SILO AI

Large Language Models (LLMs) In Industries

Dr. Shantipriya Parida
Silo AI, Finland

SILO AI

Agenda

- Overview
  - Generative AI
  - Language Model
  - Large Language Models
- LLMs in Industries
- Use Cases
- Case Study
- Conclusion

SILO AI

Generative AI

"We estimate that generative AI could add up to $20 trillion to global GDP by 2030 and save 300 billion work hours a year."

Source: Generative AI Executive Summary (Oliver Wyman Forum)

SILO AI

What is Generative AI ?

- Generative AI refers to a category of artificial intelligence (AI) that focuses on creating or generating new content, data, or information.
- Refers to algorithms and models that can generate new, realistic data.
- Key Characteristics - Creativity, Learning from Data, Autonomy

SILO AI

Capability of Generative AI

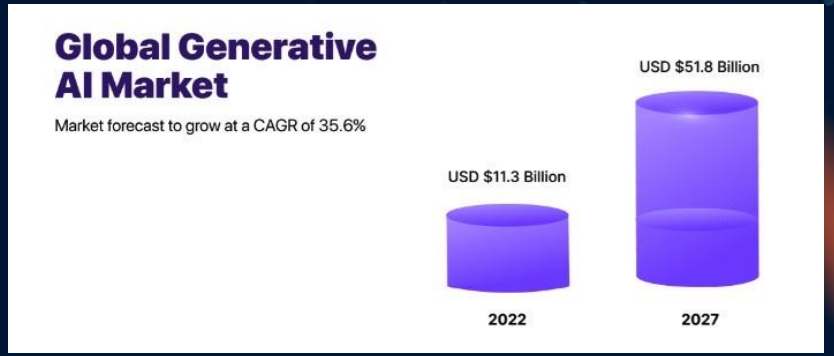
Organizations across different industries can rely on the top generative AI examples as references for creating new and effective solutions. Here are some of the notable applications of generative AI which can help you identify the true potential of generative AI.

- Content Generation
- Chatbots and Virtual Assistants
- Music Creation
- Image Creation and Editing
- 3D Modeling
- Code Generation
- Video Creation & Editing
- Art Creation
- Game Development
- Voice Generation

Source: **101blockchains.com**

Generative AI Market

- According to recent market reports, the AI market is expected to grow in value from $11.3 billion in 2023 to $51.8 billion by 2028.
- ChatGPT setting a record as the fastest-growing internet app in history.
- Large language models (LLMs) are very versatile, their business applications are wide-ranging, from stock trading to fraud detection.



Source:https://shopdev.co/blog/what-are-large-language-models

Language Model (LM)/
Large Language Model (LLM)

SILO AI

What is a Language Model (1/2)

A language model is a probability distribution over sequences of words.

- Given any sequence of words of length m, a language model assigns a probability $P(w_1, ..., w_m)$ to the whole sequence.
- Language models generate probabilities by training on text corpora in one or many languages.
- The probability intuitively tells us how “good” a sequence of tokens is.

SILO AI

What is a Language Model (2/2)

For example, for the vocabulary, V = {ate, ball, cheese, mouse, the}.

The language model might assign

- P (the, mouse, ate, the, cheese) = 0.02
- P (the, cheese, ate, the, mouse) = 0.01
- P (mouse, the, the, cheese, ate) = 0.001

SILO AI

What is a Large Language Model

- A large language model (LLM) is a language model consisting of a neural network with many parameters (typically billions of weights or more), trained on large quantities of unlabeled text using self-supervised learning or semi-supervised learning.
- They are built on neural network architectures, particularly the transformer architecture.

