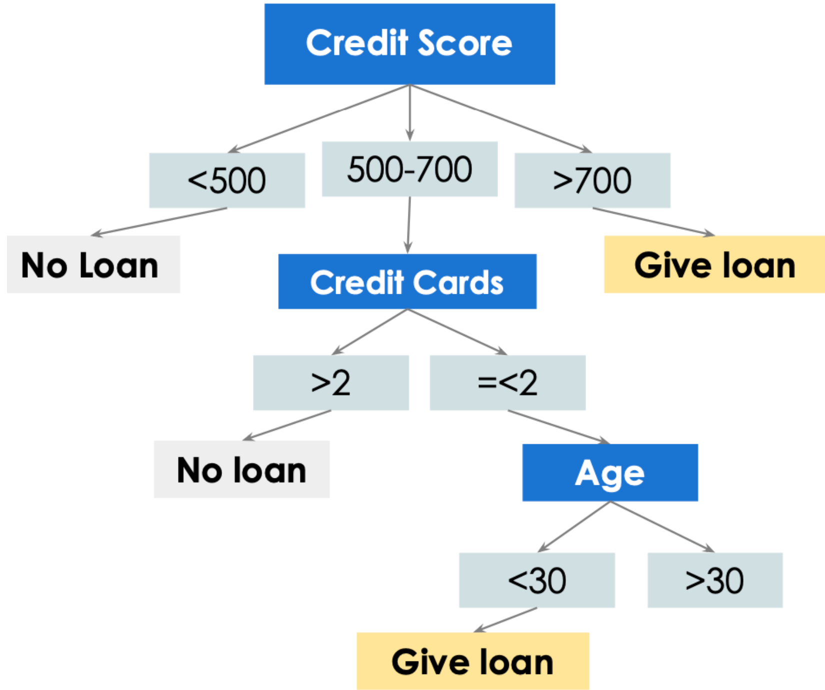




Supervised Learning

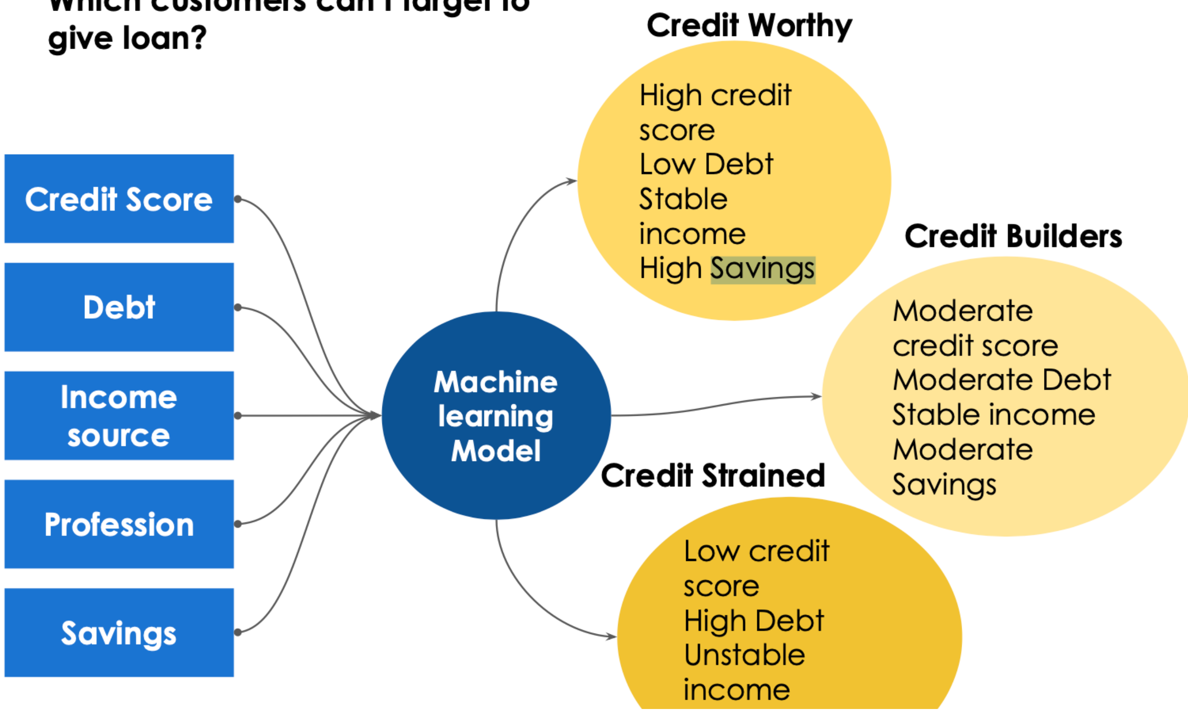
Should I give loan to this customer?

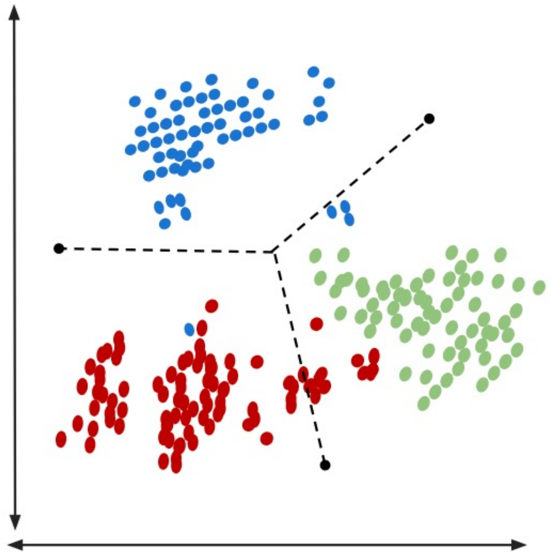




Unsupervised Learning

Which customers can I target to give loan?





Discriminative Ai Vs Gen AI

Train

Model

Learns from relationship between **labels** and data

Could this be a dog?

