

# NLP/ASR

Unit 5: Advanced Topics in  
ASR/NLP



# 5.4.2

## State-of-the-art Models

Zero-shot and few-shot learning  
for QA

# Few-Shot & Zero-Shot

- Few-Shot Learning:
  - Aims to learn from a very small number of labeled examples
  - Leverages prior knowledge from pre-training on broader tasks
  - Better at adapting to new tasks with minimal data
- Zero-Shot Learning:
  - Adapts to a new task without any explicit training examples
  - Relies heavily on pre-trained models' (ex: LLMs) ability to understand language and extract information from prompts and instructions

# Why Zero-Shot and Few-Shot Learning?

- Reduced data dependency: Addresses scenarios where obtaining or labeling data is expensive, time-consuming, or impossible
- Faster adaptation: Ideal for quick deployment of models to new domains or tasks
- Handling rare or unseen events: Useful when dealing with infrequent events for which a large dataset would be unavailable

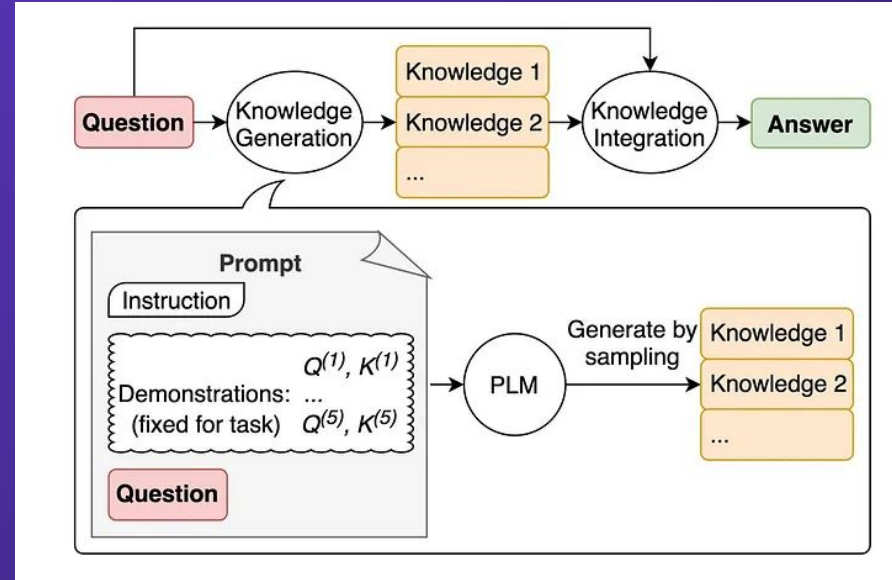
# Large Language Models (LLMs)

- Large Language Models (LLMs) like GPT-3 and T5 form the backbone of zero/few-shot learning
- Pre-trained on massive text corpora, LLMs acquire rich linguistic representations and implicit knowledge



# Zero-Shot QA

- Task Reframing: Questions are transformed into prompts compatible with the LLM's pre-training format (e.g., cloze style, text completion)
- Direct Answer Generation: The LLM processes the prompt and attempts to generate the answer directly
- No Explicit Training: There's no fine-tuning of the LLM on question-answering examples



# Few-Shot QA

- **Prompting and Demonstration:** A few labeled examples (questions, contexts, answers) are used to create prompts or templates
- **Targeted Fine-Tuning:** The LLM's weights are slightly adjusted to better align it with the specific QA task structure and domain
- **Improved Adaptability:** The fine-tuned model becomes more effective at handling similar questions

# Zero/Few-Shot QA Techniques

- Prompt Engineering: Carefully designing prompts that best leverage the LLM's capabilities
- Answer Verification: Adding mechanisms to assess the quality and reliability of generated answers
- Retrieval Augmented Generation: Incorporating external knowledge bases or structured data to enhance LLM reasoning