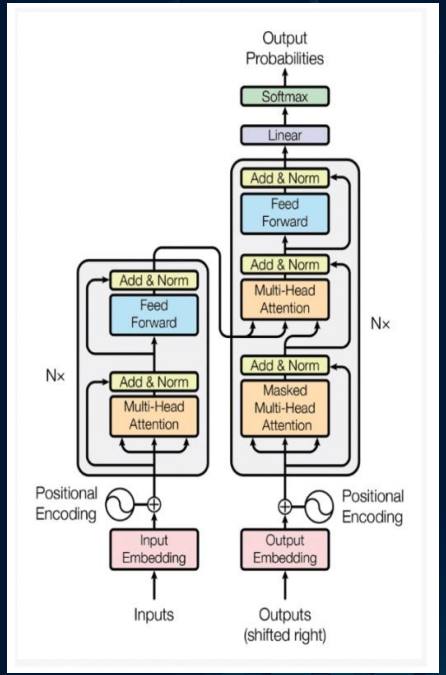


Fig: Transformer architecture

SILO AI

LLM Timeline

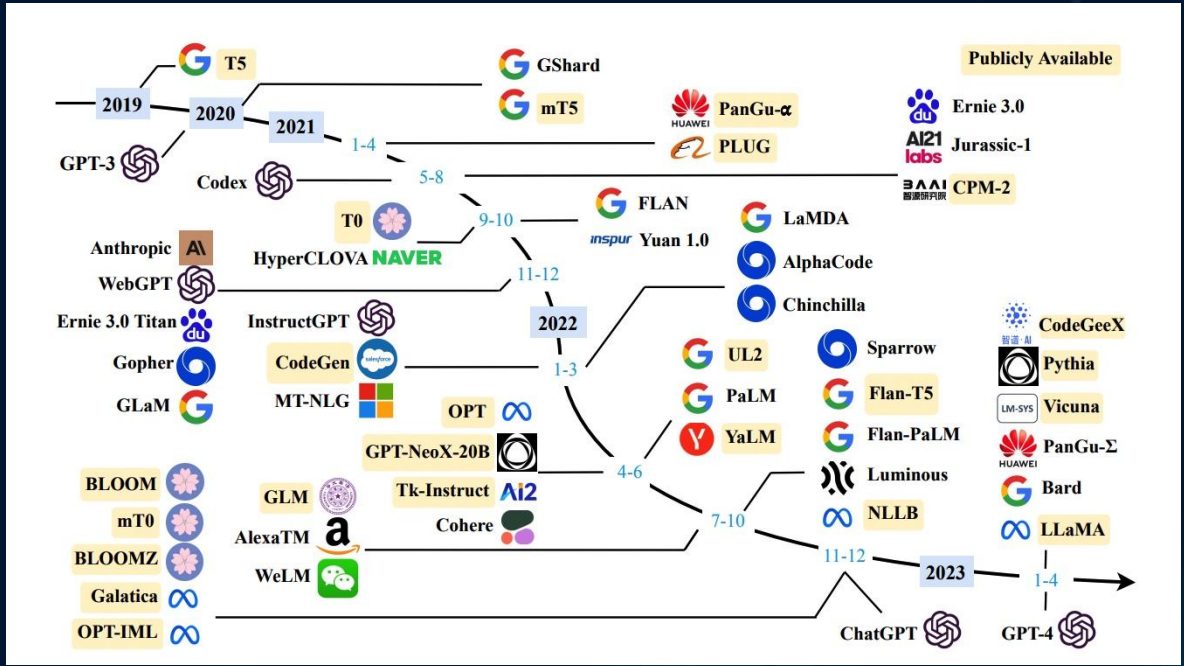


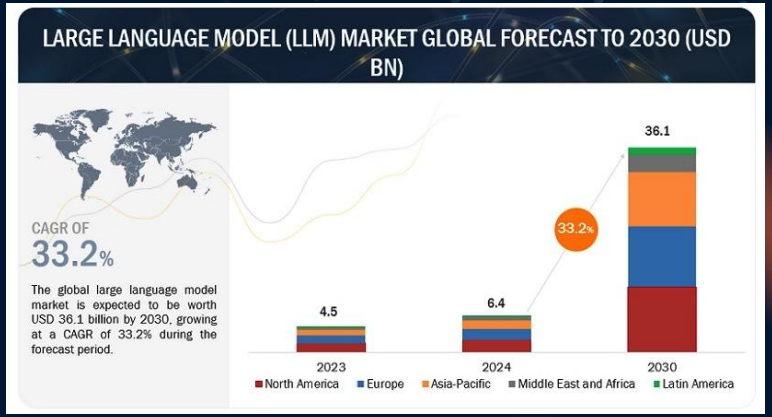
Fig: A timeline of existing large language models (having a size larger than 10B) in recent years.
Source: (Source: A Survey of Large Language Models)

