1. Prompt:

   a. Chat GPT, Bing Chat, Bard, Llama (not sure if it has a prompt or hosted api)

   b. DallE, MidJourney

2. Code Assistants: AWS Code Wisper, GitHub CoPilot, VS Intellisense (a bit of ai thr as well)

3. How does Prompting tech to integrate with clouds, Azure, AWS, GCP mainly

4. What else can prompting an LLM do...

**Prompt:**

Prompt is the text we input to the AI models when interfacing with them.

AI Prompt Writing (or Engineering) is the process of creating input (usually text) instructing the Generative AI to generate the desired response.

Different types of prompts:

            Instructions

            Questions

            Data

            Examples

**a. Chat GPT, Bing Chat, Bard, Llama (not sure if it has a prompt or hosted api)**

**Chat GPT:**

**What is ChatGPT?**

ChatGPT is a natural language processing tool driven by AI technology that allows you to have human-like conversations and much more with the chatbot. The language model can answer questions and assist you with tasks, such as composing emails, essays, and code.

**Who made ChatGPT?**

ChatGPT was created by OpenAI, an AI and research company. The company launched ChatGPT on November 30, 2022.

**What are ChatGPT's limitations?**

Despite looking very impressive, ChatGPT still has limitations. Such limitations include the inability to answer questions that are worded in a specific way, as it requires rewording to understand the input question. A bigger limitation is a lack of quality in the responses it delivers -- which can sometimes be plausible-sounding but make no practical sense or can be excessively verbose.

**Bing Chat:**

**What is Bing Chat?**

Bing Chat launched in February 2023, a new feature for Bing's search engine that let you talk to its AI chatbot rather than simply filling out search queries.

The sytem is conversational and basically let Microsoft steal a march on Google, despite the latter's normally complete dominance of the online search market.

Bing Chat sounds like it's distinct from ChatGPT, and in practical terms it is, but it's powered by the same model from OpenAI and therefore operates in much the same way, although we'd say it's easier to access and use for normal people.

**Is Bing Chat better than ChatGPT?**

This is a slightly complicated question given that ChatGPT's technology (from OpenAI) effectively underpins and powers Bing Chat, but we still feel the answer might just be "yes".

This is because Bing Chat is so, so much easier to access than most parts of ChatGPT, and its most up-to-date features aren't locked behind a paywall at all, so you get to use the newest options and features without any real barriers.

The experience you have chatting to each might not feel drastically different all the time, but Bing Chat's useability and sophistication edges it ahead, too.

**Bard**

**What is Google Bard?**

Google Bard is an AI-powered chatbot tool designed by Google to simulate human conversations using natural language processing and machine learning. In addition to supplementing Google search, Bard can be integrated into websites, messaging platforms or applications to provide realistic, natural language responses to user questions.

**How does Google Bard work?**

Google Bard is built on the Pathways Language Model 2 (PaLM 2), a language model released in late 2022.

PaLM and the model that preceded it, Google's Language Model for Dialogue Applications (LaMDA) technology, are based on Transformer, Google's neural network architecture released in 2017. Because Google released Transformer as open source, it has been the framework for other generative AI tools, including the GPT-3 language model used in ChatGPT.

Bard is designed around search. It aims to allow for more natural language queries, rather than keywords, for search. Bard's AI is trained around natural-sounding conversational queries and responses. Instead of just giving a list of answers, it provides context to the responses. It is also designed to help with follow-up questions -- something new to search.

The initial version of Bard used a lighter-model version of LaMDA that requires less computing horsepower to scale to more concurrent users. The incorporation of the PaLM language model allows Bard to be more visual in its responses to user queries.

**Who can use Google Bard?**

Users must be 18 or older and have a personal Google account. Bard is available in 180 countries and territories and plans to expand.

**What are the limitations of Bard?**

Like all AI chatbots, Bard must learn and be trained to give the correct answer from inaccurate or misleading information, as was evident during its first demo. AI training is an endless, compute-intensive process since there is always new information to learn.

**Is Bard free?**

As of this writing, Google has given no indication that it will charge for use. Google has no history of charging customers for services -- its cloud business notwithstanding. The current assumption that Bard will be integrated into Google's basic search engine indicates that it will be freely available for use.

**LLaMA (Large Language Model Meta AI)**

**What is LLaMA?**

A New Paradigm in AI Research LLaMA AI (Large Language Model Meta AI) represents a new paradigm in artificial intelligence research. It focuses on developing AI systems that possess self-learning capabilities, adaptability to new environments, and contextual awareness.