Classify

Discriminative AI

+

Other images on the internet

Train

Model

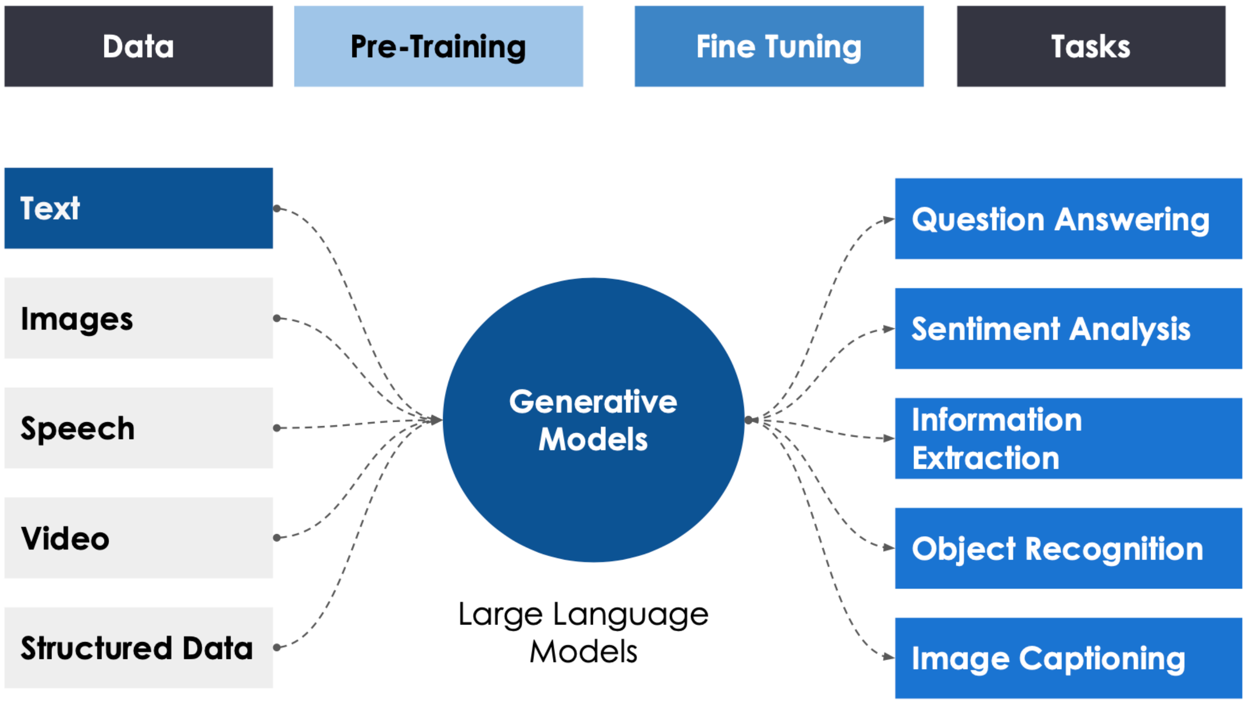
Learns from patterns in large amounts of unstructured data

Create an image of dog, having fun in a party, wearing a black tuxedo with wine in one hand

Generate

Generative AI

Generative AI Models



Large Language Models - LLM

Large, because 2 things:

1. Trained on **large amounts of data**

2. billions of **trainable parameters**

Language, because it deals with text data (takes input in text and generates output in text).

Model, because it predicts the next word/sentence/token.

So LLMs are language models consisting of a neural network with billions of parameters, trained on large quantities of unlabeled text using self-supervised learning.

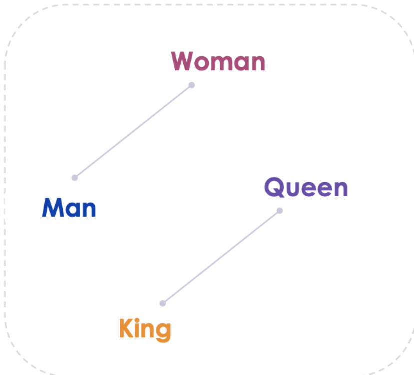
How does the model understand text?

Is there a numeric way to represent association between text or words?

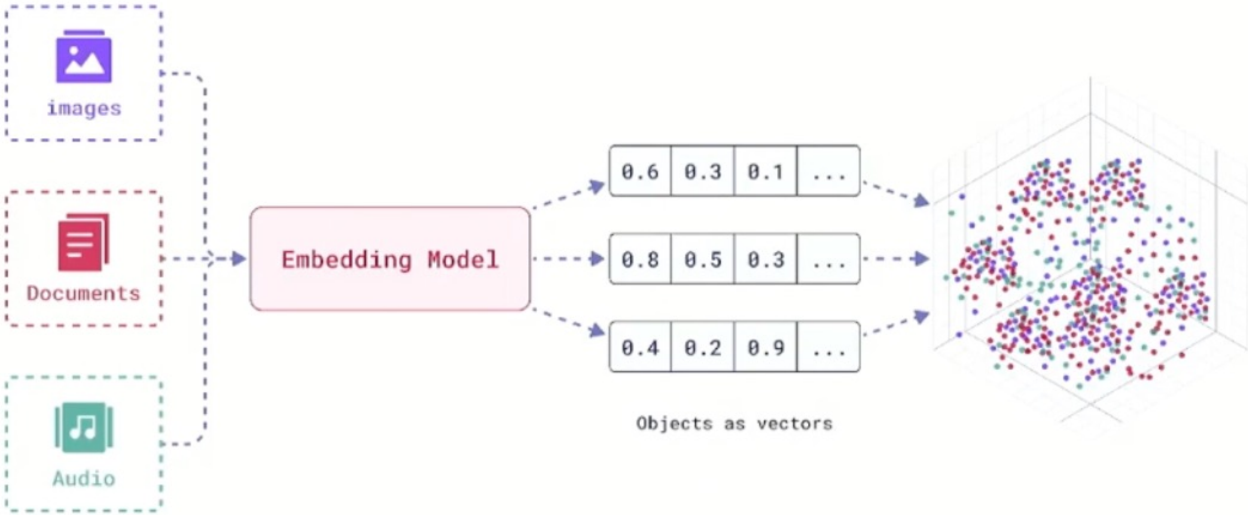
Word embeddings = semantic + syntactic relations in a vector space

similarity + rules = meaning

Let's try to understand vector space	Living Being	Human	Gender	Royalty
Man	0.8	0.8	0.8	- 0.7
Woman	0.9	0.9	- 0.9	- 0.8
King	0.8	0.7	0.7	0.7
Queen	0.7	0.8	- 0.8	0.8



The magic of Embedding !!