SILO AI

LLMs in Industries

LLMs In Industries

- Global LLM market expected to reach 36.1 billion by 2023 with a CAGR of 33.2 %.
- Major Factors:
  - Demand for automated content creation and curation
  - Demand for LLM in knowledge discovery and management
- Challenges:
  - High memory requirements
  - Training cost
  - Inference optimization



Source:
https://www.marketsandmarkets.com/Market-Reports/large-language-model-llm-market-102137956.html

Industries revolutionized by LLMs

- LLMs find applications in diverse domains
- LLMs expand the purview of AI across industries and businesses.

Source: https://datasciencedojo.com/blog/llm-use-cases-top-10/

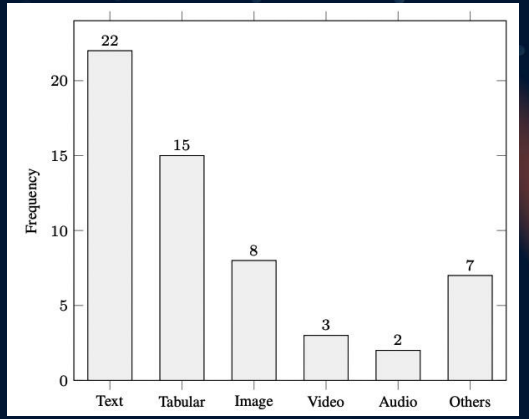
SILO AI

Applying LLM in Industries - Best Practices

- Identify the right use case
- Select the appropriate model
- Prepare and fine-tune the data
- Plan the integration with existing systems
- Monitor and evaluate performance
- Address ethical and privacy considerations
- Pay attention to scalability and maintenance
- Foster a culture of AI adoption

Data Modality Distribution

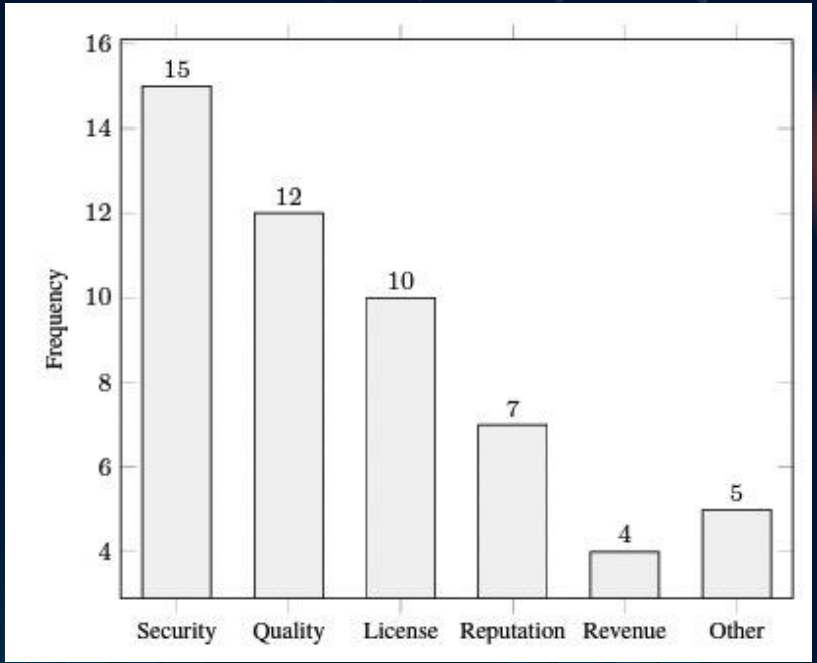
- Data modalities distribution for industrial application
  - Text, Tabular, Image, Video, Audio



Source: LLMs with Industrial Lens: Deciphering the Challenges and Prospects– A Survey