Some of the challenges of LLaMA AI are:

1. It can generate factually wrong information or hallucinate.

2. It can be misused for malicious purposes such as synthetic biology, spam or disinformation.

3. It faces competition from other AI models such as OpenAI’s ChatGPT and Google Bard.

**DALL-E :**

**What is DALL-E ?**

1. DALL-E is a generative AI model developed by Open AI that can generate images and art from a text prompt.

2. It is a text-to-image AI system, or a CLIP system (connecting text to images).

3. The name is a combination of the artist Salvador Dali and the robot WALL-E from Pixar. DALL-E is an encoder-decoder model that can encode text into a machine input, process it, and then decode it into a visible image.

or

DALL-E (stylized as DALL·E) and DALL-E 2 are text-to-image models developed by OpenAI using deep learning methodologies to generate digital images from natural language descriptions, called "prompts". DALL-E was revealed by OpenAI in a blog post in January 2021, and uses a version of GPT-3 modified to generate images. In April 2022, OpenAI announced DALL-E 2, a successor designed to generate more realistic images at higher resolutions that "can combine concepts, attributes, and styles".

**Technical limitations**

DALL-E 2's language understanding has limits. It is sometimes unable to distinguish "A yellow book and a red vase" from "A red book and a yellow vase" or "A panda making latte art" from "Latte art of a panda". It generates images of "an astronaut riding a horse" when presented with the prompt "a horse riding an astronaut". It also fails to generate the correct images in a variety of circumstances. Requesting more than three objects, negation, numbers, and connected sentences may result in mistakes, and object features may appear on the wrong object.Additional limitations include handling text - which, even with legible lettering, almost invariably results in dream-like gibberish - and its limited capacity to address scientific information, such as astronomy or medical imagery.

**Midjourney:**

Midjourney is a generative artificial intelligence program and service created and hosted by San Francisco-based independent research lab Midjourney, Inc. Midjourney generates images from natural language descriptions, called "prompts", similar to OpenAI's DALL-E and Stable Diffusion.

The tool is currently in open beta, which it entered on July 12, 2022. The Midjourney team is led by David Holz, who co-founded Leap Motion. Holz told The Register in August 2022 that the company was already profitable. Users create artwork with Midjourney using Discord bot commands.

**Functionality**

Midjourney is currently only accessible through a Discord bot on their official Discord server, by direct messaging the bot, or by inviting the bot to a third party server. To generate images, users use the /imagine command and type in a prompt; the bot then returns a set of four images. Users may then choose which images they want to upscale. Midjourney is also working on a web interface.

**Uses**

Founder David Holz says he sees artists as customers, not competitors of Midjourney. Holz told The Register that artists use Midjourney for rapid prototyping of artistic concepts to show to clients before starting work themselves.Some artists have accused Midjourney of devaluing original creative work by using it in the training set; Midjourney's terms of service includes a DMCA takedown policy, allowing artists to request their work to be removed from the set if they believe copyright infringement to be evident.

The advertising industry has been quick to embrace AI tools such as Midjourney, DALL-E, and Stable Diffusion, among others. The tools, which enable advertisers to create original content and brainstorm ideas quickly are providing new opportunities such as "custom ads created for individuals, a new way to create special effects, or even making e-commerce advertising more efficient", according to Ad Age.

**Code Assistants: AWS Code Wisper, GitHub CoPilot, VS Intellisense (a bit of ai thr as well)**

**AWS Code Wisper**

**What are AWS CodeWhisperer ?**

AWS CodeWhisperer is a machine learning-powered service that generates code recommendations based on developers’ comments and code. It is a code review tool that automates code analysis, identifies potential issues, and improves code quality. It is available as part of the AWS Toolkit extension for major IDEs, such as JetBrains, Visual Studio Code, and AWS Cloud9, and as a native feature on the AWS Lambda console. It supports Python, Java, and JavaScript.

**How do they work?**

Both AWS CodeWhisperer operate by scrutinizing pre-existing code snippets and providing suggestions based on the analysis. AWS CodeWhisperer is constructed on Amazon Web Services and utilizes AI-powered natural language processing (NLP) algorithms to comprehend the context of the code and formulate pertinent suggestions.

**2. GitHub Copilot:**

GitHub Copilot is an advanced AI tool - a programming assistant that helps developers write code more quickly and with less effort.

