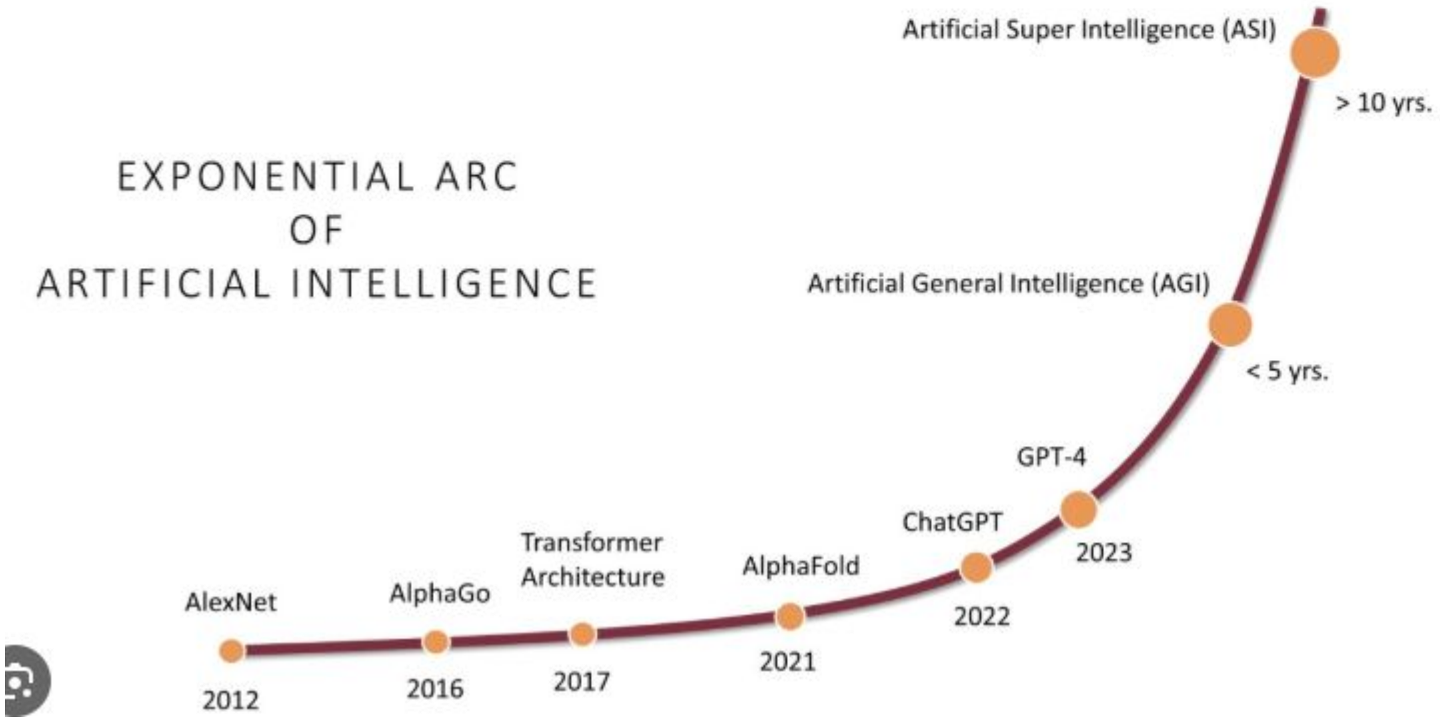
Introduction to Observability & Site Reliability Engineering (SRE) in Large Language Models (LLMs)

