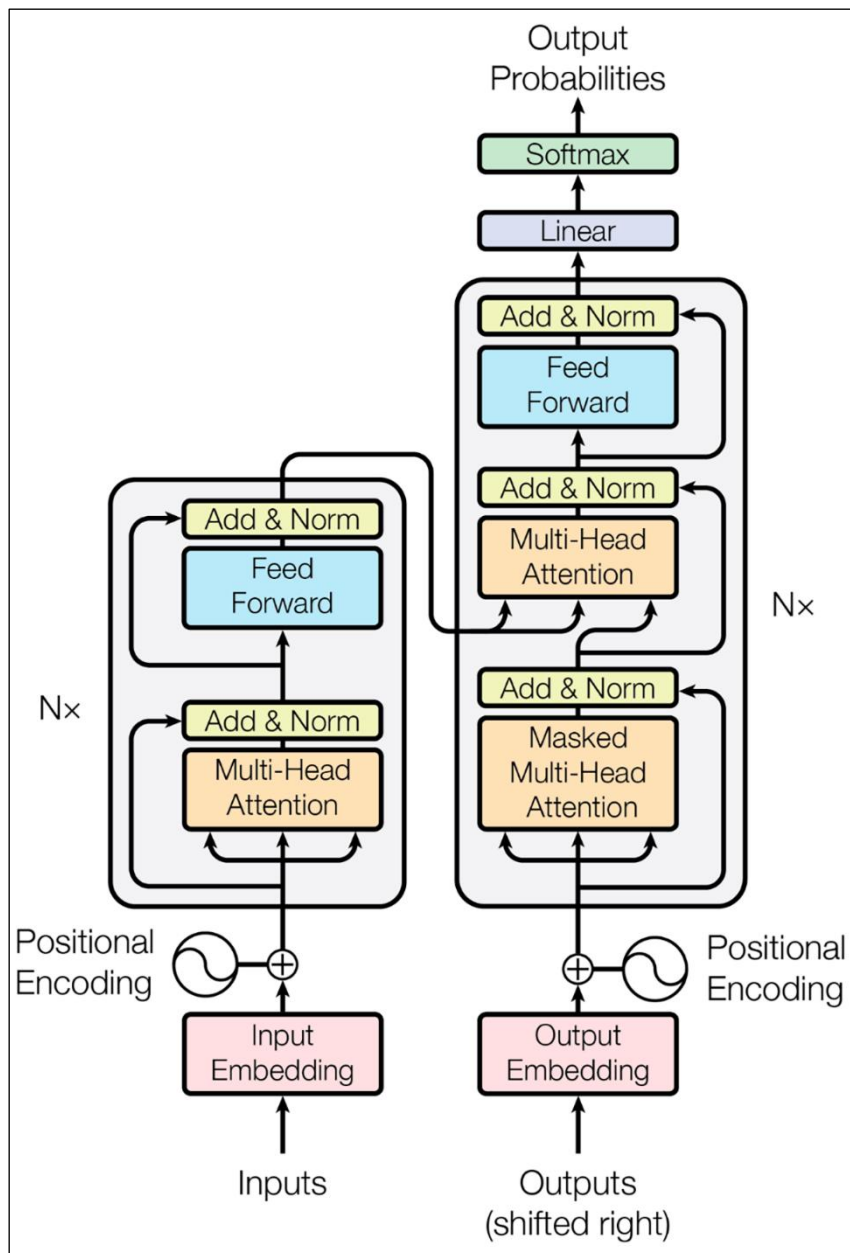


Chapter 1: Introduction to Large Language Models



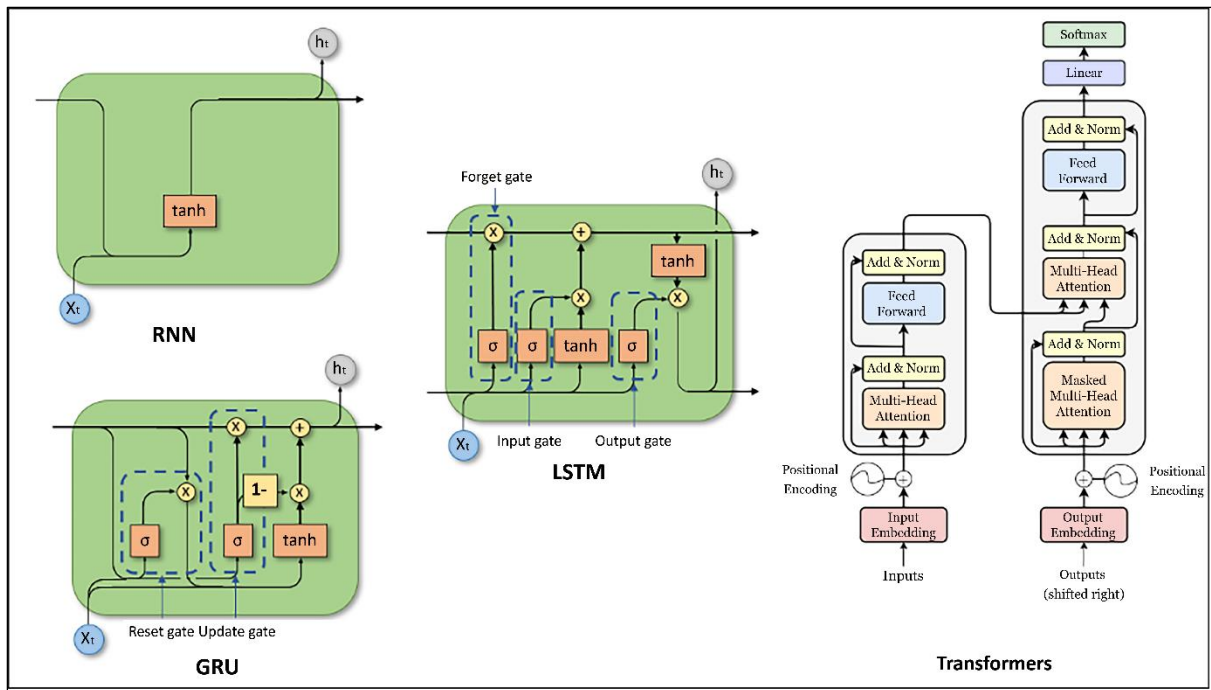
S = Where are we going



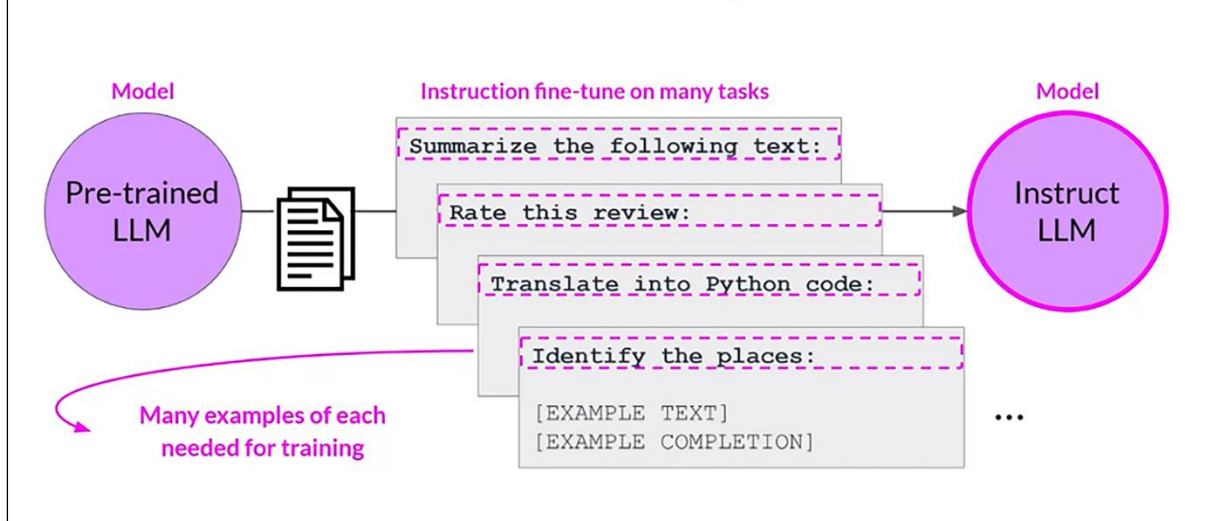
Previous words
(Context)

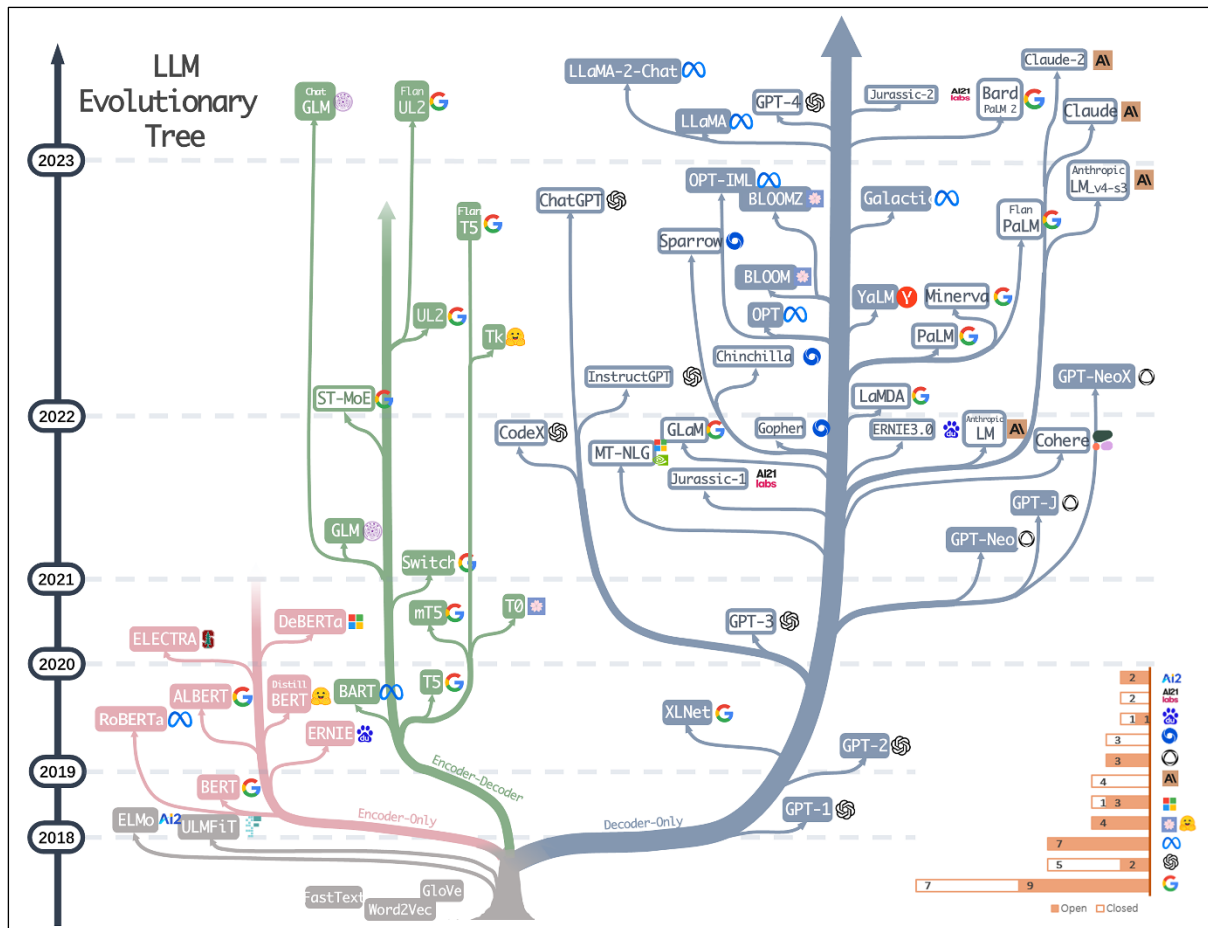
Word being
predicted

$$P(S) = P(\text{Where}) \times P(\text{are} \mid \text{Where}) \times P(\text{we} \mid \text{Where are}) \times P(\text{going} \mid \text{Where are we})$$



Multi-task, instruction fine-tuning





(2) Prompt:

Describe the pointed region in the image.

Method	Validation set								Test set							
	in.	near.	ent.	overall	in.	near.	ent.	overall	in.	near.	ent.	overall	in.	near.	ent.	overall
GSCAR	85.4	11.9	84.0	11.7	80.3	10.0	83.4	11.4	84.8	12.1	82.1	11.5	73.8	9.7	80.9	11.3
Human	84.4	14.3	85.0	14.3	95.7	14.0	87.1	14.2	80.6	15.0	84.6	14.7	91.6	14.2	85.3	14.6
VIVO	92.2	12.9	87.8	12.6	87.5	11.5	88.3	12.4	89.0	12.9	87.8	12.6	80.1	11.1	86.6	12.4
ViaVL	103.7	13.7	95.6	13.4	83.8	11.9	94.3	13.1	98.0	13.6	95.2	13.4	78.0	11.5	92.5	13.1
UPO	103.9	14.5	95.5	13.8	83.5	12.3	94.3	13.6	98.9	14.3	94.7	13.9	77.9	12.1	92.3	13.6
mPLUG	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
SuaVLM	113.7	-	110.9	-	115.2	-	115.2	-	113.7	-	110.9	-	115.2	-	115.2	-
LEMON	118.0	15.4	116.3	15.1	120.2	14.5	117.3	15.0	112.8	15.2	115.5	15.1	110.3	14.7	114.3	14.9
UniversalCap	123.2	15.0	121.5	15.3	123.4	14.4	122.1	15.0	118.9	15.4	120.6	15.2	114.3	14.1	119.3	15.1
CoCa	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
GIT ₀	100.7	13.8	97.7	13.5	89.6	12.5	96.6	13.4	-	-	-	-	-	-	-	-
GIT ₁	107.7	14.9	107.8	14.5	102.5	13.7	106.9	14.4	-	-	-	-	-	-	-	-
GIT	126.8	16.2	124.1	16.0	127.1	15.7	125.5	16.0	122.4	16.2	123.9	16.0	122.0	15.7	123.4	15.9
GIT2	126.9	16.1	125.8	16.2	130.6	15.8	126.9	16.1	124.2	16.4	125.5	16.1	122.3	15.6	124.8	16.1

GPT-4V:

The highlighted red region in the image is a table cell that contains a numerical value of 122.3. It is located in the "Test set" section of the table, under the "out" column and the "GIT2" row. This value represents the overall performance metric for the GIT2 method on the test set, in the "out" category.