GitHub Copilot uses OpenAI Codex to provide real-time code and function suggestions based on context from comments and existing code. It can be easily integrated into various editors, including Neovim, JetBrains IDEs, or Visual Studio Code.

Some of the features that GitHub Copilot has are:

Real-time code suggestions, turning natural language prompts into code.

Compatibility with dozens of languages, thanks to its training on billions of lines of code.

Context-aware recommendations that fit the project's style patterns.

Two pricing plans: Copilot for Individuals ($10/month or $100/year) and Copilot for Business ($19/user/month), with added benefits for business users.

Direct integration with popular editors like Neovim, JetBrains IDEs, Visual Studio, and Visual Studio Code.

**3.VS Intellisense:**

IntelliSense is a general term for various code editing features including: code completion, parameter info, quick info, and member lists. IntelliSense features are sometimes called by other names such as "code completion", "content assist", and "code hinting."

**What is IntelliSense in VS Code?**

IntelliSense is a general term for various code editing features including: code completion, parameter info, quick info, and member lists. IntelliSense features are sometimes called by other names such as "code completion", "content assist", and "code hinting."

Visual Studio Code IntelliSense is provided for JavaScript, TypeScript, JSON, HTML, CSS, SCSS, and Less out of the box. VS Code supports word based completions for any programming language but can also be configured to have richer IntelliSense by installing a language extension.

**3.How does Prompting tech to integrate with clouds, Azure, AWS, GCP mainly**

**ChatGPT**

Integrating prompting technology, like ChatGPT, with cloud platforms such as Azure,AWS and GCP involves several steps:

**API Integration**: Cloud platforms typically offer APIs(Application Programming Interface)that allow external applications to interact with their services. you’d need to use the appropriate API to connect your prompting tech to the cloud platform of choice.

**Authentication**: Securely authenticate your application with the cloud platform using API keys, tokens or other authentication mechanisms provided by the cloud provider.

**Data Transfer**: If your prompting tech requires access to specific data or services hosted on the cloud ,you’ll need to set up data transfer mechanisms between your application and the  cloud resources. This might involve setting up data pipelines or accessing cloud databases.

**Scaling**: Cloud platforms can dynamically scale resources based on demand. Ensure your prompting tech can take advantage of this scalability, so it can handle varying loads effectively.

**Cost Management**: Be mindful of the cost associated with cloud services. Monitor resource usage and optimize where necessary to avoid unexpected expenses.

**Error Handling** : Implement proper error handling to manage any issues that may arise when interacting with cloud services.

**Security**: Keep security in mind. Ensure that any data transferred between your application and the cloud is encrypted and that you follow best practices for securing your application.

**Monitoring and Logging**: Implement robust monitoring and logging so you can track the performance of your prompting tech when integrated with the cloud platform.

**4.What else can prompting an LLM do...**

 ChatGPT Prompting a Large Language Model (LLM) like GPT-3.5(which powers this conversation) can enable a wide range of capabilities:

**Text Generation**:  LLMs can generate human-like text in response to prompts, which can be used for content creation, creative- writing, and more.

**Language Translation**: LLMs can help translate text from one language to another, making them useful for multilingual communication.

**Information Retrieval**: You can use prompts to ask LLMs questions or request specific information, similar to a search engine.

**Code Generation**: LLMs can generate code snippets in various programming languages based on that provided prompts.

**Conversation:** LLMs can hold interactive conversations, providing responses that can simulate conversation with a human.

**Summarization:** LLMs can summarize longer pieces of text , extracting the key points and main ideas.

**Idea Generation:** LLMs can help brainstorm ideas for various topics, from product names to marketing slogans.

**Data Analysis:** LLMs can assist in analyzing textual data or identify trends, sentiments, or insights.

**Storytelling**: LLMs can create fictional stories, generate plot ideas, or assist in building narrative elements.

**Natural language understanding**: LLMs can analyze and interpret natural language, making them useful for sentiment analysis, intent recognition, and more.

**Educational Assistance:** LLMs can provide explanations, definitions, and examples to help users understand complex topics.

**Content optimization:** LLMs can suggest improvements to written content, making it more engaging, concise , or relevant.

Remember that while LLMs are powerful tools, they also have limitations. They don’t possess real-world knowledge beyond their training data, and their responses are based on patterns in that data. Therefore, critical thinking and validation of the information they provide are essential, especially in professional or educational contexts.