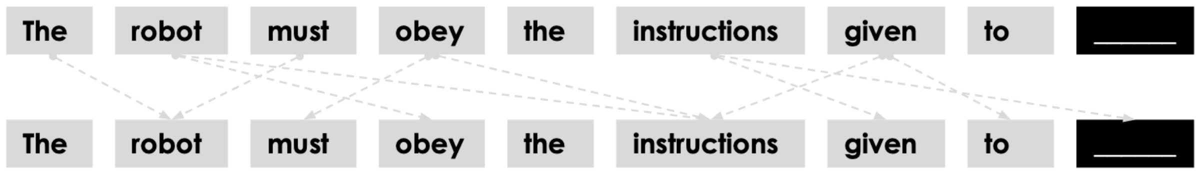
An object is known by company it keeps !

A high dimensional array of numbers that captures the context & semantic meaning

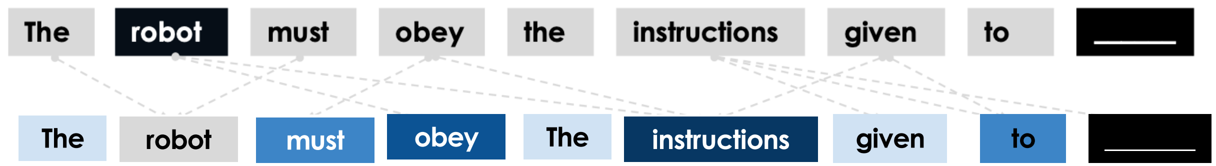
How does it predict the next word?

- **Step-1**: **Word Embeddings** - Break the sentence into words and convert them to embeddings

- **Step-2**: **Find connections** - Understand which word is related to which word more



- **Step-3**: **Giving importance/attention**: Each word is assigned a score based on how important it is to other words in the sentence



How does it predict the next word?

- **Step-4**: **Assigning Weights** - This is for "robot" - but step-3 and 4 will be repeated for all words

The	**robot**	**must**	**obey**	**the**	**instructions**	**given**	**to**	______
0.2	**-**	**0.6**	**0.7**	**0.2**	**0.8**	**0.3**	**0.7**	**Weights**

- **Step-5**: **Find Relevance** - To complete the sentence, which are the words to consider

The	**robot**	**must**	**obey**	**the**	**instructions**	**given**	**to**	**it**
0.2	**0.9**	**0.2**	**0.7**	**0.2**	**0.8**	**0.3**	**0.2**	

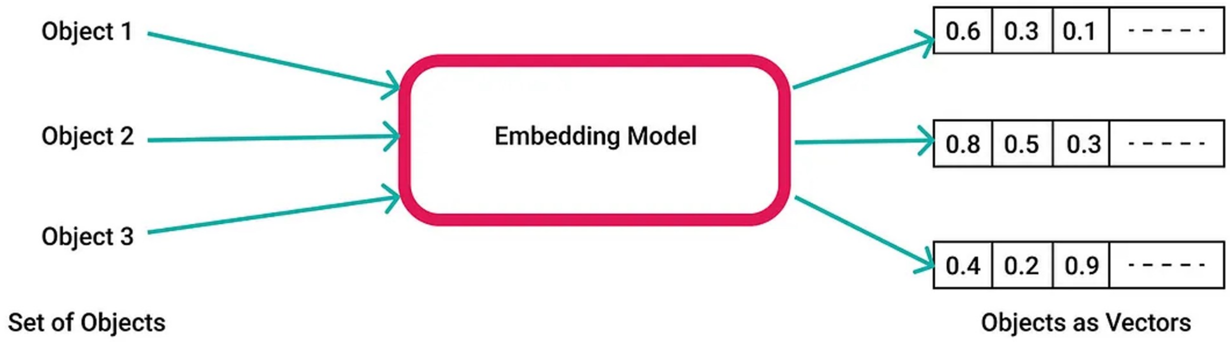
it	**0.9**
its	**0.8**
itself	**0.85**
robot	**0.5**

- What is a vector?

  A numerical representation of text/image/video, ..

- What is a embedding?

  - Mapping of data into a format (vector).
  - Goal is to capture semantics and context.
  - Processed by a ML model.



Embeddings

Take source objects (text, images, sound, movies) and create Vector Embeddings. This allows for Similarity Search on the Database finding Semantically comparable objects.

“To create a security token that can be used to log into a database, select Token Management from the User Management menu. Then, choose an appropriate role for the user, and click the Generate Token button. Copy the token details to a safe place, as the secret that is shown can never be reproduced in the Astra console for security reasons.”

Raw Text

[-0.029334254562854767, 0.06338247656822205, 0.03711941838264465, 0.06770425289869308, 0.030722564086318016, -0.03855780512094498,
0.05715630576014519, 0.011225797235965729, 0.03209076076745987, 0.018565965816378593, 0.005725057329982519, 0.003278332995250821,
0.019661232829093933, 0.008483093231916428, 0.01011530589312315, -0.06865981221199036, -0.02742725796997547, 0.004272086545825005,
0.006464742589741945, 0.033381473273038864, -0.06456394493579865, -0.00168662762735039, -0.02538292482495308, -0.013529211282730103,
0.006745325401425362, -0.09090574085712433, -0.004533097147941589, -0.011557350866496563, -0.017933078110218048, -0.013565625995397568,
0.037340205162763596, -0.013345368206501007, -0.046318963170051575, 0.018221553415060043, -0.030514687299728394, 0.06087908893823624, -
0.015947293490171432, -0.004384475760161877, 0.011510275304317474, 0.03181201219558716, 0.0004961538943462074, -0.009607694111764431, -
0.0026698557194322348, -0.02120211534202099, -0.02445143461227417, 0.018808122724294662, -0.04526928439736366, -0.03507508337497711,
0.011936023831367493, -0.04246317967772484, -0.04538712650537491, 0.0030571885872632265, -0.02645616978406906, 0.004339783918112516, -
0.004989228677004576, -0.0019529943820089102, -0.015389597043395042, -0.008066490292549133, -0.04361109808087349, 0.018591511994600296, -
0.008249715901911259, 0.031450431793928146, 0.008753931149840355, -0.06284607946872711, -0.02690013311803341, 0.061753395944833755,
0.0314679853618145, 0.005365008022636175, -0.034285105764865875, -0.06475269049406052, 0.06229010224342346, -0.016091199591755867, -
0.038237784057855606, -0.01697474904358387, 0.0023320959880948067, -0.02873411774635315, -0.07216104120016098, 0.04663623124361038,
0.023897146806120872, -0.02821142040193081, -0.03714695945382118, -0.055613692849874496, -0.0028377221897244453, -0.06574894487857819, -
0.06103818118572235, 0.06294400244951248, 0.0034130788408219814, 0.07920042425394058, 0.007338271476328373, 0.06506536900997162, -
0.02522268146276474, 0.027450041845440865, -0.01720043271780014, 0.046272337436676025, -0.05018896237015724, 0.015779439359903336, -
0.026586400344967842, -0.01974601112306118, -0.00036689057014882565, -0.016816521063447, -0.025464840233325958, 0.0007100807852111757,
0.04524853080511093, 0.0010508883278816938, -0.005472411867231131, 0.011604292318224907, -0.0427076481282711, -0.02004644088447094, -
0.06824997067451477, -0.08084388822317123, -0.08167271316051483, 0.038480401039123535, -0.04149484634399414, 0.0621405728161335, 0.01636849343776703, -
0.02775057591497898, 0.02410232089459896, 0.021344885230064392, 0.056428126990795135, 0.02979239635169506, -0.05207456275820732,
0.004299748223274946, 0.03417612612247467, 0.034210722893476486, 0.00010842653136933222, 0.011242502368986607, 0.03719356656074524, -
0.004098605364561081, 0.013202376663684845, 0.021659305319190025, 0.03850701451301575, -0.03979567810893059, 0.024909289553761482,
0.0036120524164289236, 0.030269717797636986, 0.03532775491476059, 0.04048445075750351, -0.021236592903733253, 0.05895552039146423,
0.04913758486509323, -0.047305576503276825, 0.05272332951426506, 0.012154217809438705, -0.02513653226196766, -0.0105582932010293, -
0.049685653299093246, 0.032950107008218765, -0.007436738815158606, -0.07494320720434189, -0.04471106082201004, 0.03816404938697815, -
0.029877835884690285, -0.020543526858091354, 0.02532779611647129, 0.011234065517783165, 0.07374250143766403, 0.04288359731435776,
0.03435317426919937, -0.02951200306415558, -0.09385887533426285, -0.005317367613315582, 0.01705515943467617, -0.00934696663171053,
0.01293235830962658, 0.02108096517622471, 0.030062183737754822, 0.004270109347999096, -0.005795920733362436, 0.006119553931057453, -
0.009726069867610931, -0.0016054088482633233, -0.12823154032230377, 0.005963715258985758, -0.01607099547982216,…]