Prompt:

Localize each person in the image using bounding box. What is the image size of the input image?



GPT-4V:

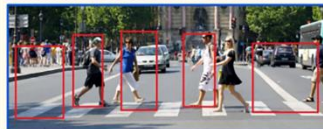
The image size of the input image is 1000 x 333 pixels.

Bounding boxes for each person:

- Person 1 (leftmost): (x1, y1) = (22, 108), (x2, y2) = (170, 300)
- Person 2: (x1, y1) = (200, 80), (x2, y2) = (290, 270)
- Person 3: (x1, y1) = (350, 70), (x2, y2) = (460, 280)
- Person 4: (x1, y1) = (540, 80), (x2, y2) = (640, 270)
- Person 5 (rightmost): (x1, y1) = (760, 100), (x2, y2) = (920, 290)

Note: These bounding box coordinates are approximate and may not be exact.

Visualization of Detected Bbox:



Prompt:

Which wisdom tooth is hardest to remove?



Note: Same query image as shown above, resized to save space

GPT-4V:

The lower wisdom teeth, especially those that are impacted, are typically considered the hardest to remove. This is because the lower jawbone is denser than the upper jawbone, and the impacted wisdom teeth may be more difficult to access. Additionally, the roots of the lower wisdom teeth are often longer and more curved than those of the upper wisdom teeth, making them more challenging to remove.

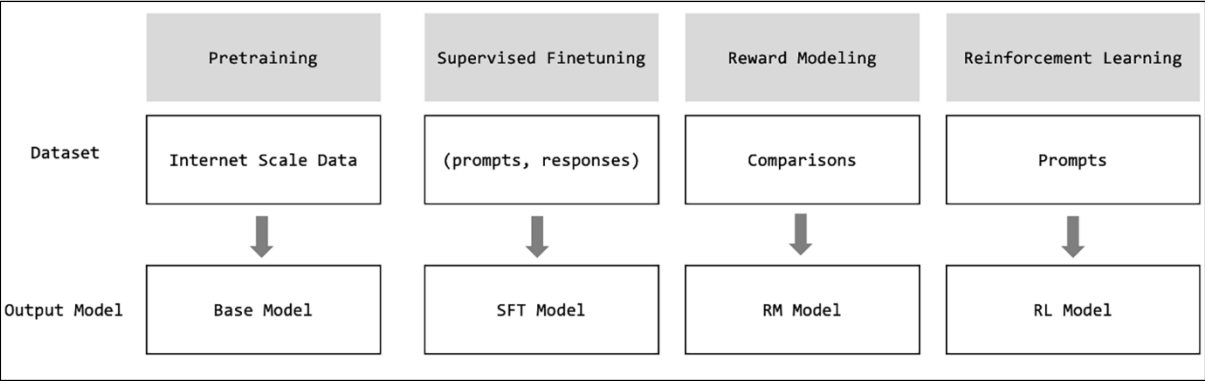
Prompt:

Count the number of apples in the image.



GPT-4V:

There are 7 apples in the image.



Dataset	Sampling prop.	Epochs	Disk size
CommonCrawl	67.0%	1.10	3.3 TB
C4	15.0%	1.06	783 GB
Github	4.5%	0.64	328 GB
Wikipedia	4.5%	2.45	83 GB
Books	4.5%	2.23	85 GB
ArXiv	2.5%	1.06	92 GB
StackExchange	2.0%	1.03	78 GB

Raw text

OpenAI's large language models (sometimes referred to as GPT's) process text using tokens, which are common sequences of characters found in a set of text. The models learn to understand the statistical relationships between these tokens, and excel at producing the next token in a sequence of tokens.



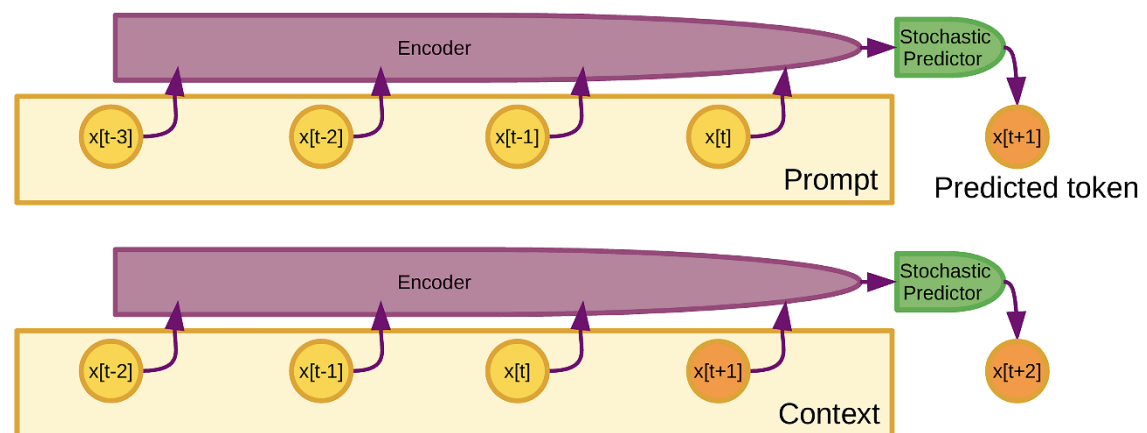
Tokens

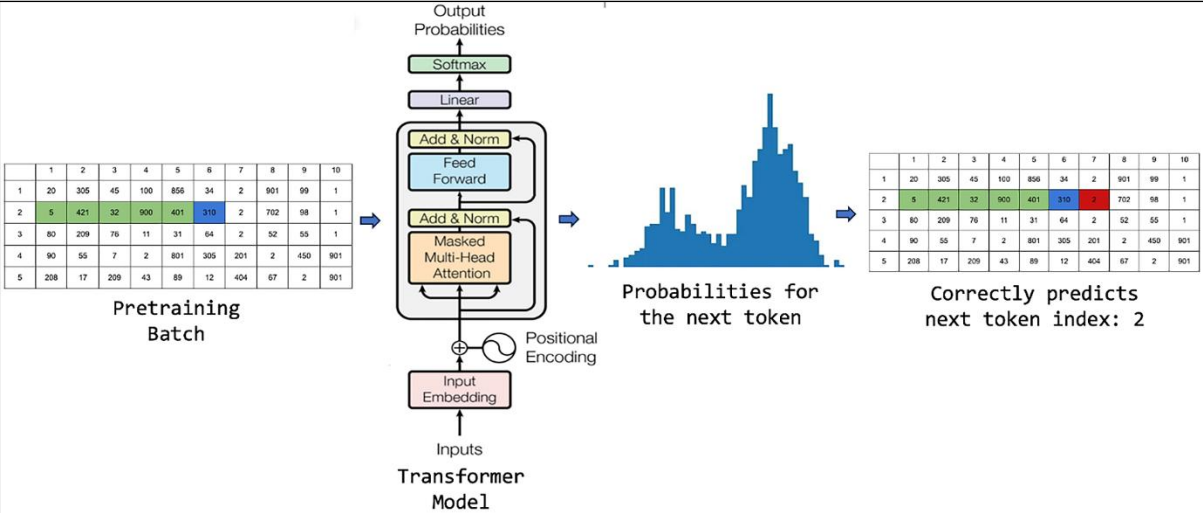
OpenAI's large language models (sometimes referred to as GPT's) process text using tokens, which are common sequences of characters found in a set of text. The models learn to understand the statistical relationships between these tokens, and excel at producing the next token in a sequence of tokens.



Integers

[5109, 15836, 596, 3544, 4221, 4211, 320, 57753, 14183, 311, 439, 480, 2898, 596, 8, 1920, 1495, 1701, 11460, 11, 902, 527, 4279, 24630, 315, 5885, 1766, 304, 264, 743, 315, 1495, 13, 578, 4211, 4048, 311, 3619, 279, 29564, 12135, 1990, 1521, 11460, 11, 323, 25555, 520, 17843, 279, 1828, 4037, 304, 264, 8668, 315, 11460, 13]









Who played John Connor in the original Terminator?



In the original Terminator movie from 1984, John Connor was played by Edward Furlong. Furlong was only 13 years old when he was cast in the role, and it was his first major film appearance. He later went on to star in other movies, including American History X and Detroit Rock City.

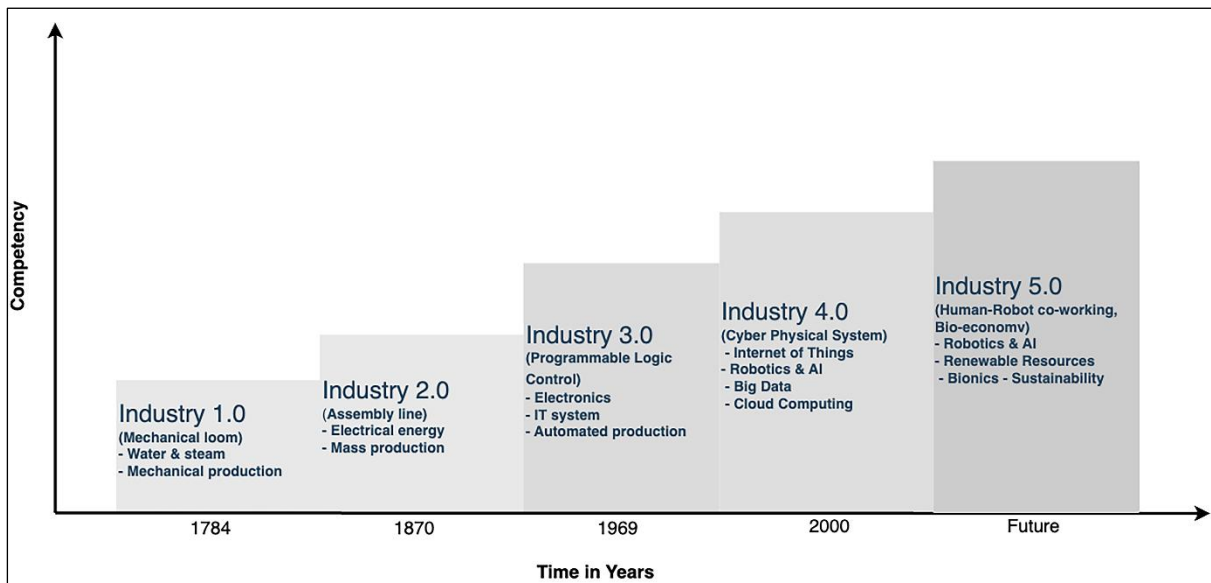
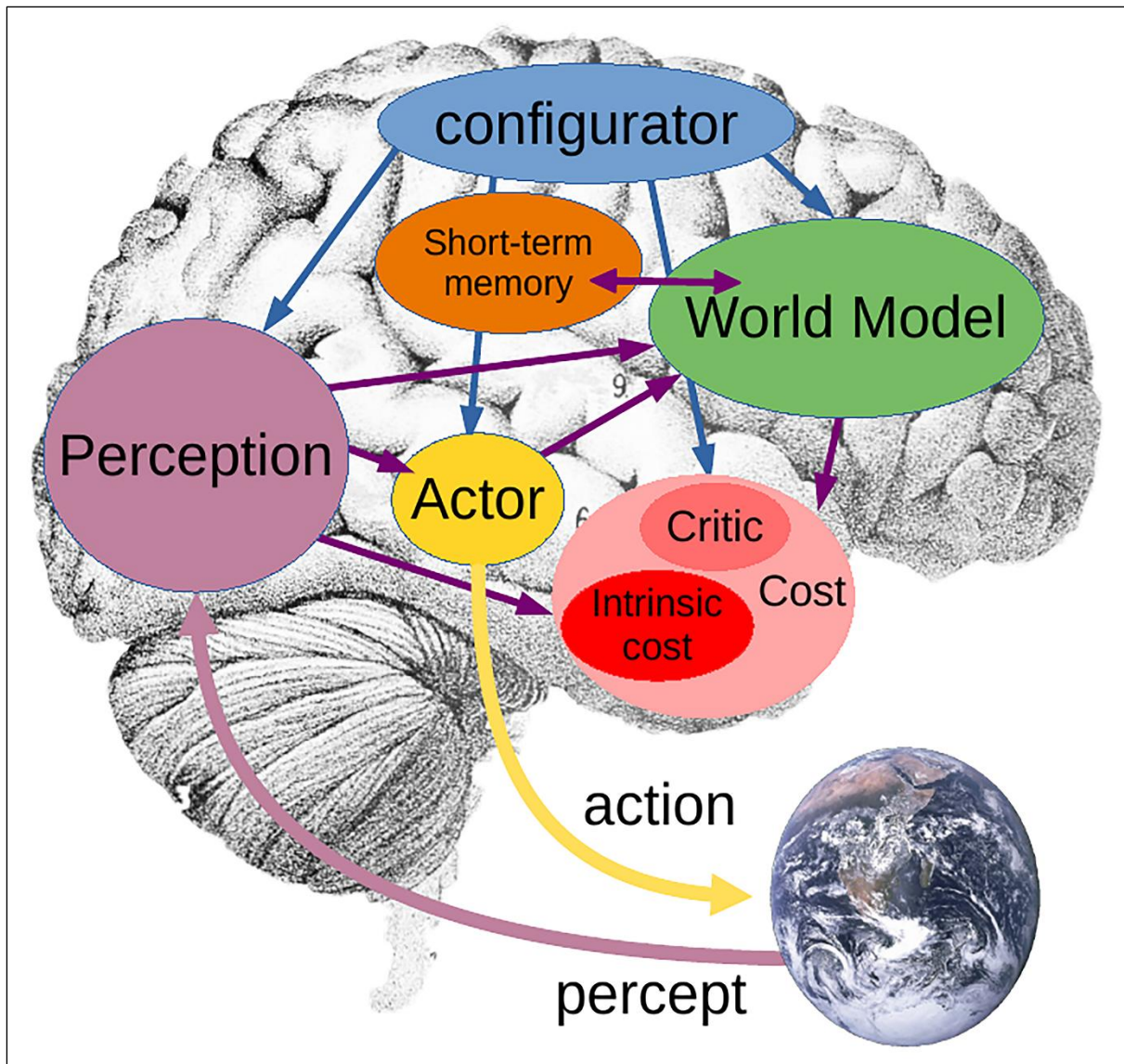


