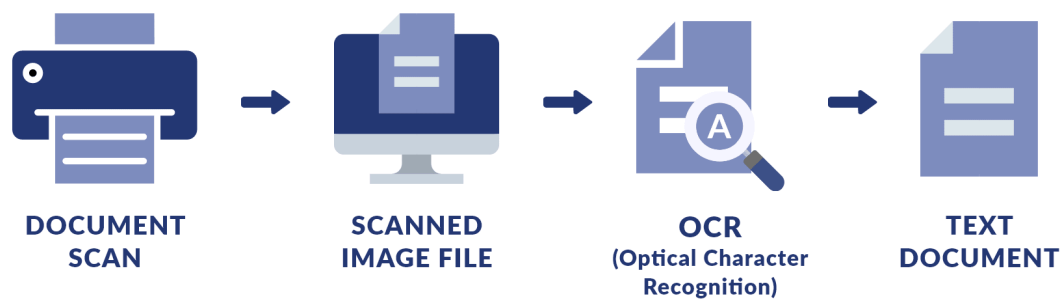# LEARNING GEN AI

-   BY SANJANA MCS

OCR AND PROMPT ENGINEERING

What is OCR?

Optical Character Recognition (OCR) is a technology that enables computers to extract and convert text from images, scanned documents, or printed materials into an editable format. This allows users to digitize physical documents and make text searchable, editable, and more accessible.

How Does OCR Work?

OCR operates through three key steps:

1. **Image Preprocessing** – The software enhances the image by removing noise, adjusting contrast, and sharpening edges to improve readability.
2. **Character Recognition** – The OCR algorithm scans the image, identifies characters (letters, numbers, symbols), and compares them against stored patterns or trained models.
3. **Text Conversion** – The recognized characters are compiled into structured text, which can be edited, copied, or processed further.

Why is OCR Important?

OCR has numerous applications, including:

- **Digitization of Documents** – Converts printed books, invoices, and records into searchable digital formats.
- **Accessibility Enhancement** – Assists visually impaired users by enabling text-to-speech conversion.
- **Data Extraction** – Automates the extraction of key information from images, receipts, or scanned forms.

Popular OCR Tools and Libraries

- **Tesseract OCR** – Open-source and widely used for various OCR applications.
- **EasyOCR** – Supports multiple languages and is easy to implement.
- **Google Vision API** – Cloud-based OCR with powerful AI-driven recognition capabilities.
- **Amazon Textract** – Optimized for large-scale document processing with deep learning capabilities.

Challenges in OCR

- **Handwriting Recognition** – Printed text is easier to process than handwritten text, which varies in style and legibility.
- **Image Quality** – Blurred, noisy, or low-resolution images reduce OCR accuracy.
- **Multilingual Processing** – Different fonts and languages require specialized models and training for better accuracy.

What is Prompt Engineering?

Prompt Engineering is the practice of designing effective prompts to guide AI models (like ChatGPT) in generating accurate and relevant responses. By carefully structuring input instructions, users can optimize AI-generated outputs for various applications.

Why is Prompt Engineering Important?

