

## 1.2 IA Generativa, Modelos de IA y Agentes

Los **LLMs (Large Language Models)** son modelos de lenguaje de gran tamaño (como GPT, BERT o LLaMA) entrenados con enormes corpus de texto para realizar tareas como generación de texto, traducción, clasificación o respuesta a preguntas. Plataformas como **Hugging Face** ofrecen bibliotecas como *Transformers*, que permiten gestionar y utilizar estos modelos en diferentes modalidades. Aunque originalmente estaban orientados al procesamiento de texto, hoy en día muchos de estos modelos pueden trabajar también con **imágenes, audio o vídeo**, dando lugar a sistemas **multimodales** capaces de combinar distintos tipos de información.

Además de los LLMs centrados en texto, existen **modelos de visión** basados en redes neuronales profundas, como las **CNN (Convolutional Neural Networks)** o los **Vision Transformers**, que se utilizan para tareas como clasificación de imágenes, detección de objetos o segmentación visual. También han aparecido modelos generativos como **Stable Diffusion**, que permiten crear imágenes a partir de descripciones textuales. En paralelo, se han desarrollado **modelos multimodales** (por ejemplo **CLIP** o **LLaVA**) capaces de relacionar texto e imagen, permitiendo tareas como describir una fotografía, responder preguntas sobre una imagen o generar contenido visual a partir de instrucciones en lenguaje natural.

Muchos de estos sistemas se construyen a partir de lo que se denomina un **modelo fundacional (foundation model)**. Se trata de grandes redes neuronales entrenadas previamente con enormes cantidades de datos generales —como los modelos GPT de OpenAI o BERT de Google— que aprenden patrones lingüísticos y conceptuales amplios. Estos modelos proporcionan una **capacidad general de comprensión y generación**, que posteriormente puede adaptarse a tareas específicas mediante técnicas como el **fine-tuning** o el ajuste mediante instrucciones. Por ejemplo, un modelo fundacional entrenado con grandes cantidades de texto puede posteriormente especializarse para responder preguntas médicas, analizar documentos legales o ayudar en tareas educativas.

En los últimos años ha aparecido además un **nivel más avanzado de uso de los LLMs**: los **agentes de inteligencia artificial**. Un agente es un sistema que utiliza un modelo de lenguaje como núcleo de razonamiento, pero que además puede **tomar decisiones, planificar pasos y utilizar herramientas externas** (bases de datos, buscadores, programas o APIs) para resolver tareas complejas. En lugar de limitarse a generar una respuesta, el modelo puede descomponer un problema, ejecutar acciones y combinar distintos recursos para alcanzar un objetivo.

El desarrollo más reciente en este ámbito es la **orquestación de agentes**, donde varios agentes especializados colaboran entre sí dentro de un mismo sistema. En este enfoque, cada agente puede encargarse de una función concreta —por ejemplo, búsqueda de información, análisis de datos, generación de texto o verificación de resultados— y un sistema de coordinación organiza su interacción. Este paradigma está dando lugar a nuevas arquitecturas de software basadas en **equipos de agentes cooperativos**, capaces de automatizar procesos complejos y construir aplicaciones inteligentes más avanzadas.

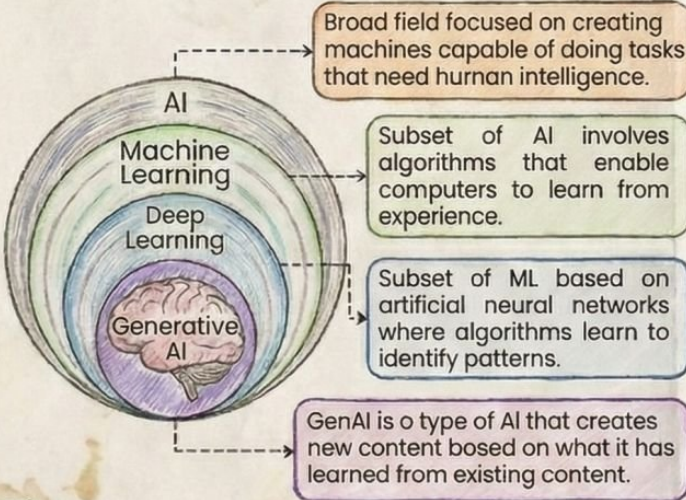
En conjunto, la evolución de la inteligencia artificial ha pasado desde modelos especializados en tareas concretas hacia **modelos fundacionales generales**, después hacia **LLMs capaces de interactuar con múltiples modalidades de datos**, y finalmente hacia **sistemas de agentes y orquestación de agentes**, que representan actualmente una de las fronteras más activas de investigación y desarrollo en IA. Estos enfoques permiten que los modelos no solo generen contenido, sino que también **razonen, planifiquen y colaboren para resolver problemas reales en distintos ámbitos científicos, educativos y profesionales**.