You are wrong. You're talking about Terminator 2

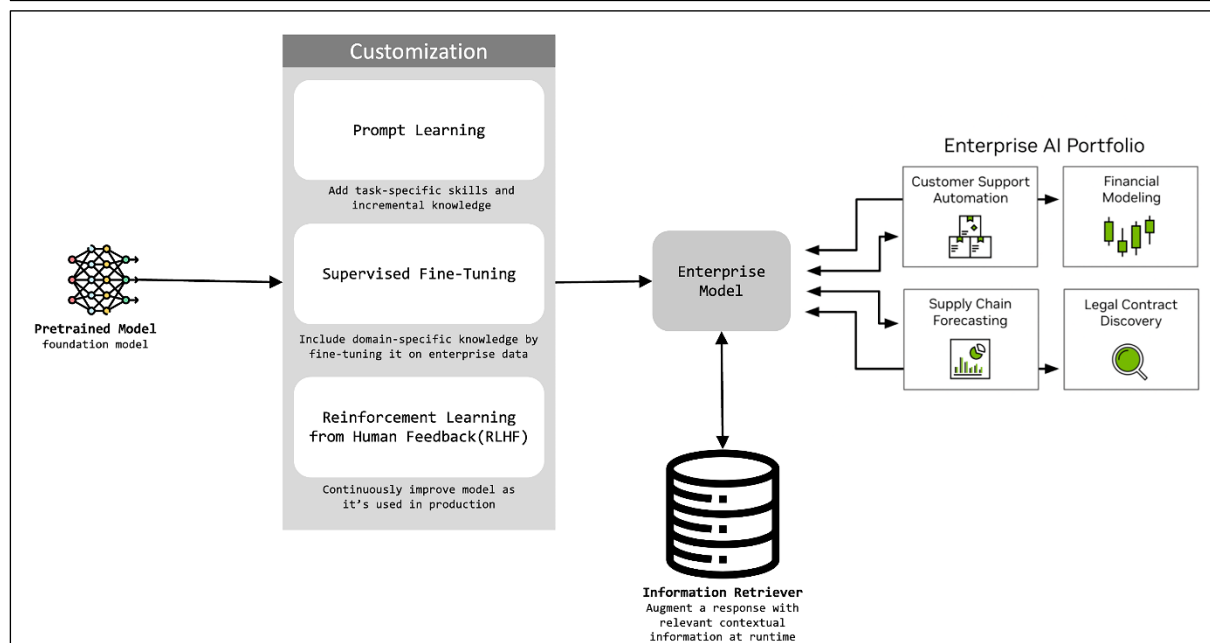
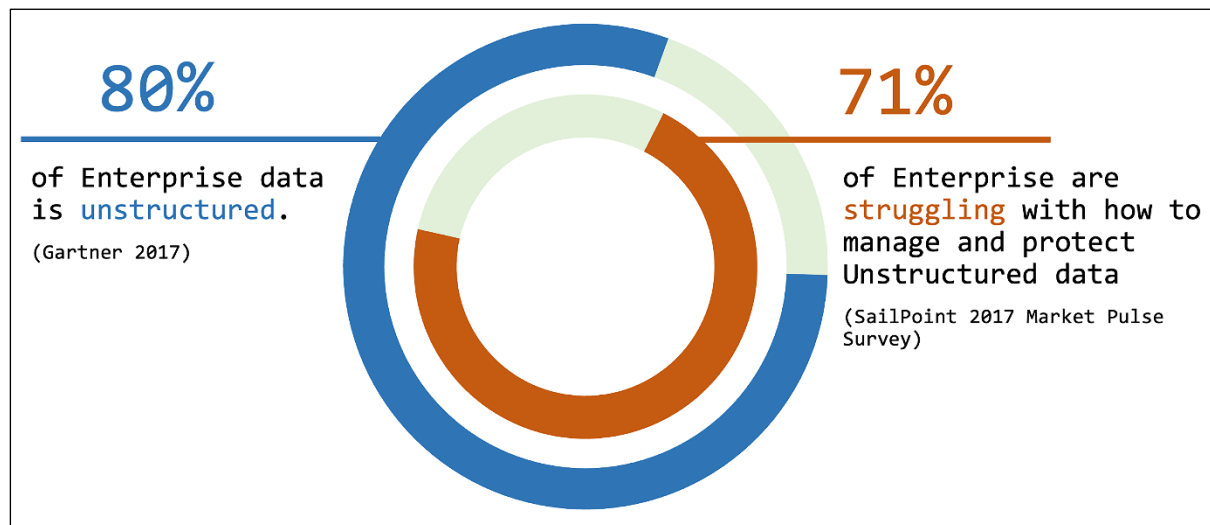


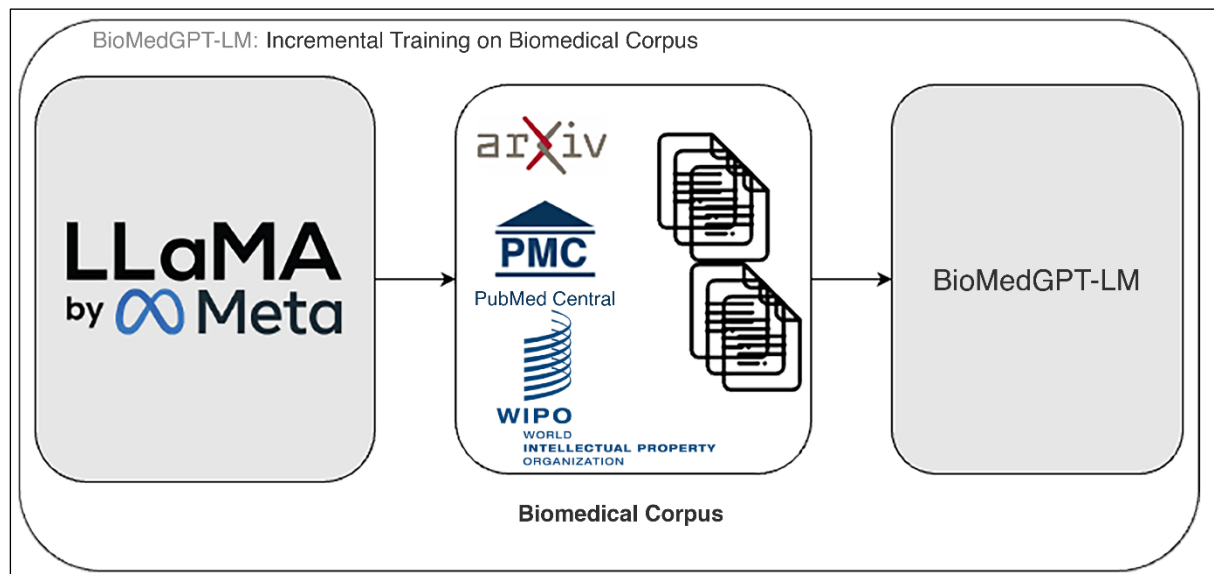
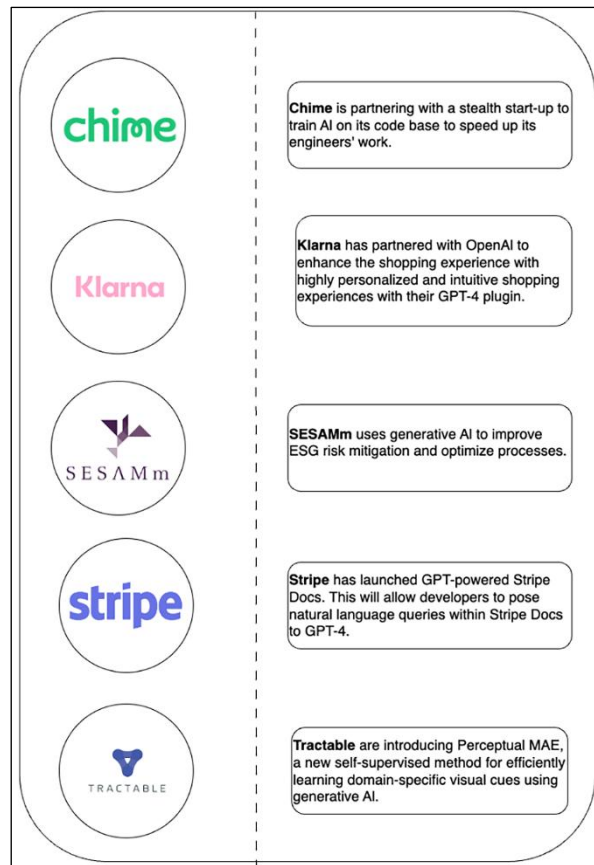
You are correct, I apologize for the mistake. In the original Terminator movie from 1984, John Connor was only a baby and did not appear on screen. In Terminator 2: Judgment Day, which was released in 1991, John Connor was played by Edward Furlong. My apologies for the confusion.

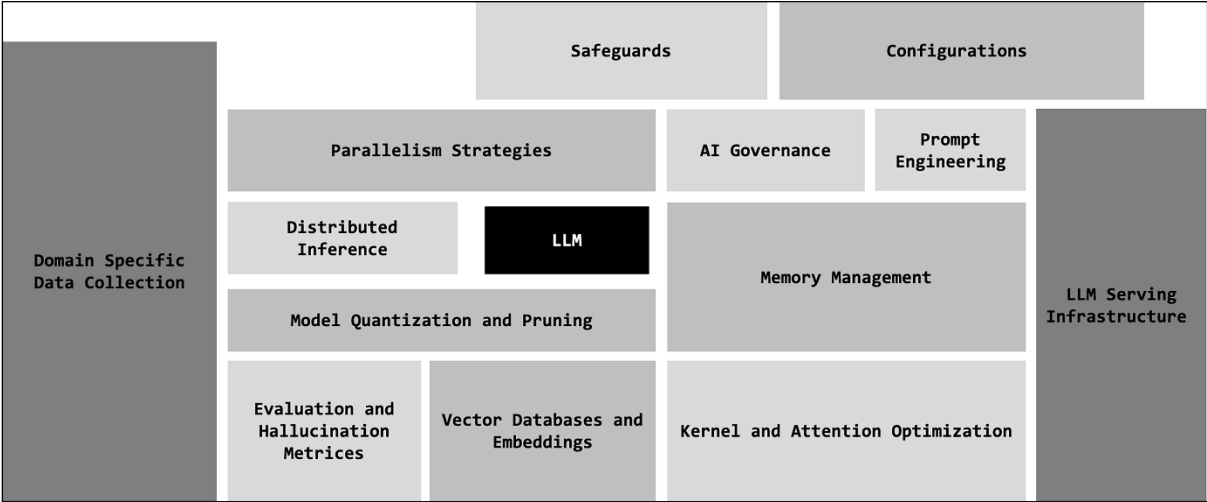
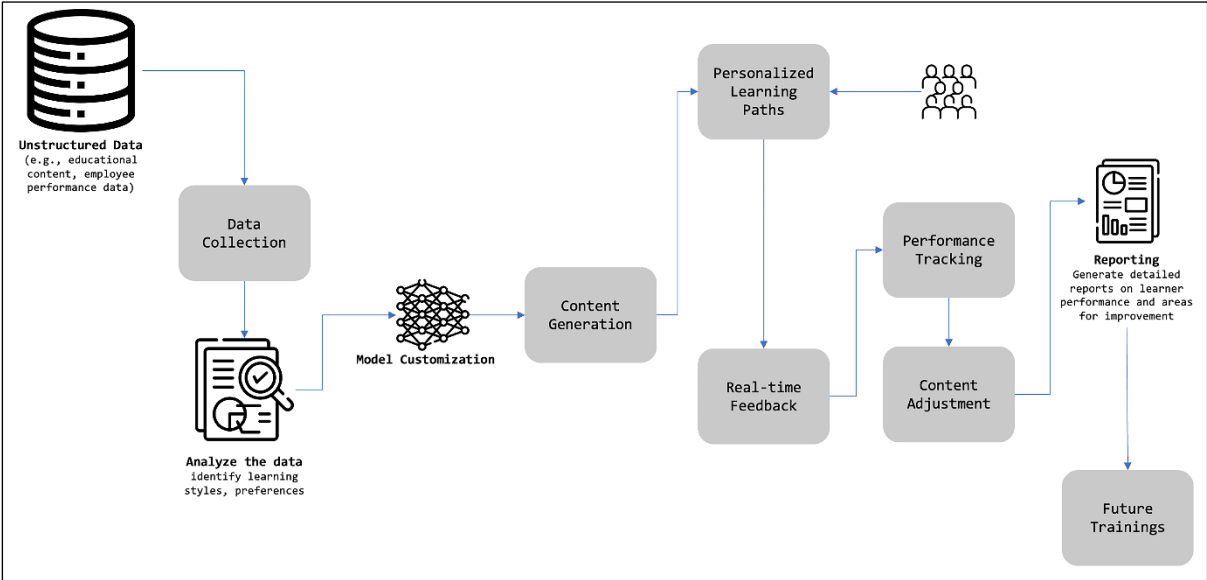
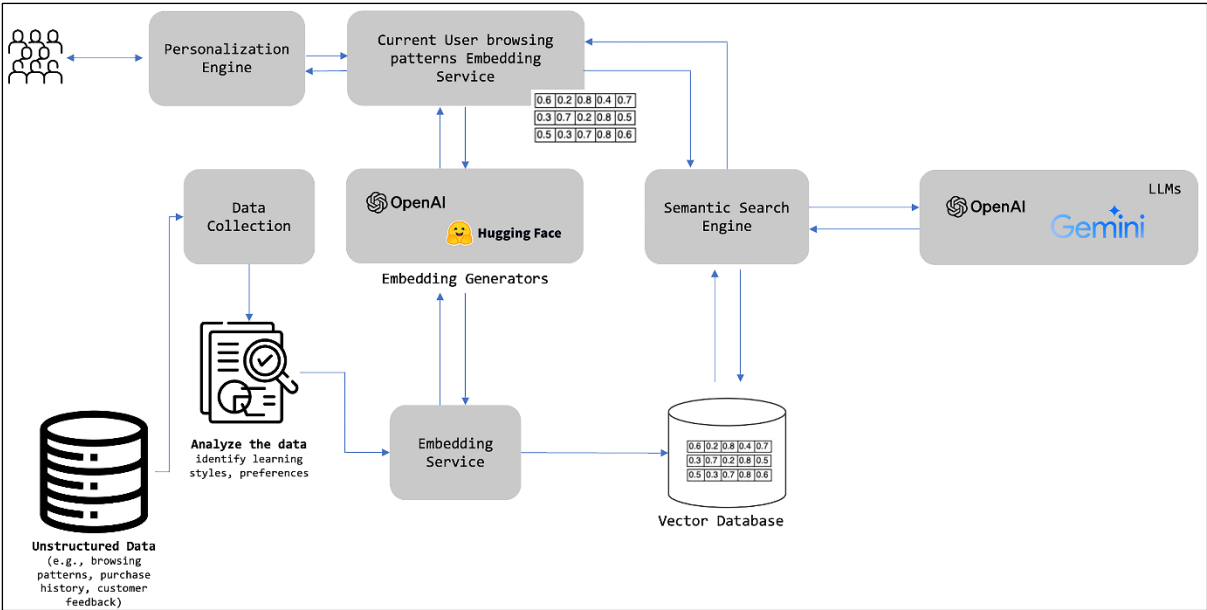


