AZURE AI: BUILDING TRUST IN THE GENERATIVE AI ERA

Mr. Sunil Kumar

About me

https://www.linkedin.com/in/sunilkumardigi/

- Cloud Instructor
- Previously worked at GlobalLogic, AKTU (State Technical University), U.P., India
- Tech enthusiast
- Tech Community Builder
- YouTube: https://www.youtube.com/@sunilkumarOnCloud
- Medium blog: https://medium.com/@suniel.vns
- Topmate: https://topmate.io/sunilkumar7
- Twitter: @techwithsunil
- Toastmasters International member
- Runner

Table of content

- What is Generative AI?
- What are Language Models?
- Using Language Models
- LLM Vs SLM
- MICROSOFT AI / RESPONSIBLE AI Principles
- Responsible Generative AI
  - Plan
  - Identify
  - Mitigate
  - Operate
- Demo – Implementation of content filters in Azure AI Studio

Generative AI

- Generative AI describes a category of capabilities within AI that create original content.

Ex. Chat applications, Microsoft copilot

- Gen AI applications take in natural language input and return appropriate responses in a variety of formats such as natural language, images , code and more.
- Example of some prompts:

‘’write a cover letter for a person with a bachelor’s degree in cloud computing”,

“create a logo for a book seller business”,

“write python code for addition of two numbers”

Language models

- Generative AI applications are powered by language models, which are specialized type of machine learning model that you can use to perform natural language processing tasks including-

-determine sentiment

-summarizing text

-generating new natural language etc.

Using language models

- GPT- Generative Pre-trained Transformer
- DALL-E model for image generation
- OpenAI
- Huggingface
- Mistral
- Meta and other

LLM Vs SLM

LLM

- Trained with vast quantities of text that represent wide range of general subject matter
- When trained , LLM’s have many billions of parameters
- Able to exhibit comprehensive language generation capabilities
- Fine-tuning the model with additional data to customize its subject expertise can be time consuming and expensive

SLM

- SLM’s are trained with smaller, more subject focused datasets
- Fewer parameters than LLM’s
- This focused vocabulary makes them very effective in specific conversational topics
- Fine-tuning can be potentially be less time-consuming and expensive

Microsoft AI / Responsible AI Principles

- **Fairness -**AI systems should treat all people fairly
- **Reliability and safety -**AI systems should perform reliably and safely
- **Privacy and security -**AI systems should be secure and respect privacy
- **Inclusiveness** -AI systems should empower everyone and engage people
- **Transparency** -AI systems should be understandable
- **Accountability** -People should be accountable for AI systems

Plan a responsible generative AI solution

Four-stage process to develop and implement a plan for responsible AI when using generative models

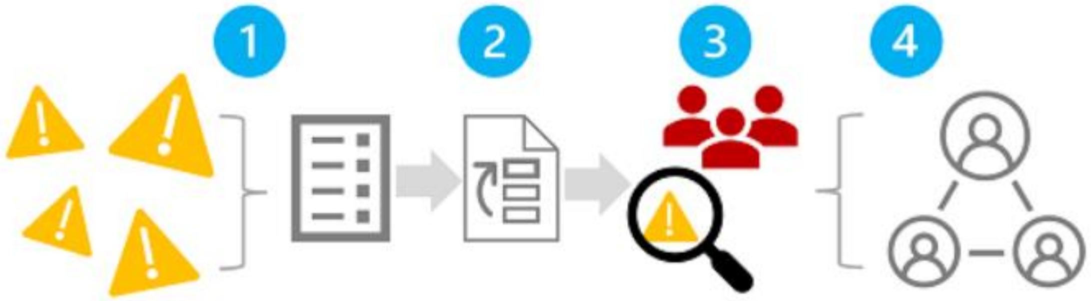
1. Identify potential harms

2. Measure the presence of theses harms

3. Mitigate the harms at multiple layers

4. Operate

Step1: Identify potential harms



There are four steps in this stage:

i. Identify potential harms

(offensive, discriminatory, factual inaccurate, encourages or supports illegal or unethical behaviour)

ii. Prioritize identified harms

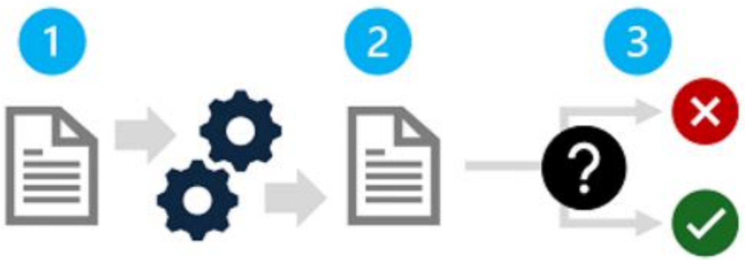
(likelihood of its occurrence & the resulting level of impact)

iii. Test and verify the prioritized harms- Red team

iv. Document and share details of harms

Image credit: Microsoft learn

Step2: Measure potential harms



It consists three steps:

i. Prepare a diverse selection of input prompts that are likely to result in each potential harms that you have documented for the system

ii. Submit the prompts to the system and retrieve the generated output

iii. Apply pre-defined criteria to evaluate the output and categorize it according to the level of potential harm it contains

Example category of **“harmful”** or **“not harmful”**

Image credit: Microsoft learn

Step3: Mitigate potential harms

A layered approach

1. Model layer
2. Safety system

Ex. Content filters in Azure AI foundry

3. Metaprompt and grounding
4. User experience

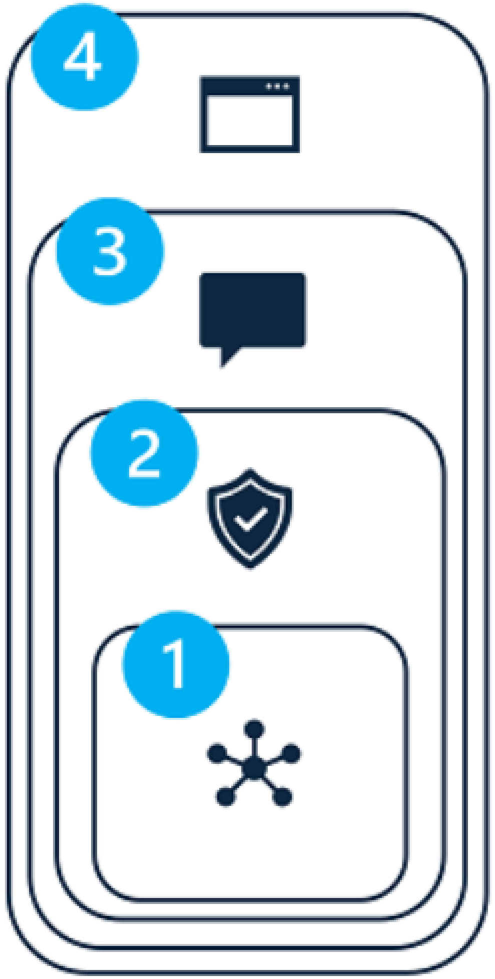


Image credit: Microsoft learn

Step4: Operate a responsible Gen AI solution

Common compliance review include:

- Legal
- Privacy
- Security
- Accessibility

Release and operate the solution:

- Devise a phased delivery plan
- Create an incident response plan
- Create a rollback plan

References

- https://learn.microsoft.com/
- https://www.microsoft.com/en-us/ai/responsible-ai#tools

Demo

- https://microsoftlearning.github.io/mslearn-ai-studio/Instructions/06-Explore-content-filters.html