



# LLMs from zero to hero

Build using open source





Large language models

Large language models

Large language models

Large language models



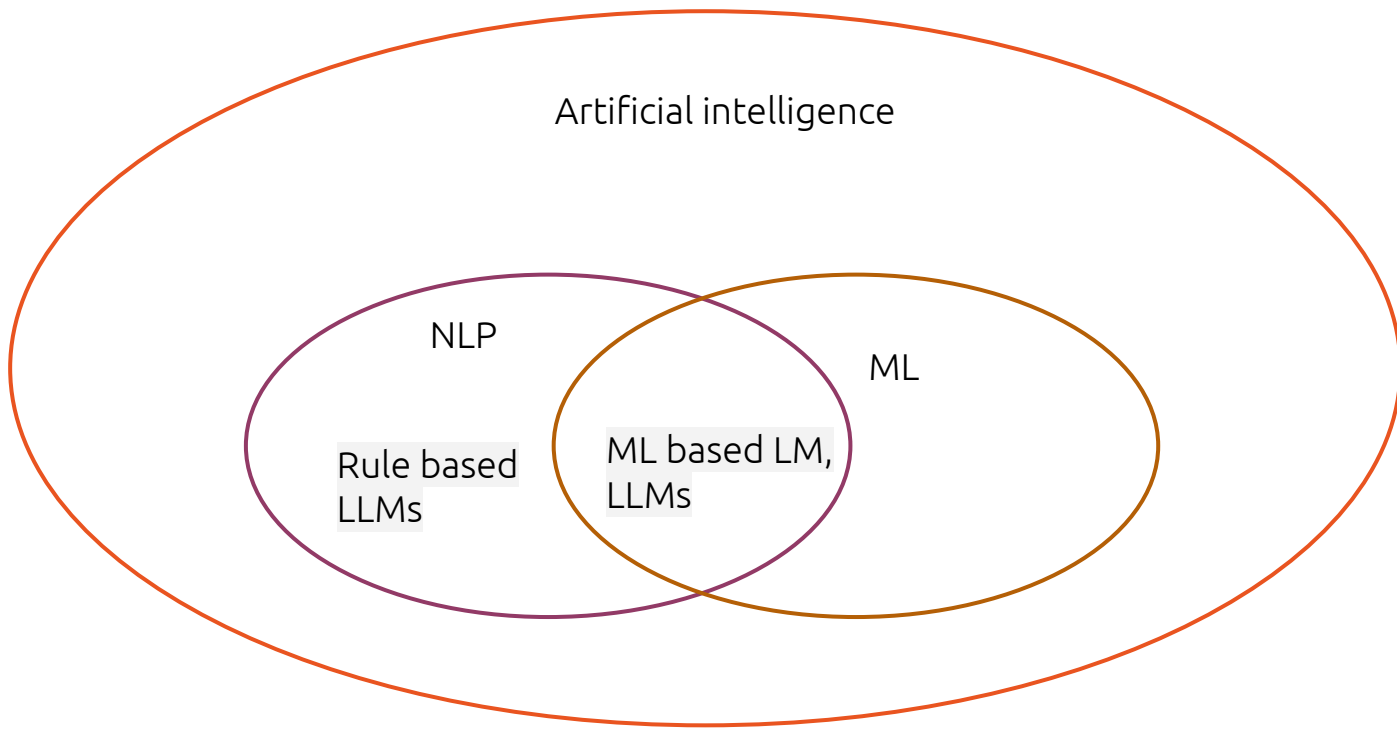
# Who needs LLMs?