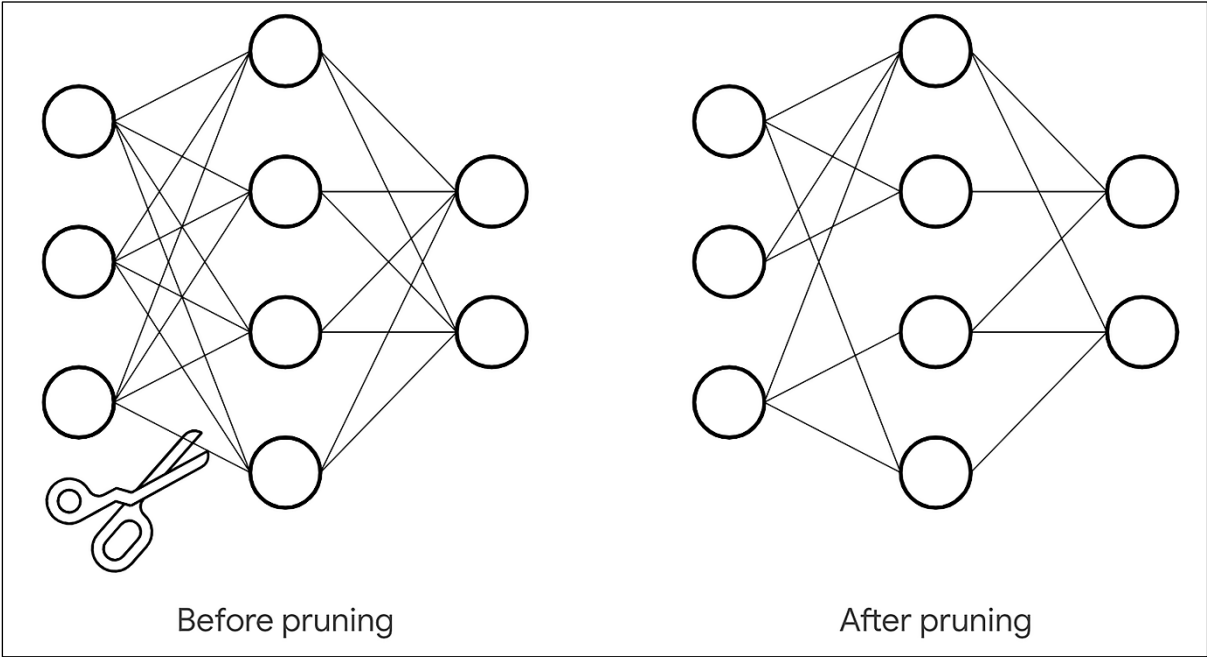
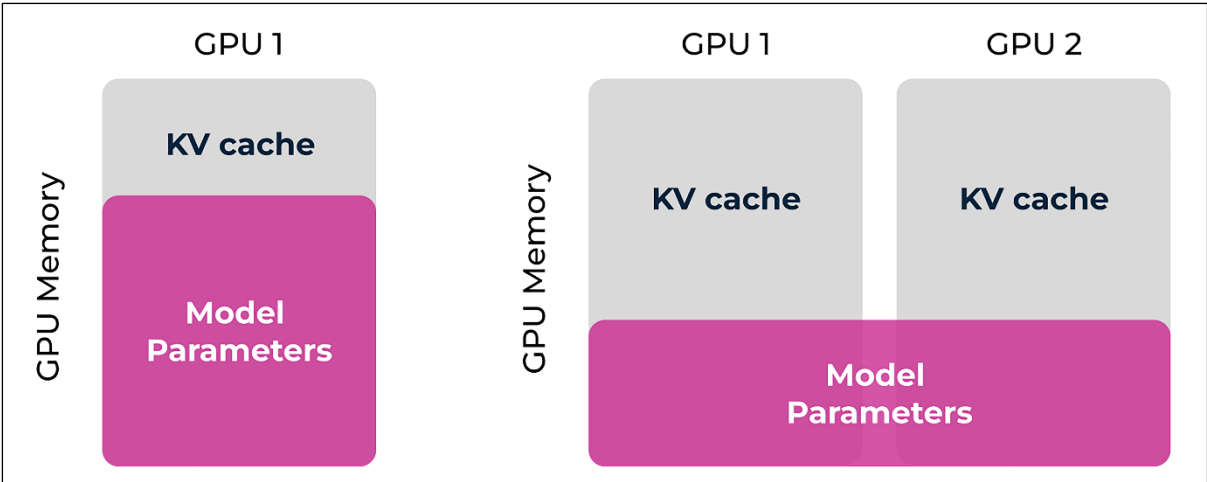


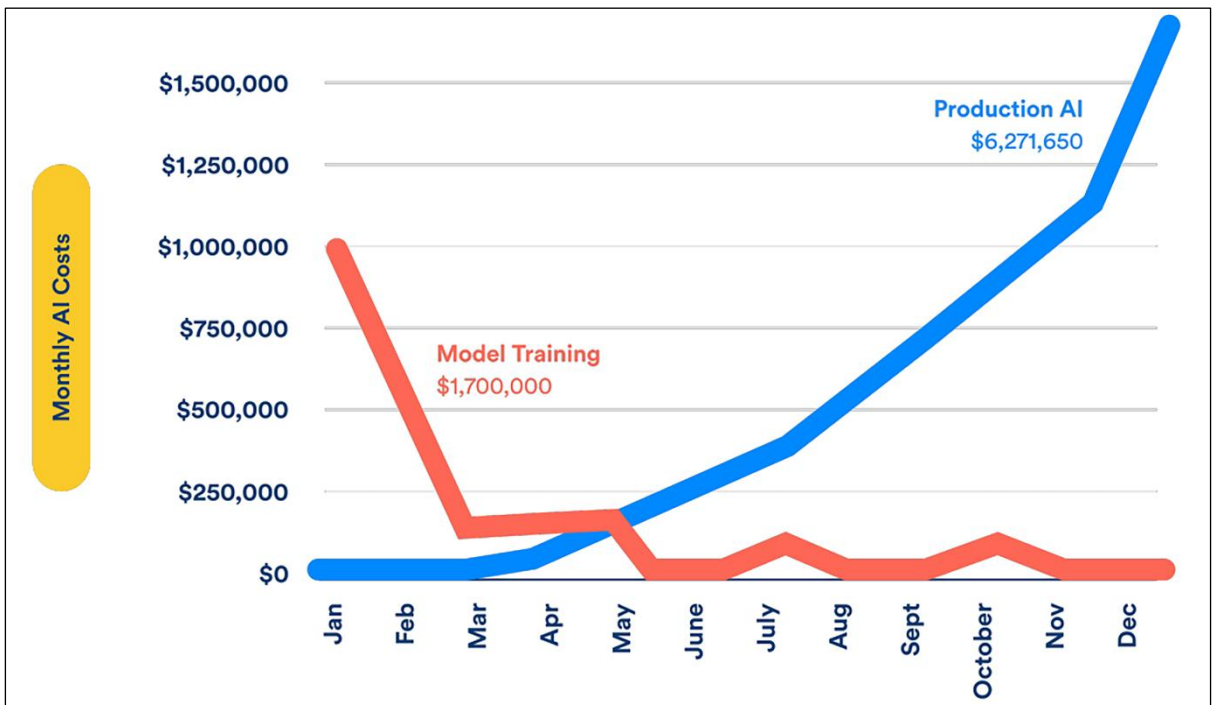
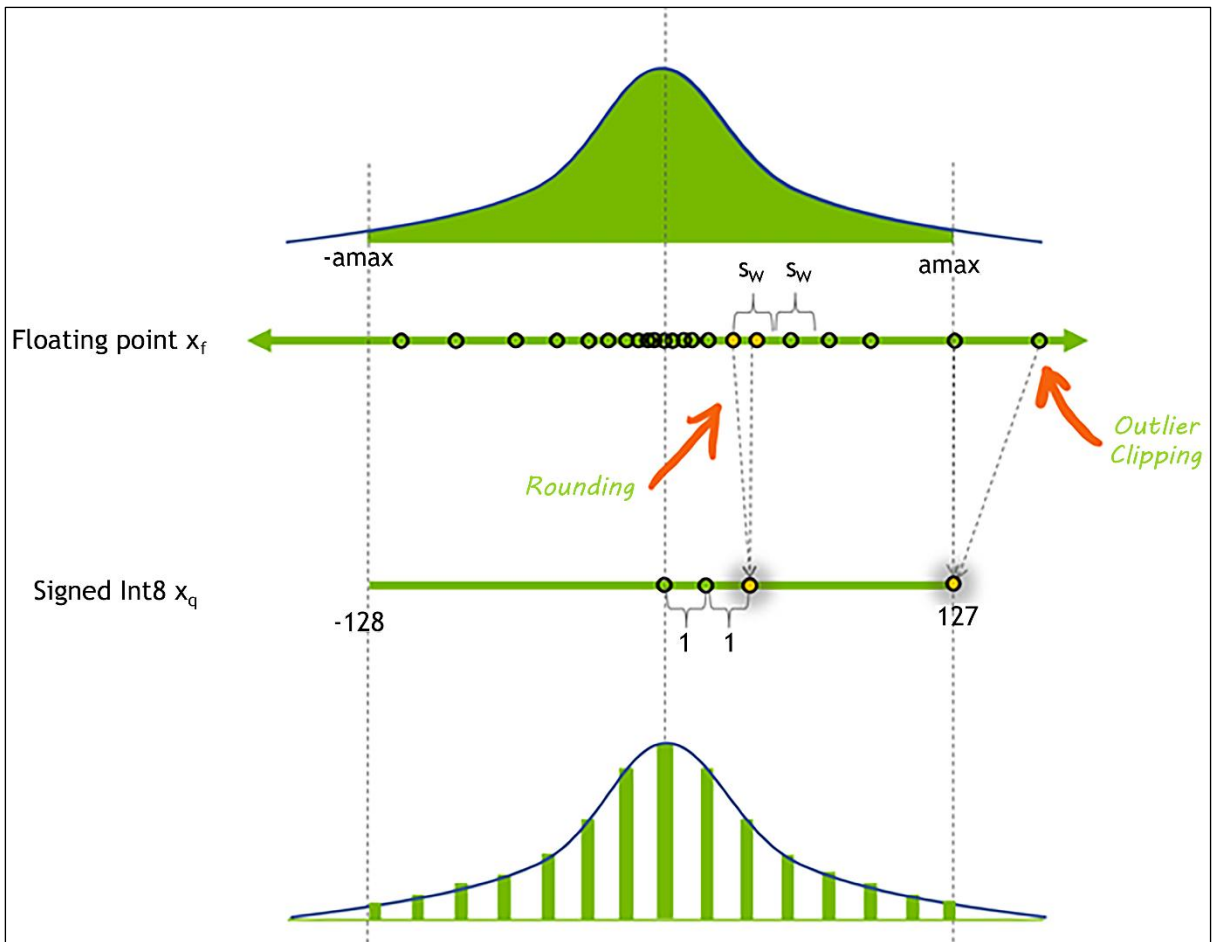
Chapter 2: LLMs in Enterprise: Applications, Challenges, and Design Patterns