Embedding

Embeddings

Take source objects (text, images, sound, movies) and create Vector Embeddings.
This allows for Similarity Search on the Database finding Semantically comparable objects.

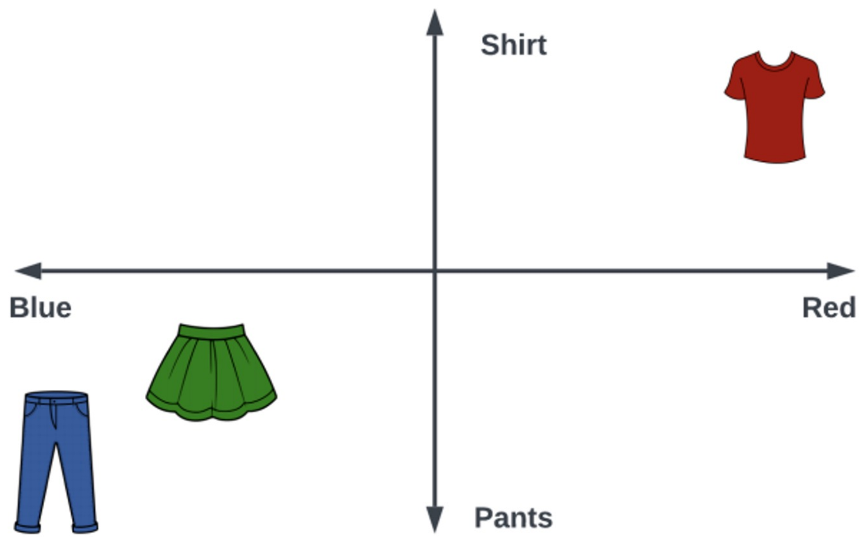
[-0.029334254562854767, 0.06338247656822205, 0.03711941838264465, 0.06770425289869308, 0.030722564086318016, -0.03855780512094498,
0.05715630576014519, 0.011225797235965729, 0.03209076076745987, 0.018565965816378593, 0.005725057329982519, 0.003278332995250821,
0.019661232829093933, 0.008483093231916428, 0.01011530589312315, -0.06865981221199036, -0.02742725796997547, 0.004272086545825005,
0.006464742589741945, 0.033381473273038864, -0.06456394493579865, -0.00168662762735039, -0.02538292482495308, -0.013529211282730103,
0.006745325401425362, -0.09090574085712433, -0.004533097147941589, -0.011557350866496563, -0.017933078110218048, -0.013565625995397568,
0.037340205162763596, -0.013345368206501007, -0.046318963170051575, 0.018221553415060043, -0.030514687299728394, 0.06087908893823624, -
0.015947293490171432, -0.004384475760161877, 0.011510275304317474, 0.03181201219558716, 0.0004961538943462074, -0.00960769411764431, -
0.0026698557194322348, -0.02120211534202099, -0.02445143461227417, 0.018808122724294662, -0.04526928439736366, -0.03507508337497711,
0.011936023831367493, -0.04246317967772484, -0.04538712650537491, 0.0030571885872632265, -0.02645616978406906, 0.004339783918112516, -
0.004989228677004576, -0.0019529943820089102, -0.015389597043395042, -0.008066490292549133, -0.04361109808087349, 0.018591511994600296, -
0.008249715901911259, 0.031450431793928146, 0.008753931149840355, -0.06284607946872711, -0.02690013311803341, 0.06175339594483755,
0.0314679853618145, 0.005365008022636175, -0.034285105764865875, -0.06475269049406052, 0.06229010224342346, -0.016091199591755867, -
0.03823778405785606, -0.0169747490435838 7, 0.0023320959880948067, -0.02873411774635315, -0.07216104120016098, 0.04663623124361038,
0.02389714680612087 2, -0.0282114204019308 1, -0.03714695945382118, -0.055613692849874496, -0.0028377221897244453, -0.06574894487857819, -
0.06103818118572235, 0.06294400244951248, 0.0034130788408219814, 0.07920042425394058, 0.007338271476328373, 0.06506536900997162, -
0.02522268146276474, 0.027450041845440865, -0.017200432717800 14, 0.046272337436676025, -0.05018896237015724, 0.015779439359903336, -
0.026586400344967842, -0.019746011123061 18, -0.00036689057014882565, -0.016816521063447, -0.025464840233325958, 0.0007100807852111757,
0.0452485308051109 3, 0.0010508883278816938, -0.005472411867231131, 0.011604292318224907, -0.0427076481282711, -0.02004644088447094, -
0.06824997067451477, -0.08084388822317123, -0.08167271316051483, 0.038480401039123535, -0.04149484634399414, 0.0621405728161335, 0.0163684934377670 3, -
0.02775057591497898, 0.02410232089459896, 0.021344885230064392, 0.05642812699079513 5, 0.02979239635169506, -0.05207456275820732,
0.004299748223274946, 0.03417612612247467, 0.034210722893476486, 0.00010842653136933222, 0.011242502368986607, 0.03719356656074524, -
0.004098605364561081, 0.013202376663684845, 0.021659305319190025, 0.03850701451301575, -0.03979567810893059, 0.02490928955376148 2,
0.003612052416428923 6, 0.030269717797636986, 0.03532775491476059, 0.04048445075750351, -0.021236592903733253, 0.05895552039146423,
0.04913758486509323, -0.047305576503276825, 0.05272332951426506, 0.012154217809438705, -0.025136532261967 66, -0.0105582932010293, -
0.049685653299093246, 0.032950107008218765, -0.007436738815158606, -0.07494320720434189, -0.04471106082201004, 0.0381640493869781 5, -
0.029877835884690285, -0.020543526858091354, 0.02532779611647129, 0.011234065517783165, 0.07374250143766403, 0.04288359731435776,
0.03435317426919937, -0.029512003064155 58, -0.09385887533426285, -0.005317367613315582, 0.01705515943467617, -0.00934696663171053,
0.01293235830962658, 0.02108096517622471, 0.030062183737754822, 0.004270109347999096, -0.005795920733362436, 0.006119553931057453, -
0.009726069867610931, -0.0016054088482633233, -0.12823154032230377, 0.005963715258985758, -0.01607099547982216,…]