- **Enhances AI Response Quality** – Well-crafted prompts yield clearer and more useful answers.
- **Improves Efficiency** – Reduces trial and error when interacting with AI.
- **Customizes Output** – Allows users to tailor responses based on specific needs and contexts.
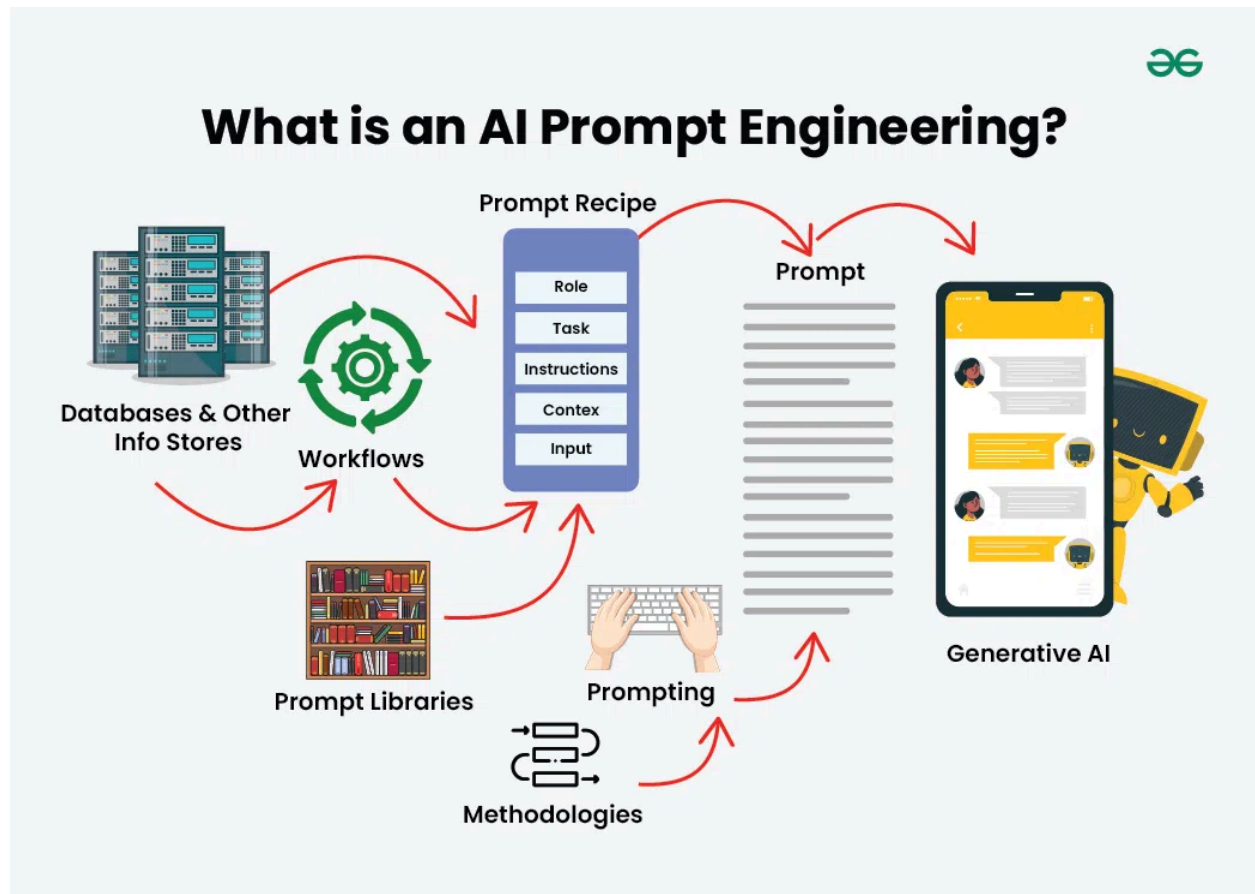
Types of Prompting Techniques

Simple Prompting – Asking direct questions.
Example: "What is the capital of Japan?"

Instruction-Based Prompting – Providing clear instructions for better output.
Example: "Summarize the history of the Internet in 100 words."

Few-Shot Prompting – Giving AI multiple examples before asking a question.
Q: What is the capital of Germany?
A: Berlin.
Q: What is the capital of Canada?
A:

Step-by-Step Prompting – Encouraging AI to explain its reasoning.
Example: "Solve 125 × 3 step by step."

Chain-of-Thought Prompting – Encouraging AI to explain complex reasoning in steps.
Example: "Explain how climate change affects sea levels step by step."

Zero-Shot Prompting – Asking AI to generate an answer without prior examples.
Example: "Translate 'Hello, how are you?' into French."

Persona-Based Prompting – Instructing AI to respond as a specific personality or expert.
Example: "Act as a historian and explain the significance of the Renaissance."

Multi-Turn Prompting – Keeping context across multiple interactions.
User: "Who discovered gravity?"
AI: "Isaac Newton."
User: "When did he discover it?"
AI: "In 1687."

Context-Aware Prompting – Providing AI with relevant background information.
Example: "Given that I have a vegan diet, suggest a high-protein meal plan."

Conversational Prompting – Making AI interact naturally in a dialogue.
Example: "You are my travel assistant. Suggest a one-week itinerary for Italy."

Common Applications of Prompt Engineering

- **Content Generation** – Writing blog posts, summaries, and reports.
- **Programming Assistance** – Generating or debugging code.
- **Customer Support** – Automating chatbot interactions.
- **Education & Research** – Creating structured learning materials and explanations.

Challenges in Prompt Engineering

- **Ambiguous Prompts** – Poorly worded prompts lead to inaccurate or vague responses.
- **Bias in AI Responses** – AI models may generate biased or contextually incorrect outputs.
- **Iterative Refinement** – Crafting an optimal prompt often requires multiple adjustments and testing.

Integrating OCR with Prompt Engineering

By combining OCR and prompt engineering, users can:

- Extract text from scanned images and process it using AI.
- Summarize, translate, or analyze extracted text for insights.
- Automate workflows such as invoice processing, legal document analysis, or historical record digitization.

Example Workflow

1. **Image Capture** – A user takes a picture of a document or handwritten note.
2. **OCR Processing** – The system extracts text from the image.
3. **AI Interaction** – A well-structured prompt guides AI to summarize, translate, or interpret the text.

Conclusion

OCR and prompt engineering are transformative technologies that bridge the gap between physical and digital text processing. OCR enables efficient text extraction, while prompt engineering enhances AI interactions, making text-based AI systems more intelligent and responsive. Understanding and leveraging these concepts can unlock new possibilities in automation, accessibility, and artificial intelligence-driven solutions.