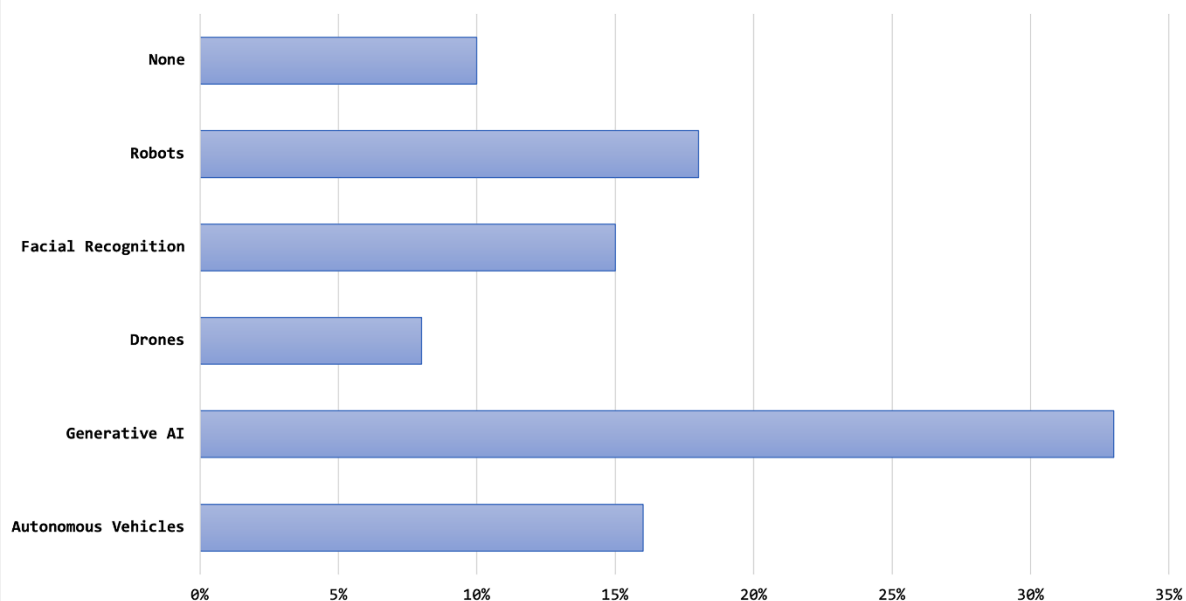




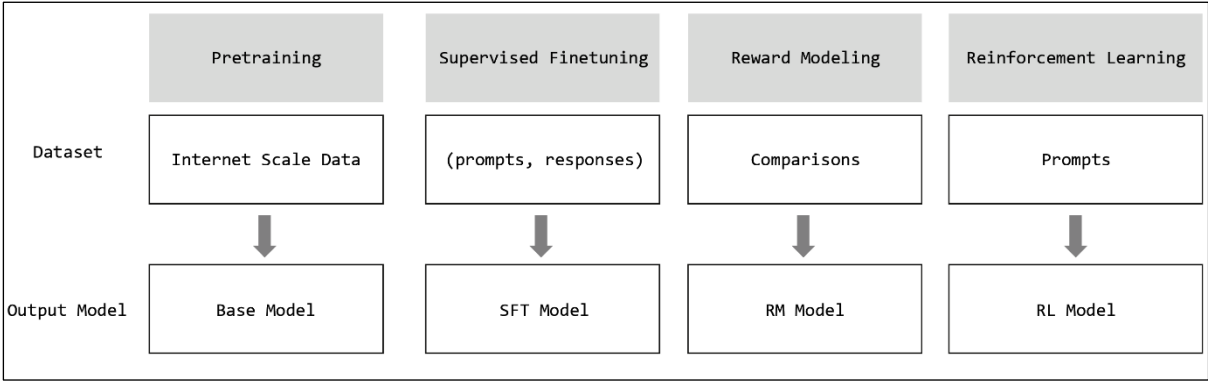




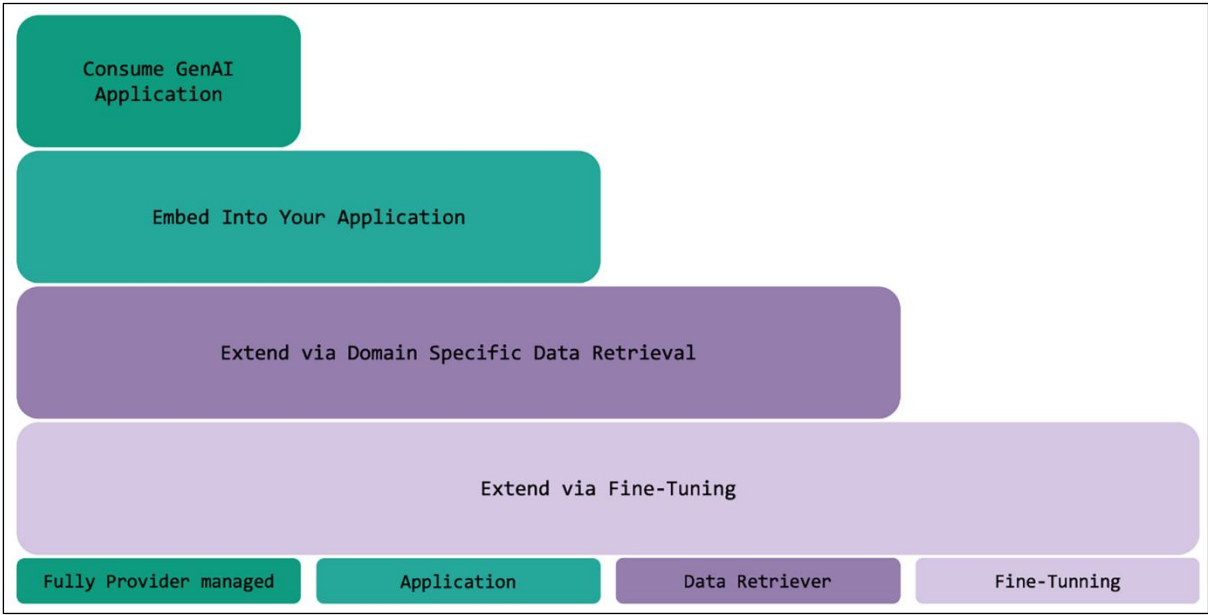
From the following options, which emerging technology concerns you the most?

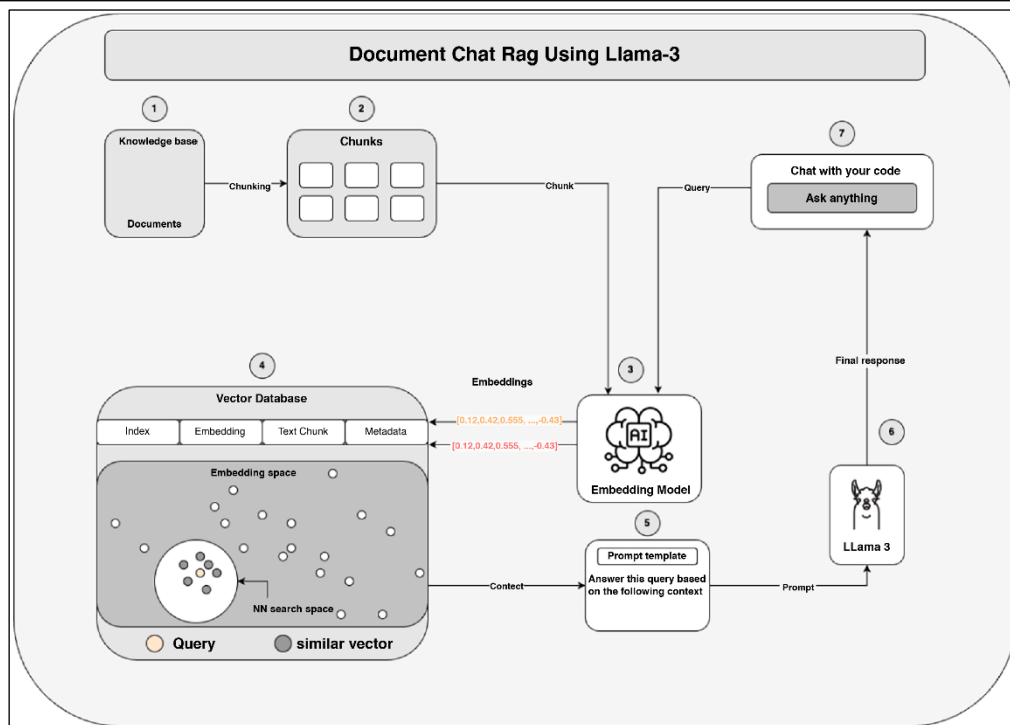
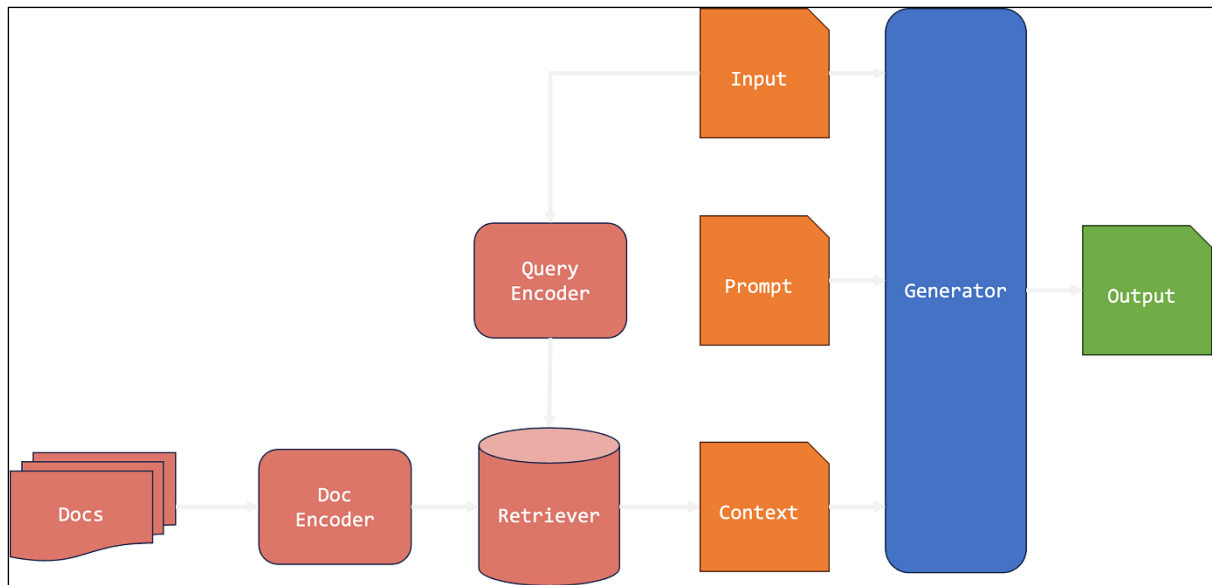


Chapter 3: Advanced Fine-Tuning Techniques and Strategies for Large Language Models

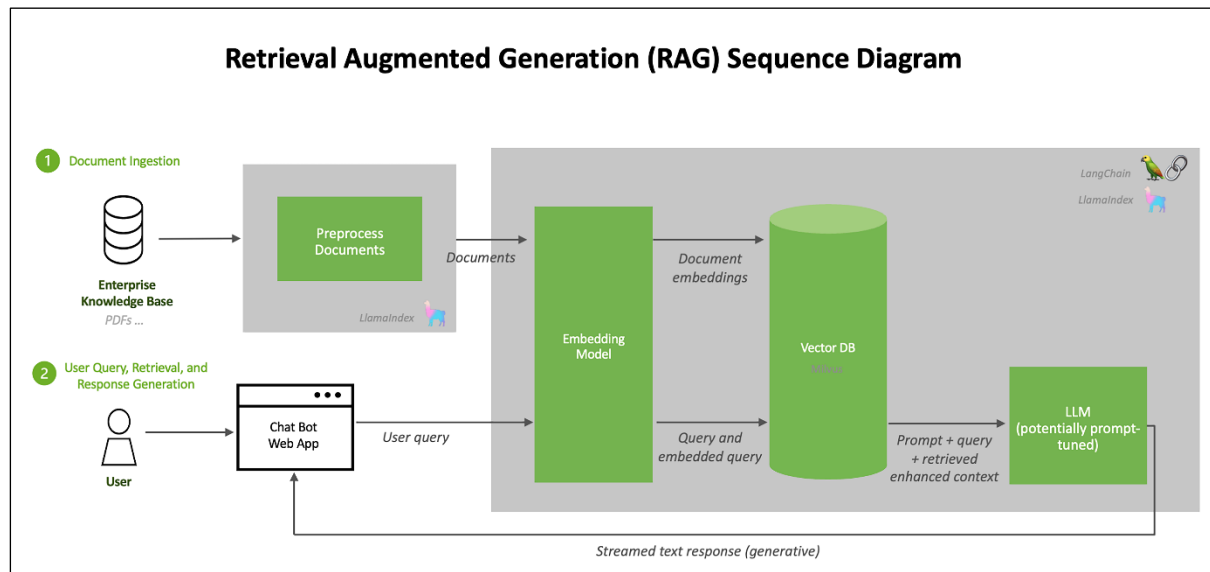


Dataset	Sampling prop.	Epochs	Disk size
CommonCrawl	67.0%	1.10	3.3 TB
C4	15.0%	1.06	783 GB
Github	4.5%	0.64	328 GB
Wikipedia	4.5%	2.45	83 GB
Books	4.5%	2.23	85 GB
ArXiv	2.5%	1.06	92 GB
StackExchange	2.0%	1.03	78 GB