Image or Audio

Embedding

What is Vector Search

- Vector search finds objects that have similar meaning
- Vector search understands MEANING
- Vectors created from EXISTING data through EMBEDDINGS



AI Governance

Ethan Mollick • Following
Associate Professor at The Wharton School. Author of Co-...
1h •

After talking to many firms, I have come to believe a key factor in successful AI adoption is whether the executive team actually experiments with AI to try to get work done themselves. Those who do tend to feel urgency and push for transformation.

Most c-level folks (the vast majority of executives I talk to) still haven't even tried LLMs. Changing that should be the highest priority of anyone who wants their company to succeed at AI.

12:19 5G

Ethan Mollick • Following
Associate Professor at The Wharton School. Author of Co-Intelligence
1h •

Big issue in organizations: They have put together elaborate rules for AI use focused on negative use cases (and some of the focus of these rules, like privacy, have actually become much less important over the last year as AI companies have updated their policies).

As a result, employees are too scared to talk about how they use AI, or to use corporate LLMs; they are afraid of punishment for ill-defined mistakes. But they keep using AI. They just become secret cyborgs, using their own AI & not sharing knowledge.

AI Governance

Integrating GenAI into Your Org

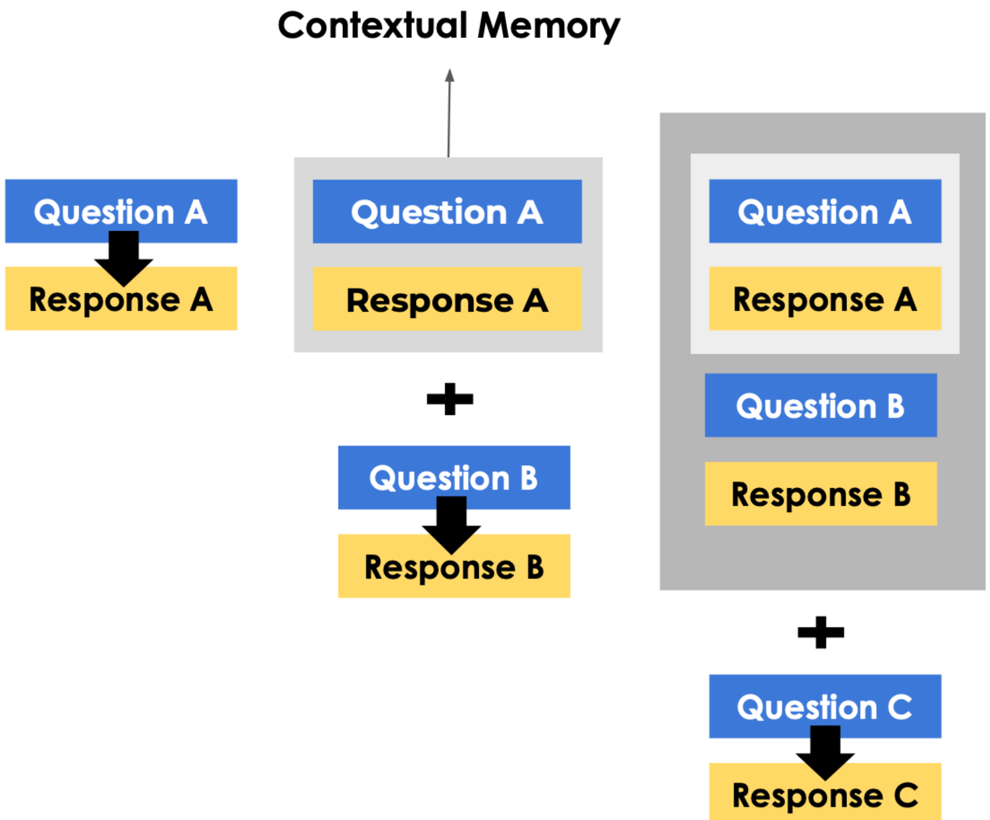
1. Organizational Assessment
2. Executive Education and Alignment
3. AI Council Formation
4. Policies and Guidelines
5. Training
6. GenAI Integration
7. Measuring the Value

Self Awareness for AI

5 simple rules

1. Good AI newsletters
2. Try stuff
3. Integrate on your daily routine
4. Improve your prompting skills
5. Stay alert

How is it able to remember conversations?