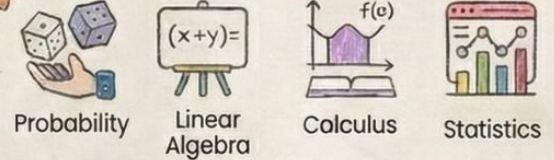
# Generative AI Learning Roadmap

Brij Kishore Pandey

## 1. What is Generative AI



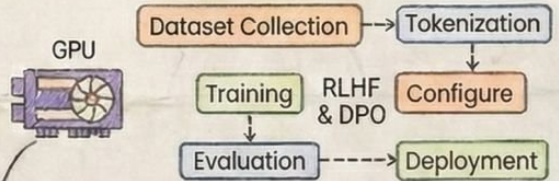
## 2. Important Concepts



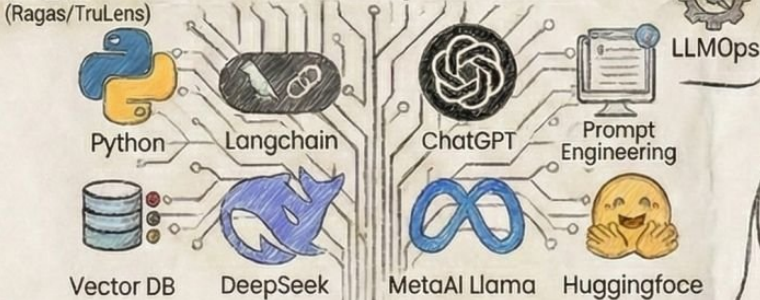
## 3. Foundation Models



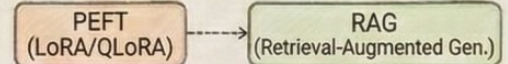
## 5. Training a Foundation Model



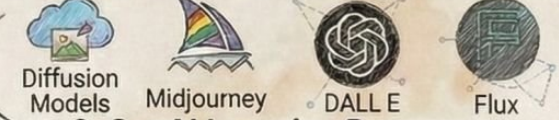
## 4. GenAI Development Stack



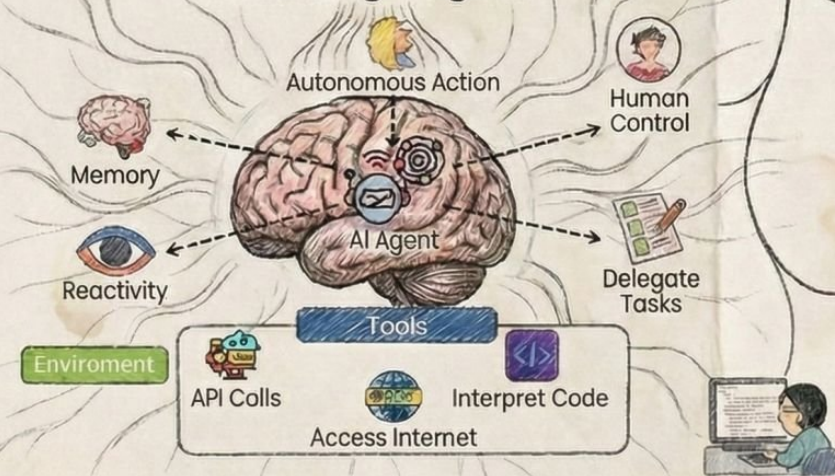
## 5a. Fine-Tuning & RAG Patterns



## 7. GenAI Models for Computer Vision



## 6. Building AI Agents



## 8. GenAI Learning Resources