Risk of LLMs for Industrial Applications

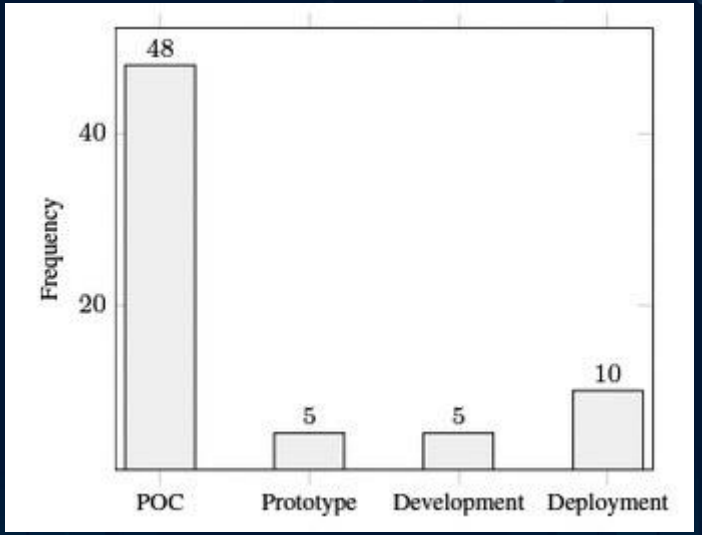
- Risks associated with LLMs for industrial application
  - Security, Quality, License, Reputation, Revenue



Source: LLMs with Industrial Lens: Deciphering the Challenges and Prospects– A Survey

State of LLMs in Industrial Applications

- Current state of the industrial applications utilizing the LLMs;
  - POC, Prototype, Development, Deployment



Source: LLMs with Industrial Lens: Deciphering the Challenges and Prospects– A Survey

SILO AI

LLMs Use Cases Across Industries (1/3)

- Customer Experience and Support
  - Chatbots
  - Personalized recommendation
- Banking and Finance
  - Financial analysis and research
  - Fraud detection
  - Risk assessment

SILO AI

LLMs Use Cases Across Industries (2/3)

- E-commerce and Retail
  - Product description and reviews
  - Inventory management and demand forecasting
- Healthcare
  - Clinical documentation automation
  - Patient assistant
  - Compliance management

LLMs Use Cases Across Industries (3/3)

- Cybersecurity
  - Threat detection and analysis
  - Incident response
- Marketing and advertising
  - Personalized marketing
  - Generating creative text

SILO AI

Custom LLMs

- Fine-tuning trains the LLM on task-specific or domain-specific data, thereby enhancing its performance in those areas.
  - Custom LLMs in Marketing
  - Custom LLMs in Healthcare
  - Custom LLMs in banking & finance
  - Custom LLMs in legal cases

SILO AI

Case-study Discussion

Case-study Discussion

- E-commerce
  - Building an Intelligent Shopping Assistant for customer.
  - Able to answer all the product information, reviews, comparison, product navigation questions.
  - Can suggest products only from the e-commerce website.
  - Browser plug-in.

Proposed Solution (1/4)

- Collect data (product information, e.g. e-comm website)
- Get embedding
- Store embedding in the vector db
- Fetch relevant context (Query + Context)
- Get response from LLM
- Evaluate the performance

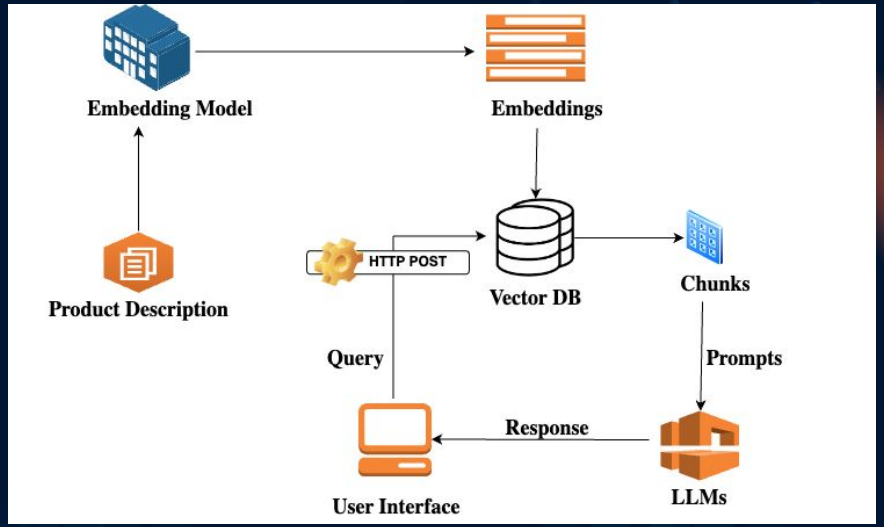


Fig: Overall Architecture

SILO AI

Proposed Solution (2/4)