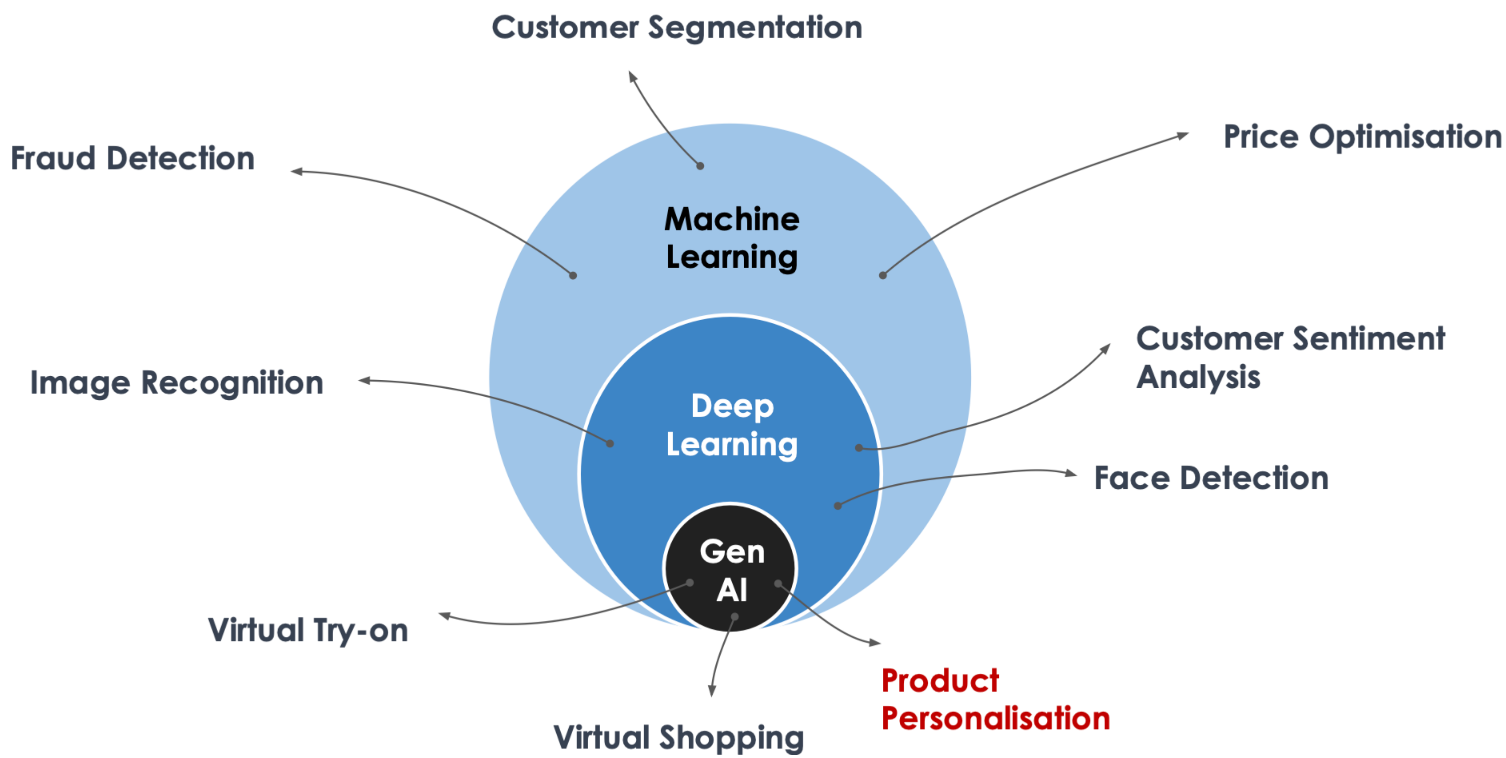


Why do Language Models Hallucinate?

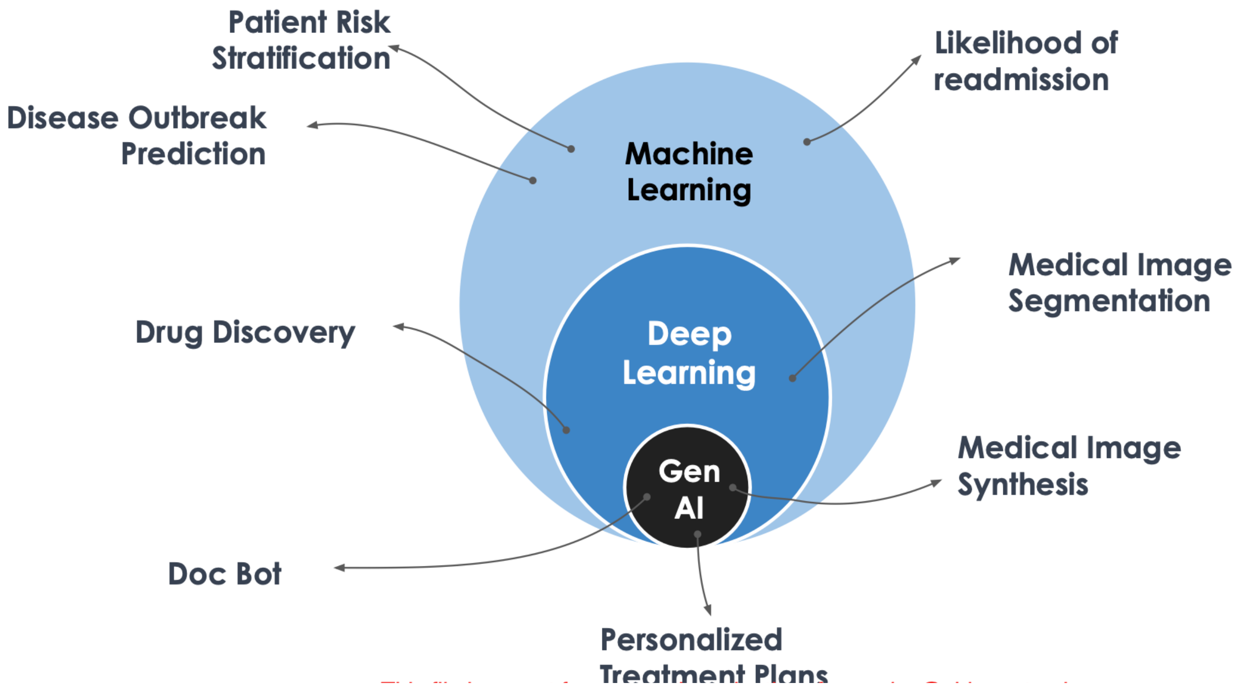
Why does this happen?