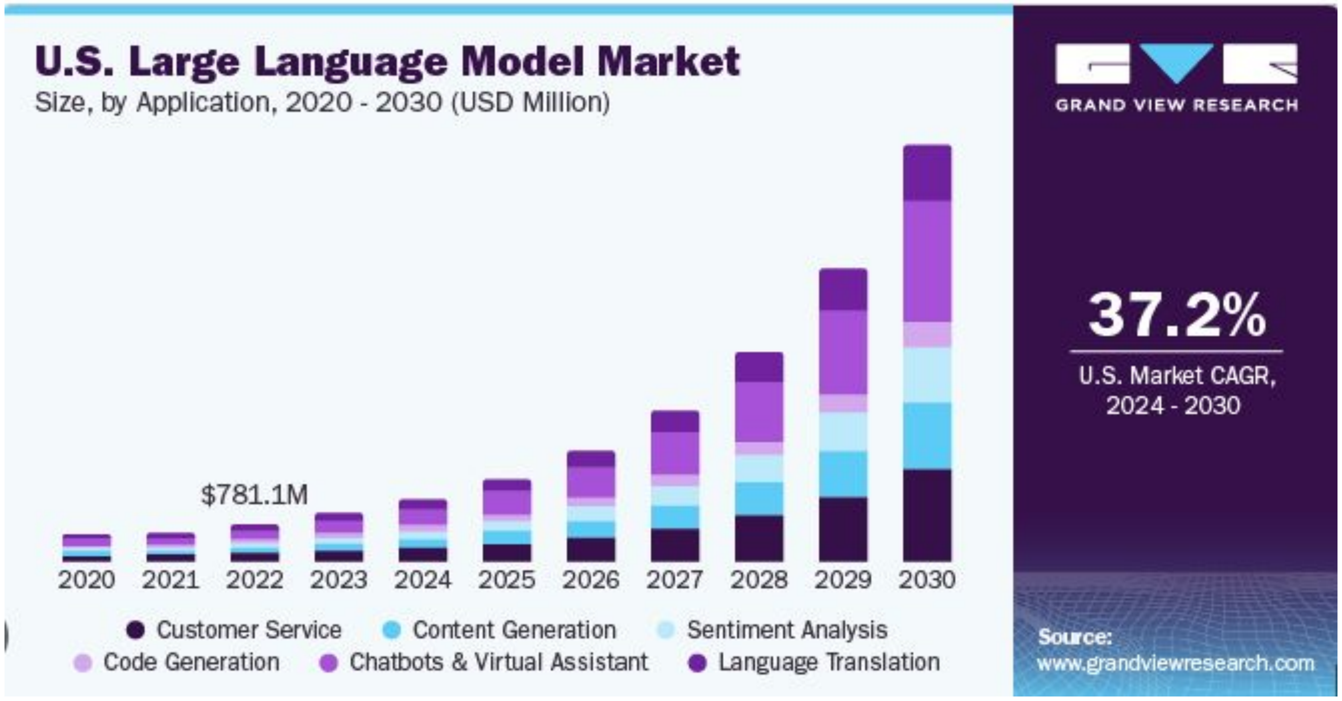
Presented by Ankush Sharma
https://www.linkedin.com/in/ankushsharmaa

Agenda

AI Journey

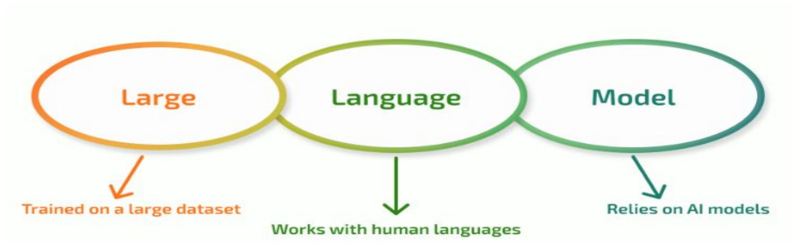

EXPONENTIAL ARC
OF
ARTIFICIAL INTELLIGENCE
AlexNet
2012
AlphaGo
2016
Transformer Architecture
2017
AlphaFold
2021
ChatGPT
2022
GPT-4
2023
Artificial General Intelligence (AGI)
< 5 yrs.
Artificial Super Intelligence (ASI)
> 10 yrs.

LLM market trend



Introduction to Large Language Models (LLMs)

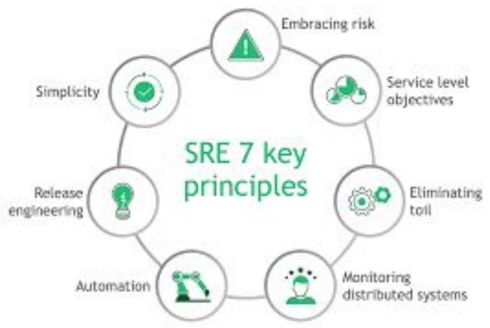
- Large Language Models (LLMs) are advanced AI systems trained on vast datasets to understand and generate human-like text.