- Data Source
  - Specific websites to scrape (csv, json, etc).
- Extracting Data
  - CSV, JSON data.
  - Web scraping using selenium.
  - FIX which data to store (product_id, product_name, description, reviews etc).
  - Bulk web scraping.
- Data Store
  - Store in JSON format (Mongo DB)
- Vector DB Store
  - Store embedding in DB.(select fields - product_decription, review etc)
  - Check for multiple collections,
  - Real time ingestion.

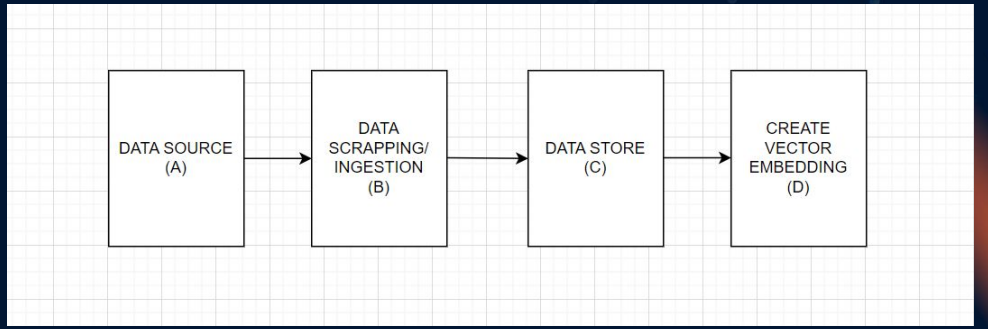


Fig: Data Pipeline

SILO AI

Proposed Solution (3/4)

- Fetch relevant context (query + context)
  - Simple Retrieval Augmented Generation (RAG) + Retriever.
  - Self Query RAG + Ensemble Retriever.
  - React Agent + Ensemble Retriever + Multi Query + Hyde
  - Langraph Agents
- Vector DB
  - Qdrant
  - Qdrant + real time ingestion
  - Qdrant + Redis (caching)
- Embedding Model
  - BGE embedding
  - Togather.ai
  - BGE embedding + Cohere + openai

SILO AI

Proposed Solution (4/4)

- LLM
  - Mistral 7x8b togather.ai API
  - Mistral 7x8b togather.ai API + openai GPT 3.5 + Claude
  - Open source LLM + In House deployment
  - LLM optimization
- Store Data
  - JSON store locally
  - JSON store local + SQLite
  - MongoDB
  - Data Lake
- Client
  - Gradio
  - React based UI

SILO AI

Open Question

SILO AI

Can large language models replace humans?

- People in industries are going to lose their jobs ?
  - adopt a human-in-the-loop approach rather than replacing everyone with an LLM
  - LLMs often referred to as black boxes and important to continually evaluate and test the results of LLMs.
  - subject matter experts (SMEs) in their respective fields and are the best candidates to judge the quality of an LLM's output.

Conclusion and Future

- LLMs highlight remarkable proficiency in comprehending, generating, and enhancing human-like text.
- The intersection of advanced Artificial Intelligence (AI) techniques and the Industrial Internet of Things (IIoT) is reshaping the world.

Source:https://medium.com/@abhi080497/industry-4-5-the-fusion-of-large-language-models-llms-and-industrial-iot-iiot-9b22ecf67730

SILO AI

References

[1] 10 practical applications of large language models in business
[2] Large Language Models Use Cases Across Various Industries
[3] LLMs with Industrial Lens: Deciphering the Challenges and Prospects – A Survey
[4] How Custom LLMs are Revolutionizing Industries
[5] LLM Market
[6] LLM Use-Cases: 10 industries revolutionized by large language models
[7] Industry 4.5: The Fusion of Large Language Models (LLMs) and Industrial IoT (IIoT)
[8] How to Implement Large Language Models in Your Business

SILO AI

Thank You