- Did not understand the context or intent behind the question / prompt
- Information needed to give the relevant answer was absent in the training data
- These are probabilistic models - leading to inconsistency/randomness in its outputs

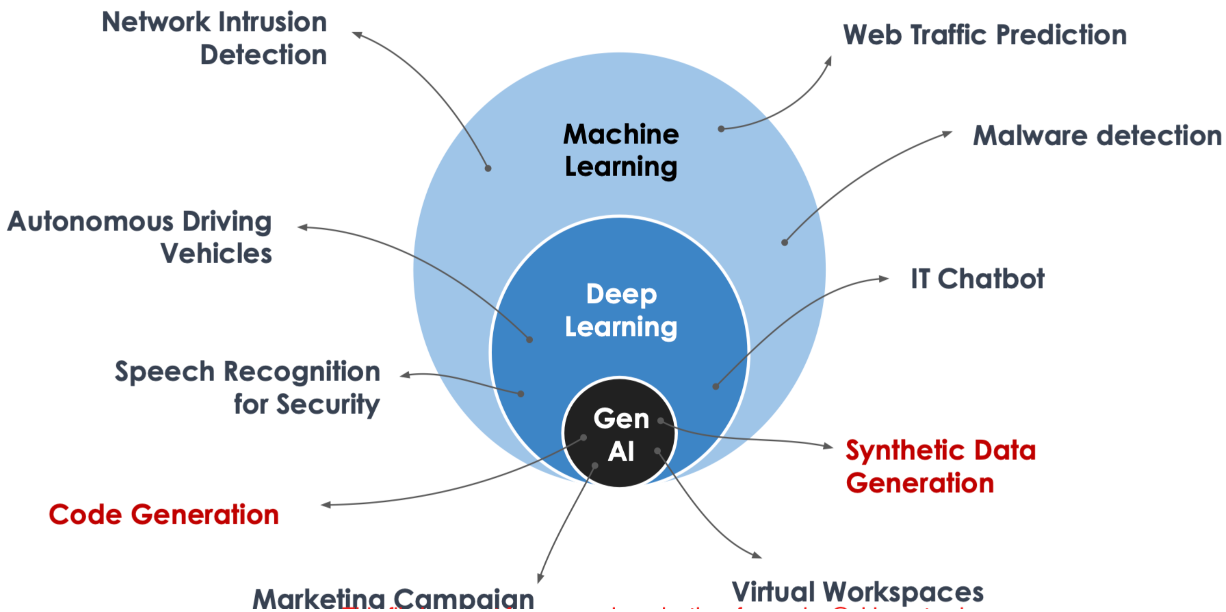
Business problems solved by Gen AI - Retail



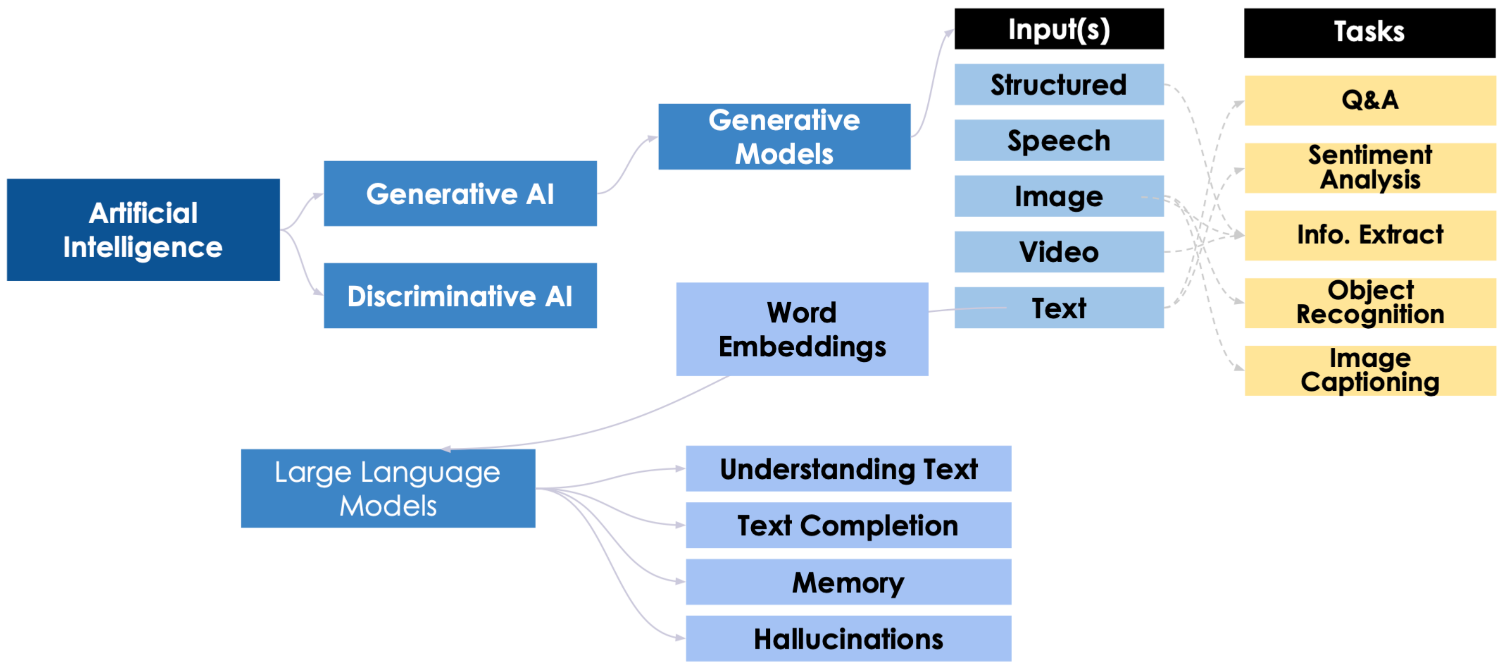
Business problems solved by Gen AI - Health