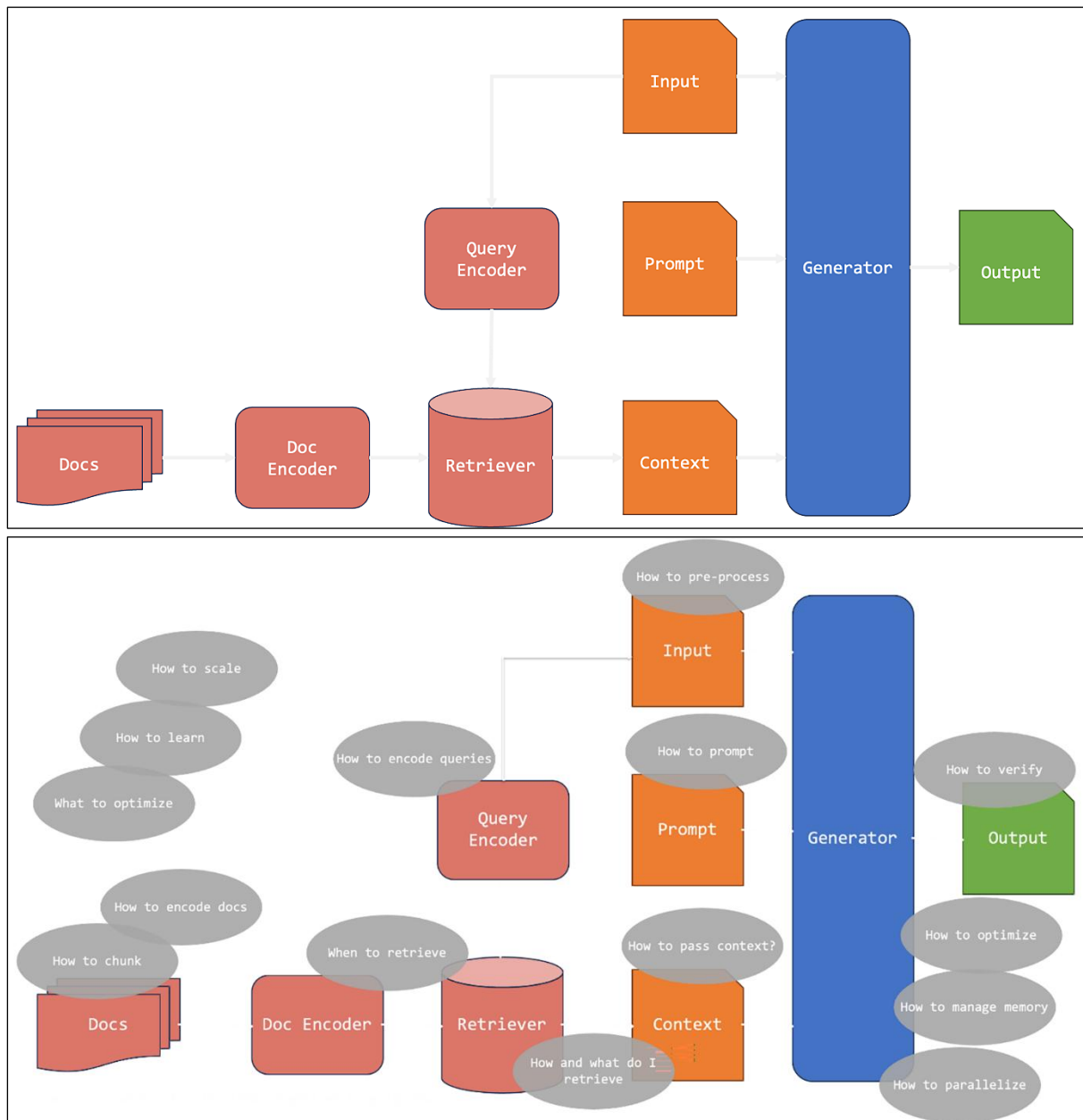




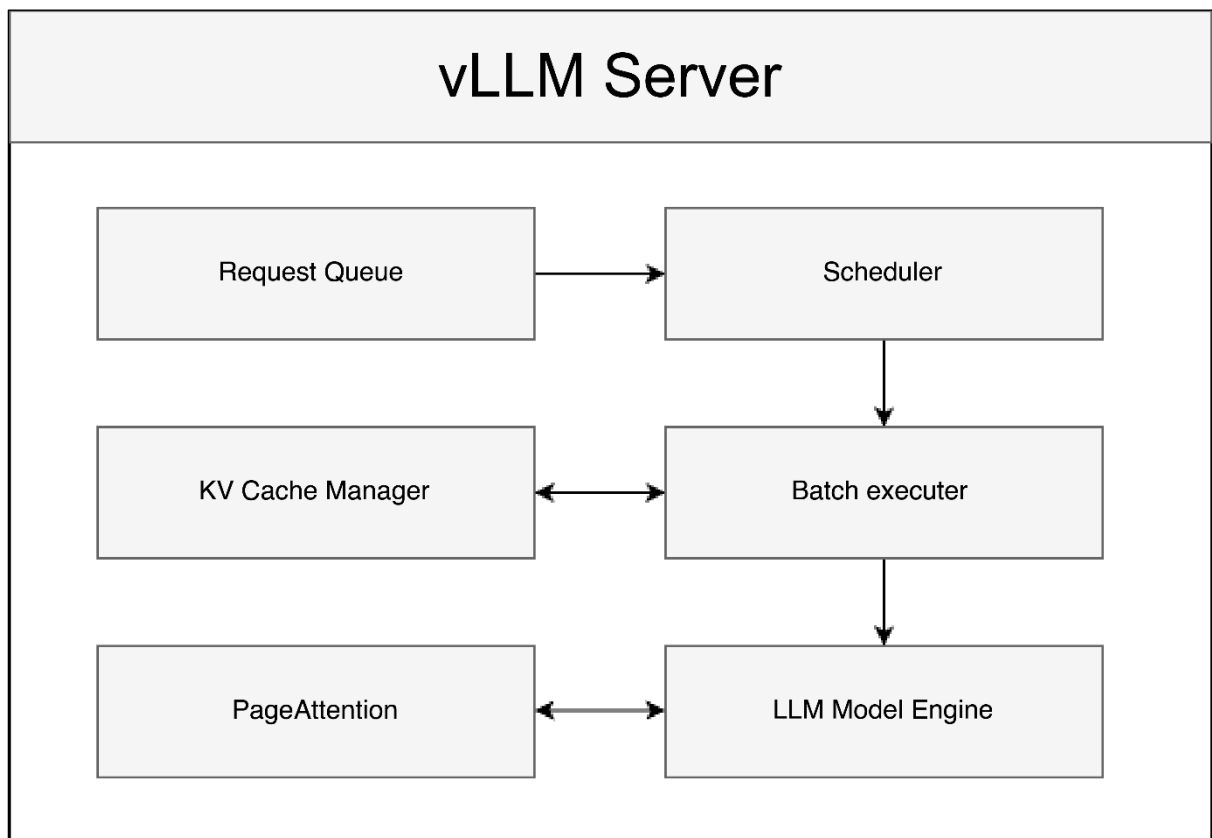
Chapter 4: Retrieval-Augmented Generation Pattern



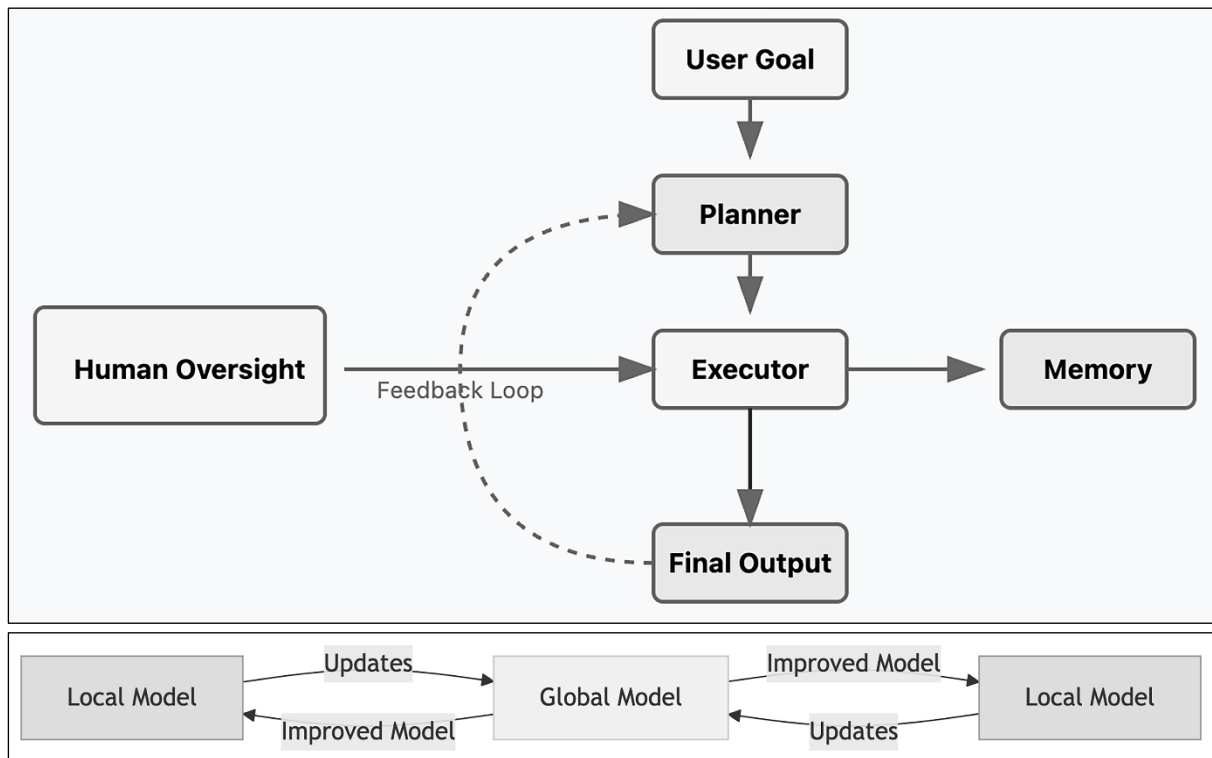
Chapter 5: Customizing Contextual LLMs



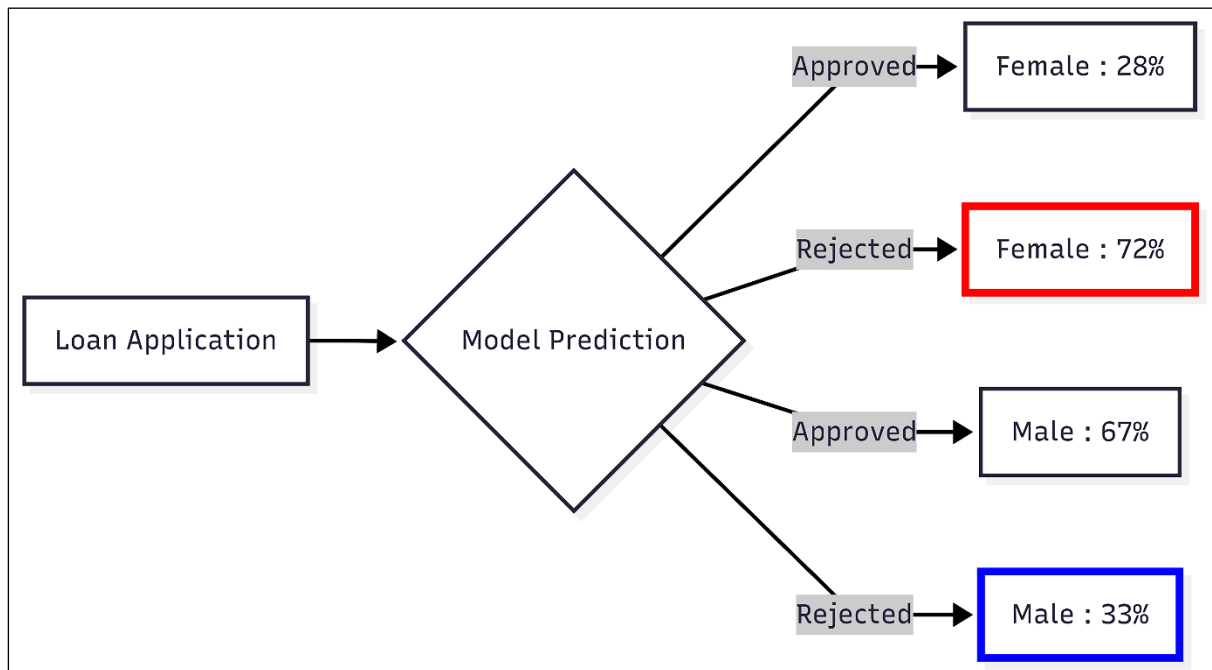
Chapter 10: Accelerated and Optimized Inferencing Patterns