- Applications: content generation, customer support, research, and coding assistance.
- Examples include GPT-4, BERT, and PaLM.
- LLMs have transformed industries such as healthcare, education, and software development.

Overview of Site Reliability Engineering (SRE)

- SRE ensures reliability, scalability, and efficiency in software systems.
- Key practices include proactive monitoring, automation, and incident response.
- Principles like Service Level Objectives (SLOs) and error budgets guide SRE teams.



- Focused on balancing reliability with innovation velocity.

Observability in AI Systems

- Observability ensures that system behavior is measurable and understandable.
- Metrics: Quantitative values like CPU usage, memory utilization, and API response times.
- Logs: Detailed, time-stamped records of system events, critical for debugging.
- Traces: Visualize the flow of requests through various services to identify bottlenecks.

Challenges in Observing LLMs

- LLMs operate on enormous datasets, creating challenges in performance monitoring.
- Dynamic behavior: LLMs adapt based on inputs, making their responses non-deterministic.
- Transparency: Black-box nature of LLMs complicates understanding their internal decision processes.
- Scalability: Monitoring and managing LLMs require robust and scalable infrastructure.

Implementing Observability in LLMs

- Step 1: Define Key Metrics

  - Monitor prediction latency, throughput, and error rates.

- Step 2: Integrate Advanced Logging

  - Collect logs for inputs, outputs, and model decisions.

- Step 3: Use Distributed Tracing

  - Map end-to-end interactions for pinpointing failures.

- Step 4: Deploy Real-Time Dashboards

  - Use tools like Grafana to visualize performance metrics.

Role of SRE in Managing LLMs

- SRE's responsibilities in LLMs include:

  - Ensuring 24/7 availability through automated failover systems.

  - Developing runbooks for common incident scenarios.
- Automation is key:

  - Automate scaling up resources during high demand periods.
- Collaboration with AI Teams:

  - Provide feedback on model performance for iterative improvements.

Case Studies and Applications

- Case Study 1: AI Chatbots

  - Monitoring response accuracy and latency to enhance user experience.
- Case Study 2: Healthcare AI

  - Ensuring compliance with medical guidelines and reducing false positives.
- Case Study 3: Fraud Detection

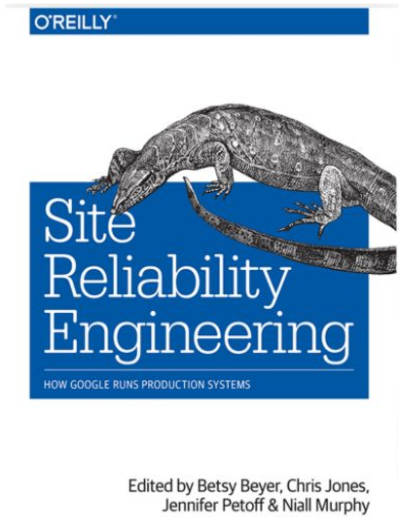
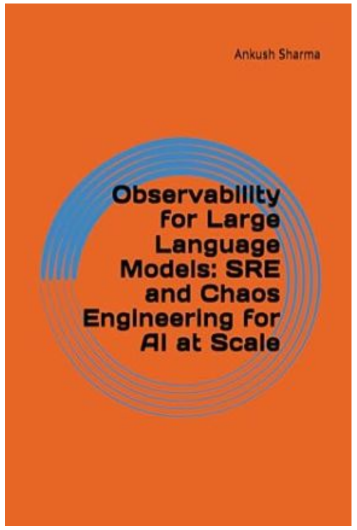
  - Continuous monitoring of prediction accuracy in identifying fraudulent transactions.

Conclusion

- Observability and SRE are vital for ensuring the reliability and scalability of Large Language Models (LLMs).
- As AI systems evolve, adopting innovative practices and prioritizing reliability practices will be key to building resilient and trustworthy AI systems that drive value across industries.

Thank You

- Recommended Reading:
  - https://www.amazon.in/Observability-Large-Language-Models-Engineering-ebook/dp/B0DJSR65TR
  - https://sre.google/books/


O'REILLY
Site Reliability Engineering
HOW GOOGLE RUNS PRODUCTION SYSTEMS
Edited by Betsy Beyer, Chris Jones, Jennifer Petoff & Niall Murphy

Ankush Sharma
Observability for Large Language Models: SRE and Chaos Engineering for AI at Scale