# LLMs vs traditional ML





# Can we do it?

Training is resource heavy



# Can we do it?

Training is resource heavy

Distributed training, GPU sharing, fine tuning and p-tuning



# Can we do it?

Training is resource heavy

Distributed training, GPU sharing, fine tuning and p-tuning

Costs are high



# Can we do it?

Training is resource heavy

Distributed training, GPU sharing, fine tuning and p-tuning

Costs are high

Open source, automation, hybrid-cloud



# Can we do it?

Training is resource heavy

Distributed training, GPU sharing, fine tuning and p-tuning

Costs are high

Open source, automation, hybrid-cloud

It is not reproducible



# Can we do it?

Training is resource heavy

Distributed training, GPU sharing, fine tuning and p-tuning

Costs are high

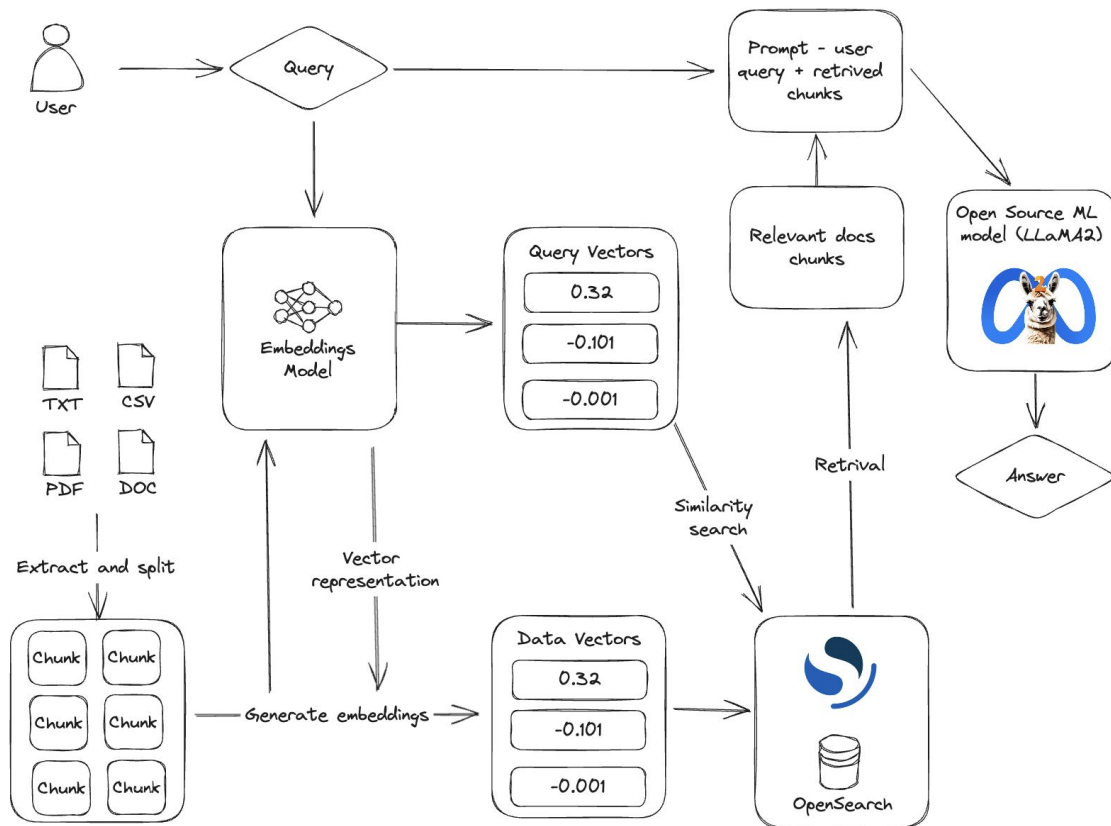
Open source, automation, hybrid-cloud

It is not reproducible

MLOps pipelines, Kubeflow, MLflow



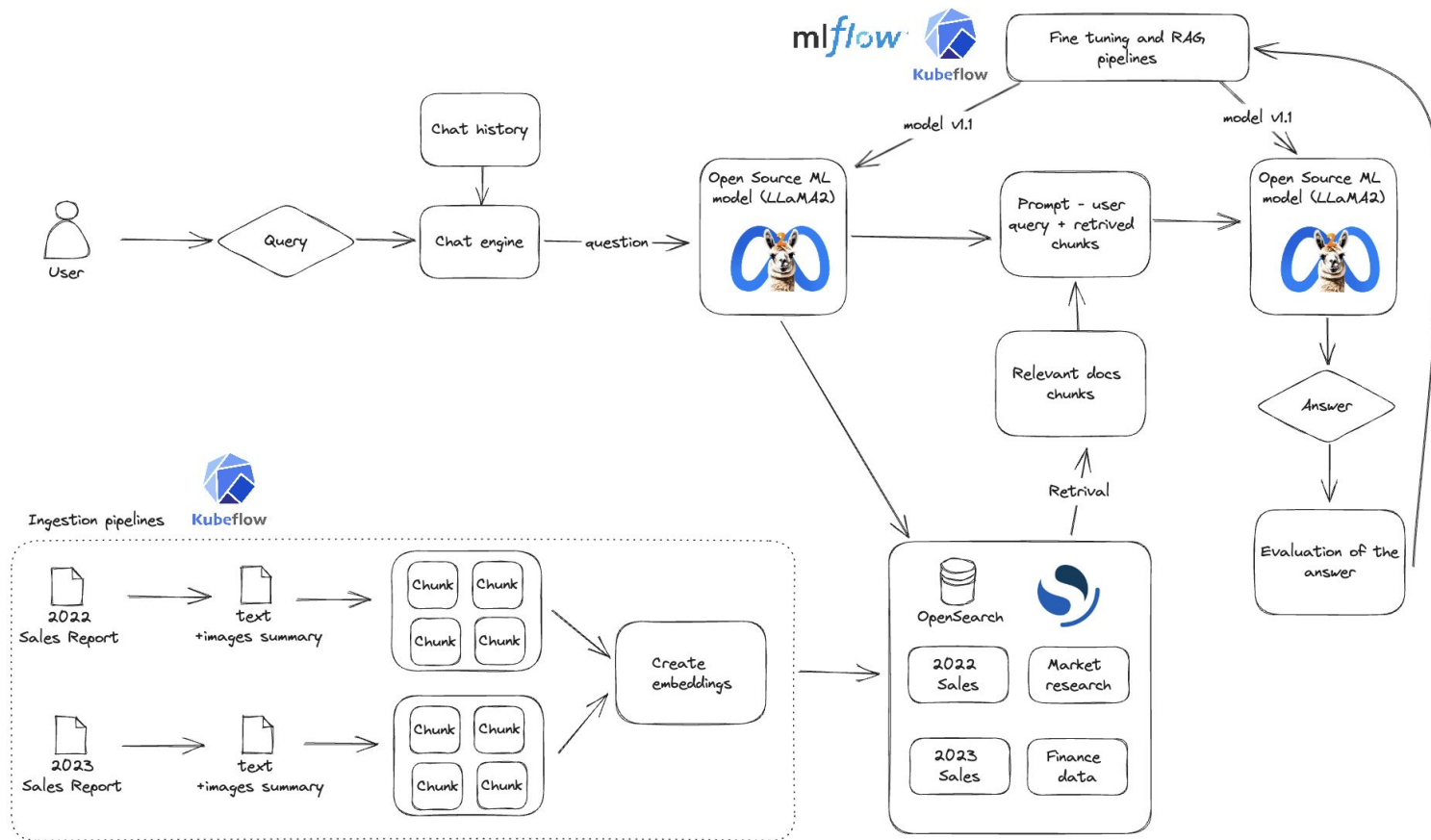
# Basic architecture





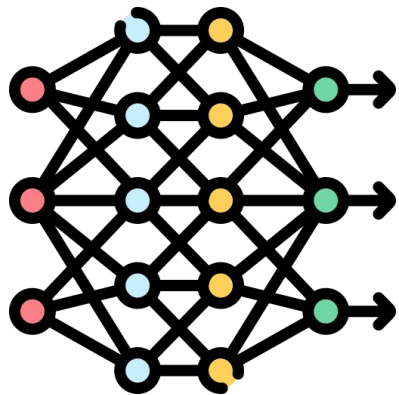


# Real world example





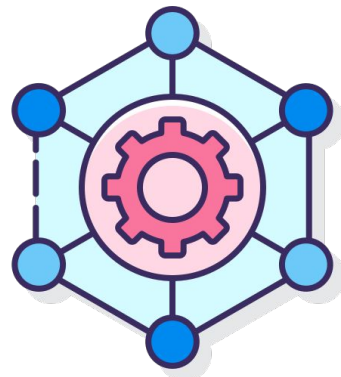
# LLMs projects with open source



Models





Tools



Frameworks



# LLMs projects with open source

	
Easy to get started	Security and privacy concerns
Cost-efficiency	Lack of support
Customisation	Stepper learning curve
Community support	Documentation
Ability to contribute	Bug fixing can take longer



# To wrap it up...

Get started with a use case



Scale with integrated solutions

- Open source
- Reproduce experiments
- Runs on any environment
- Enterprise support



# Innovate at speed with open source AI

**Andreea Munteanu**

AI Product Manager



**Maciej Mazur**

AI Director



Booth **A24!**



# Canonical AI Roadshow

Innovate at speed with  
open source AI

Meet Canonical experts, discover  
generative AI capabilities for  
enterprises using open-source, and  
ask any related AI/ML question.

## Open Source Summit

19-21 September, Bilbao, Spain

## Open Source Experience

6-7 December 2023

Paris, France

## World Summit AI

11-12 October 2023

Amsterdam, Netherlands

## Ubuntu Summit

3-5 November

Riga, Latvia

## Kubecon NA + DOK Day

6-9 November

Chicago, USA

## LATAM AI

21-22 November 2023

São Paulo, Brazil

## MLOps World

25-26 October 2023

Austin, USA

## GITEX Technology

16-20 October 2023

Dubai, UAE