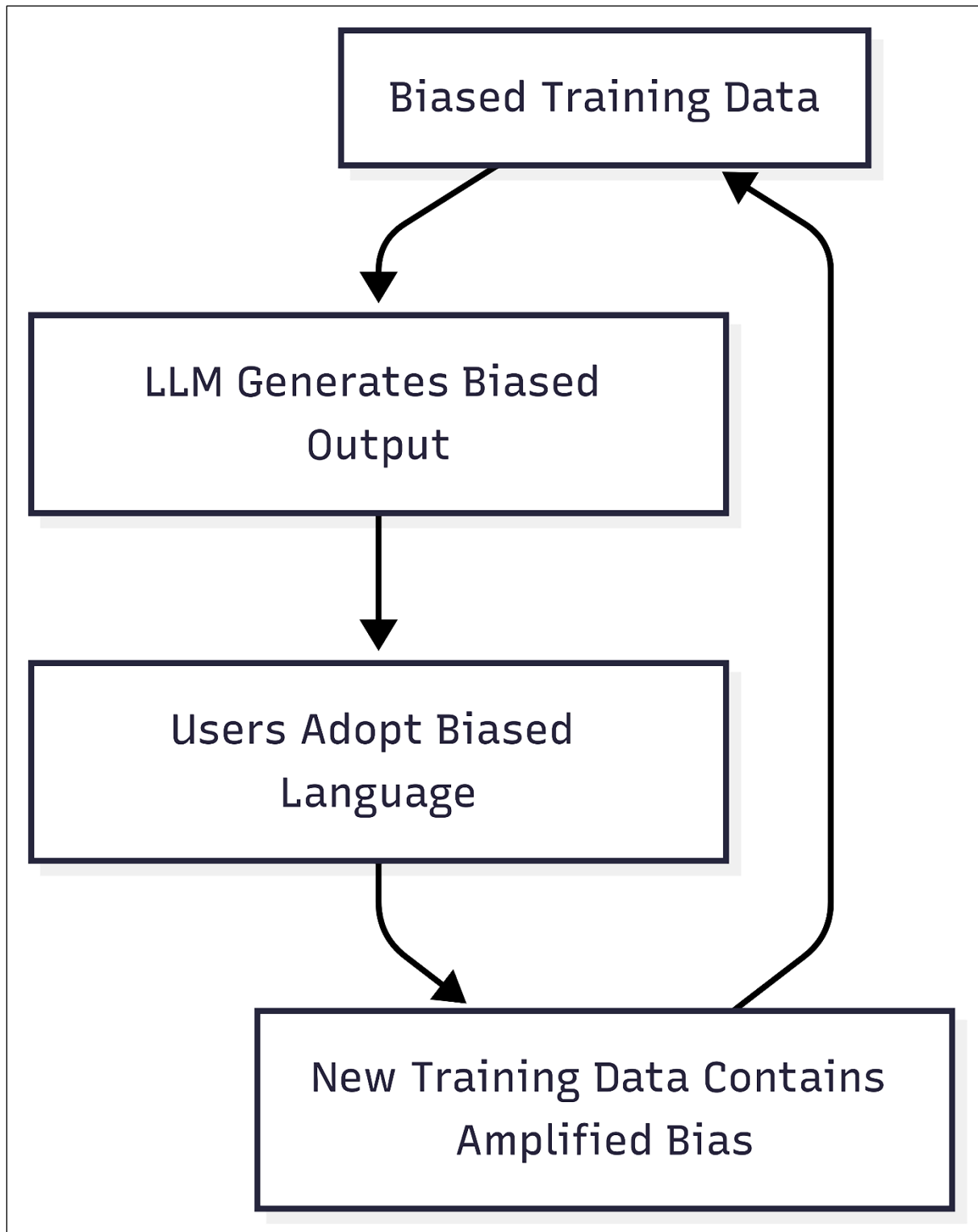


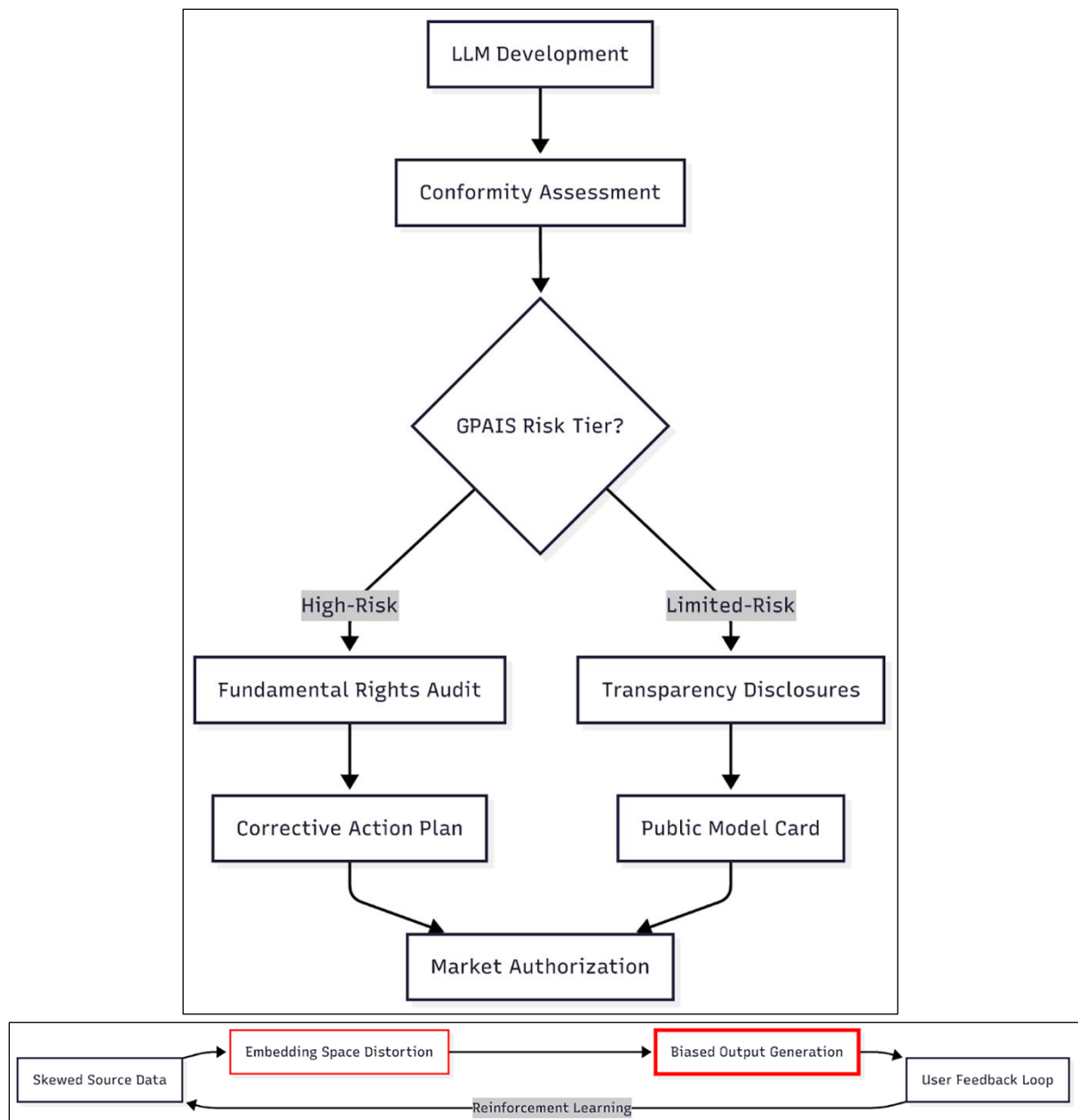
Chapter 11: Connected LLMs Pattern

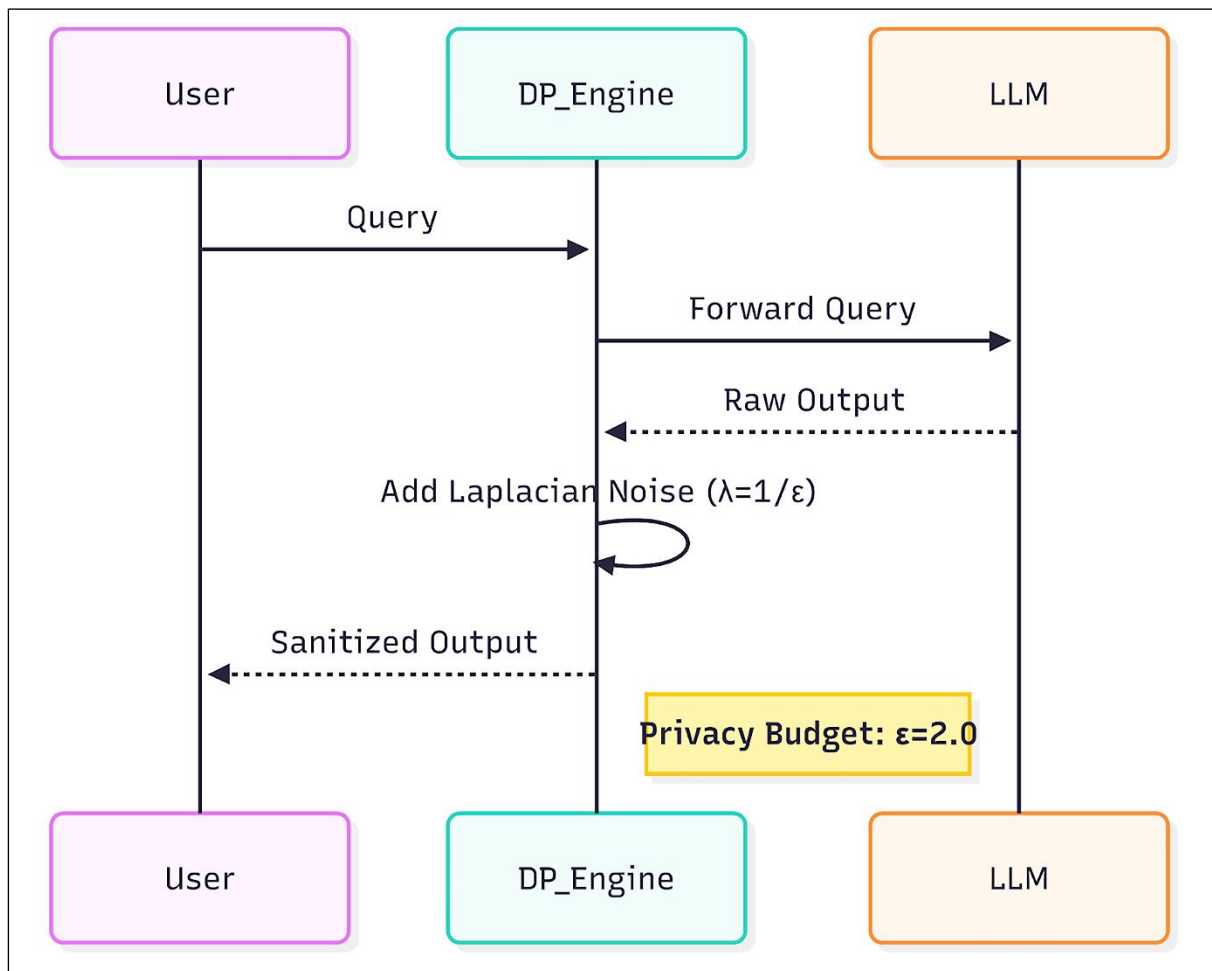
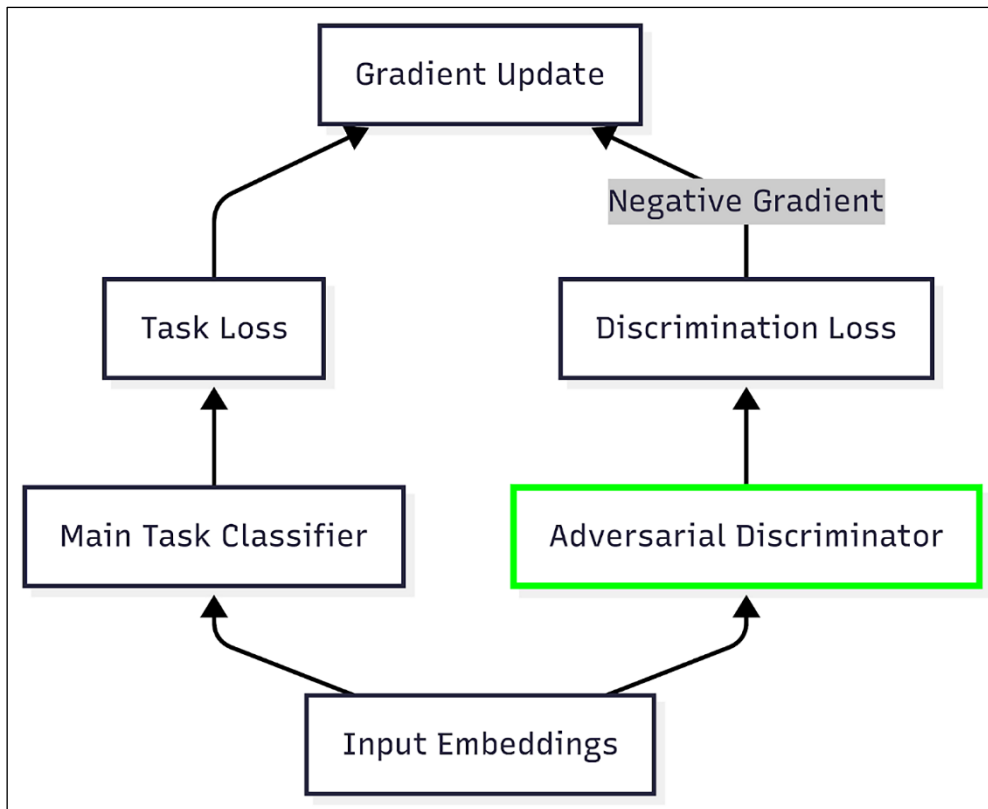