Business problems solved by Gen AI - Tech

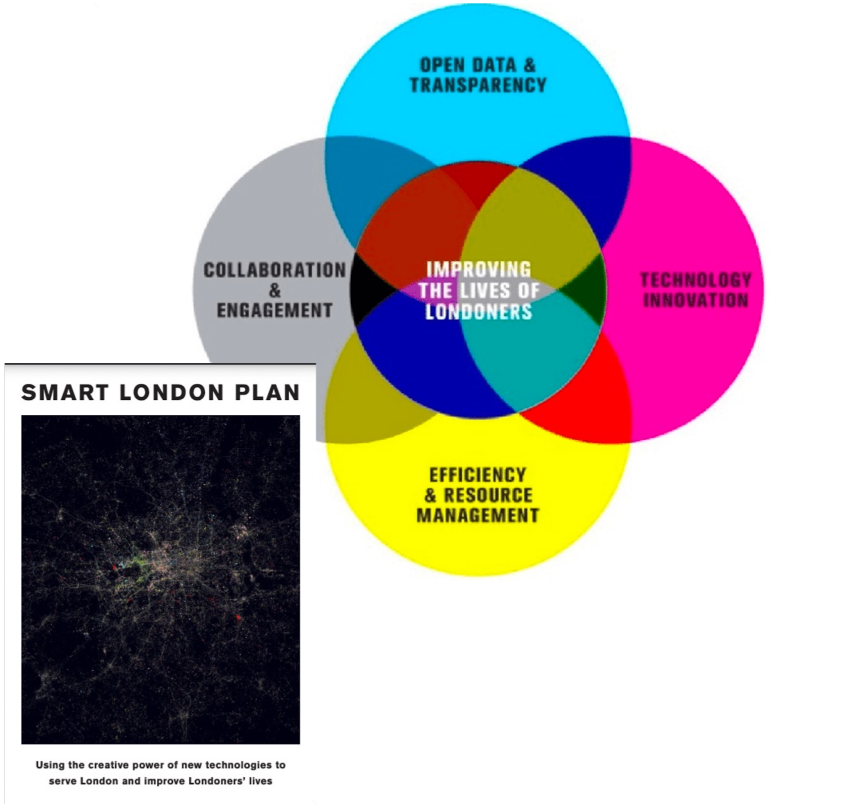


MindMap


Artificial Intelligence
Generative AI
Discriminative AI
Generative Models
Input(s)
Structured
Speech
Image
Video
Text
Tasks
Q&A
Sentiment Analysis
Info. Extract
Object Recognition
Image Captioning
Word Embeddings
Large Language Models
Understanding Text
Text Completion
Memory
Hallucinations

London Smart Plan

Uses generative AI to optimize energy usage, improve transportation systems, and manage waste.



Generative AI

San Francisco Urban Festival

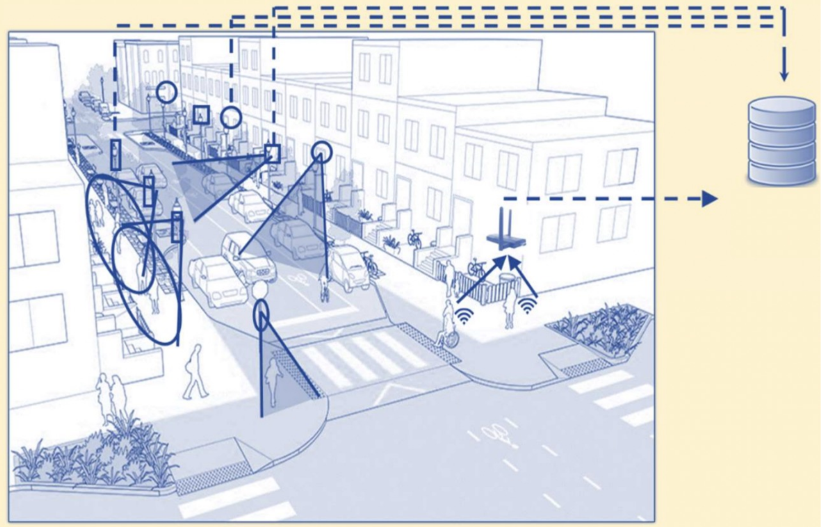
Uses generative AI to develop and test sustainable urban design prototypes.

Generative AI

Cities

Portland Urban Data Lake

Uses generative AI to analyze and manage data related to energy consumption, transportation, and waste management.



Urban data storage system
Public-private partnership
Address public's concerns

Generative AI

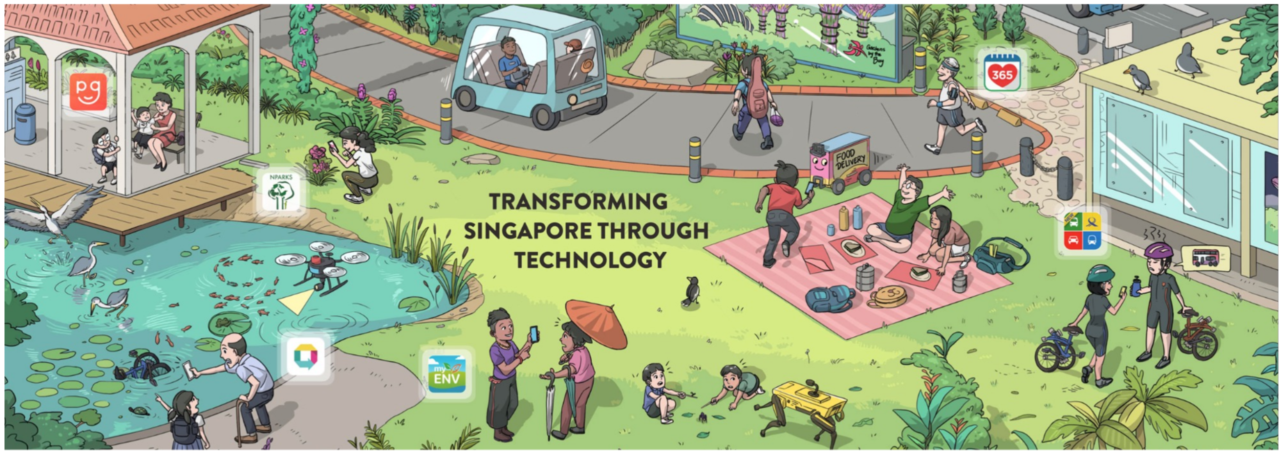
Cities

Smart Nation Singapure

Use of generative AI to optimize energy consumption, improve transportation systems, and manage waste.

Smart Nation
SINGAPORE

https://www.smartnation.gov.sg/



Generative AI

Cities

Los Angeles MDS

Uses generative AI to optimize the management of dockless e-scooters and bicycles, which have become increasingly popular modes of transportation in many cities.

The platform enables the city to collect real-time data about the location, usage, and charging status of these vehicles, and use this data to manage the fleet more efficiently and effectively.

MOBILITY DATA SPECIFICATION

OPEN MOBILITY FOUNDATION

https://cities-today.com/how-los-angeles-took-control-of-its-mobility-data/

Generative AI can help cities optimize their transportation systems, reduce congestion, and improve mobility for residents and visitors.

15 min break