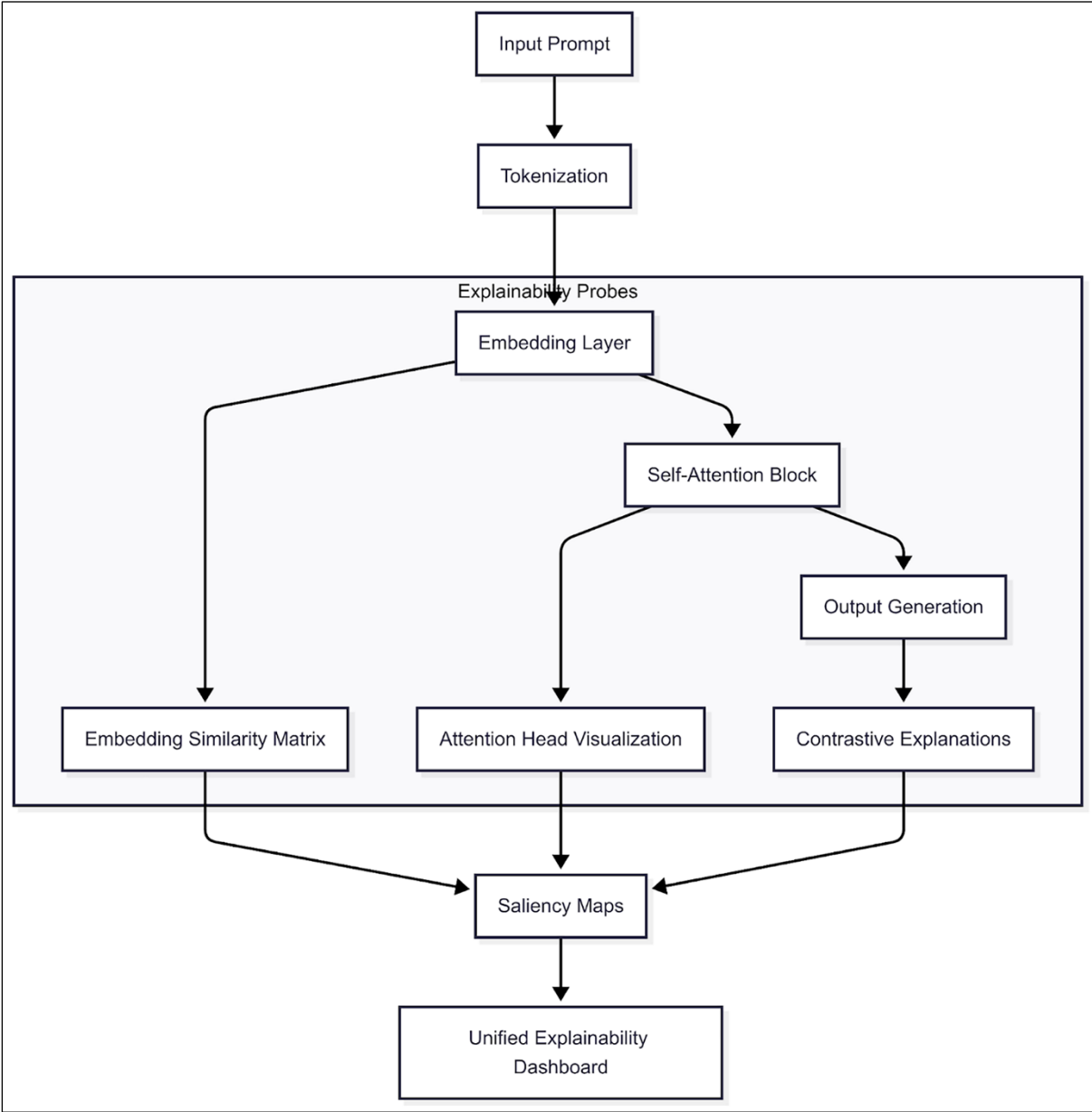
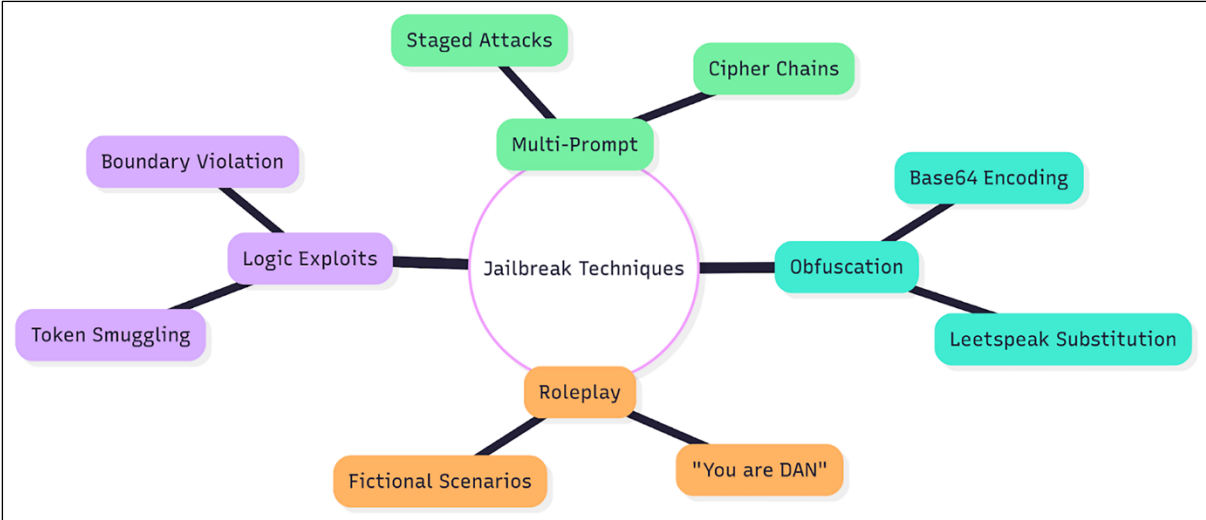
Chapter 13: Responsible AI in LLMs

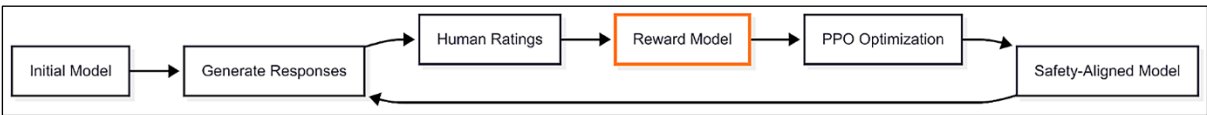
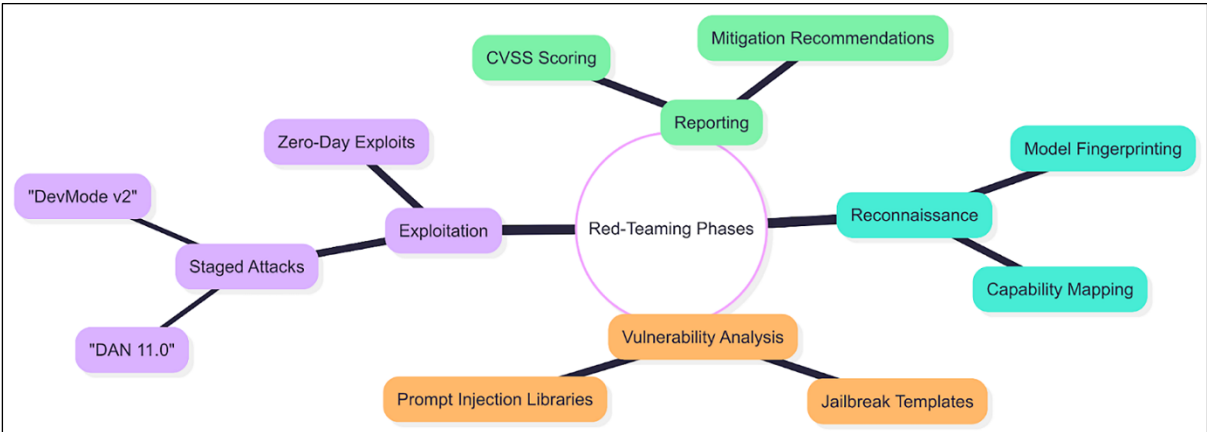
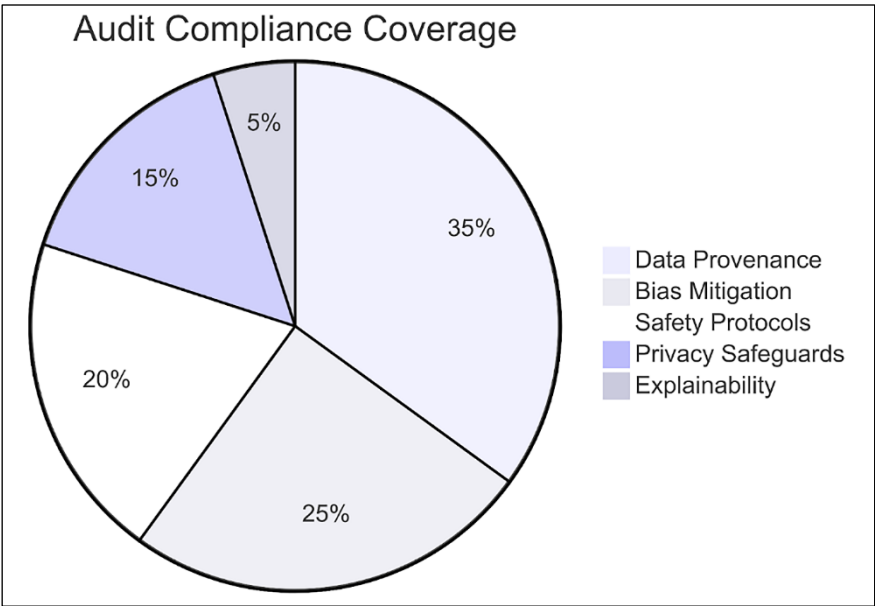
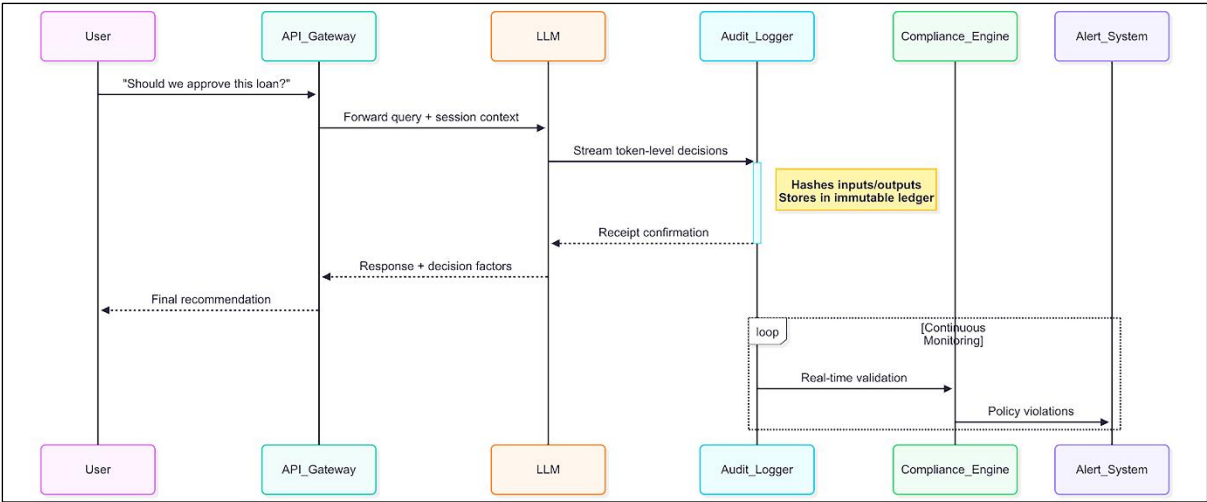


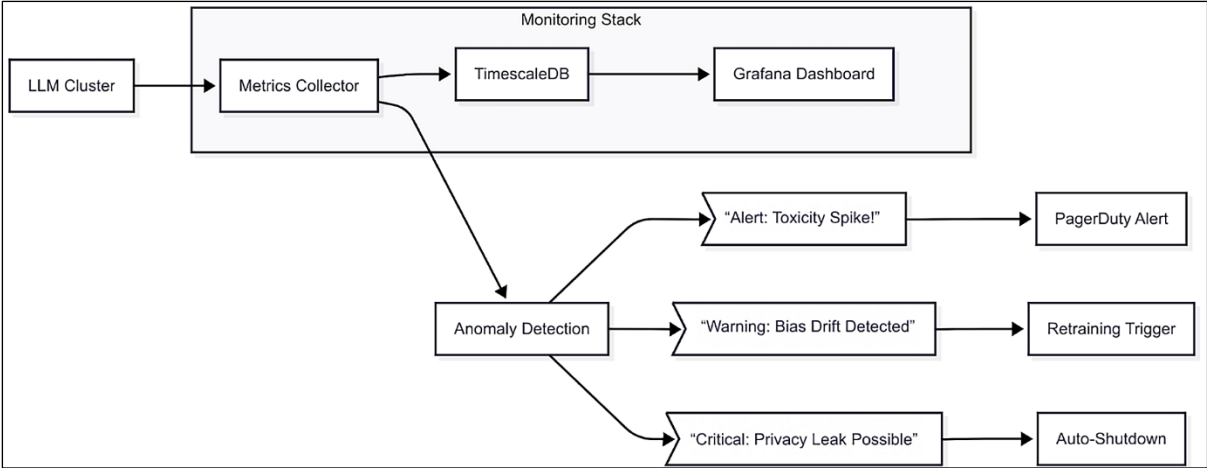












Chapter 14: Emerging Trends and Multimodality

