**Day 2**

1.

What is the primary difference between foundation models and traditional AI models?

Foundation models are trained on specific data for a single task, while traditional models are trained on diverse data for various tasks.

**Foundation models are trained on massive amounts of diverse data for various tasks, while traditional models are trained on specific data for a single task.**

Foundation models cannot be adapted to new tasks, while traditional models can.

Foundation models are only trained on text data, while traditional models use images and code.

Correct! This is the key difference between foundation models and traditional AI models.

check

2.

How do foundation models and prompt engineering work together to create value in generative AI?

Foundation models ensure the ethical use of generative AI, while prompt engineering focuses on improving the quality and creativity of outputs.

Prompt engineering trains foundation models on specific tasks, allowing them to generate highly specialized content and insights.

**Foundation models offer a vast knowledge base, and prompt engineering guides the model to use this knowledge in responses.**

Foundation models provide the computing power for generative AI, while prompt engineering directs that power to complete specific tasks.

Correct! Foundation models are trained on massive amounts of text and code, giving them a broad understanding of the world and how language works. Prompt engineering involves crafting specific instructions and questions to guide the model's use of this knowledge to generate the desired output.

check

3.

Which of the following is NOT a key feature of foundation models?

Trained on diverse data.

Adaptable to new domains and tasks.

Flexible to support various use cases.

**Specialized to specific tasks.**

Correct! Foundation models are designed to be more general purpose and adaptable to various tasks.

close

4.

What is the purpose of a prompt in the context of foundation models?

**To provide input to the model and trigger an output.**

To train the model on new data.

To evaluate the model's performance.

To fine-tune the model for a specific task.

Incorrect. **Prompts are used to get outputs from a trained model, not to fine-tune the model** itself. Fine-tuning involves training the model on additional data to specialize it for a particular task.

check

5.

Which of the following best defines a foundation model?

Traditional machine learning algorithms that rely on explicitly defined rules.

Hardware infrastructure used to train and deploy AI models.

Small, very specialized AI models trained on narrow datasets in order to perform specific tasks.

**Large AI models trained on a vast quantity of data, capable of adapting to a variety of tasks.**

Correct! Foundation models, including LLMs, are trained on massive datasets and can be adapted to various applications. They differ from traditional AI models that are designed for specific tasks.

Since 1998, Google has been at the forefront of technological innovation, pioneering advancements in AI to create groundbreaking products like Search, Gmail, and Android.

**2023: Gemini**

Google introduces Gemini, a multimodal AI model capable of understanding and integrating various information types. The following year, Google expands the Gemini ecosystem with Gemini 1.5, bringing it to more products and launching Gemini Advanced, which provides access to Google's most capable AI models.

**2023: Bard**

Google releases Bard, an early experiment that lets people collaborate with generative AI.

**2018: Smart Compose in Gmail**

Google announces Smart Compose, a new feature in Gmail that uses AI to help users more quickly reply to their email. Smart Compose builds on Smart Reply, another AI feature.

**2016: TPUs**

Google announces the Tensor Processing Unit, custom data center silicon built specifically for machine learning.

**2015: Google Search**

Google integrates **RankBrain** into Google Search, using machine learning to analyze search queries and deliver more relevant results to users.

**2006: Google Translate**

Google launches Google Translate using machine learning to automatically translate languages.

**2001: ML for spell check**

Google begins using machine learning to help with spell check at scale in Google Search.

**Individual productivity and efficiency**

Gen AI has been built into tooling across Google’s comprehensive ecosystem to support your organization in their day-to-day work. Here are some examples:

* **Google Search:** Leverage the power of Gemini in Google Search to find information faster and with more accuracy.
* **Gemini for Google Workspace:** Draft emails in Gmail, generate presentations in Slides, summarize meeting notes in Docs, and automate tasks in Sheets with integrated gen AI features.
* **Gemini App:** Access the power of Gemini directly through a dedicated app for personalized assistance and creative exploration.
* **Gemini for Google Cloud:** Build applications and services on Google Cloud using powerful Gemini models.

**Continuous improvement**

Even if you build based on the latest gen AI technology today, you still need to stay on top of all the new advancements coming. Security enhancements, performance enhancements, new feature updates. By using Google’s ecosystem, Google will help support you in staying up to date on the latest advancements. Here are some examples:

* **Automatic model upgrades:** Benefit from continuous improvements to Gemini models without needing to manage infrastructure or retrain models.
* **Access to new features:** Gain early access to new Gemini features and capabilities as they are released.
* **Security patches and updates:** Rely on Google's expertise to keep your AI systems secure and up-to-date.

**Responsible AI**

Using gen AI responsibly, ethically, and securely is extremely important for all organizations. Google has a commitment to building AI responsibly by building an ecosystem that is built securely and ethically and, at the same time, providing you educational resources for doing the same as you build your own solutions. Here are some examples:

* Secure AI Framework (SAIF): A comprehensive suite of tools and best practices for building secure AI systems.
* Mandiant: Leverage Mandiant's threat intelligence and expertise to protect your AI systems from cyberattacks.
* AI Principles: Google's published AI Principles guide the development and deployment of AI technologies.
* Responsible AI Toolkit: Resources and tools to help developers and organizations build AI systems that are fair, unbiased, and socially beneficial.

**Enterprise ready**

Building gen AI for an enterprise involves thinking about security measures, data governance, compliance certifications, and more. Through using Google Cloud’s ecosystem, you know you are building off an enterprise-ready system. Here are some examples:

* Vertex AI: A unified platform for building and deploying machine learning models, including Gemini. It provides enterprise-grade security, scalability, and compliance.
* Google Cloud's security infrastructure: Benefit from Google's world-class security infrastructure, including data encryption, access control, and network security.
* Compliance certifications: Google Cloud offers services that support compliance with a wide range of industry standards and regulations, including ISO 27001, SOC 2, and HIPAA.

**Open-source approach**

Google has an open approach so you are not stuck to any one solution and you have the flexibility to choose what works best for your organization. Here are some examples:

* **Contributions to TensorFlow and PyTorch:** Google actively contributes to popular open source machine learning frameworks.
* **Open models and datasets:** Google releases pre-trained models and datasets to the research community.
* **Support for open standards:** Google supports open standards for AI interoperability and data exchange.

**Implementing a multi-directional strategy**

Successful gen AI implementation requires a combination of top-down and bottom-up approaches. For both approaches, you want to be strategizing on:

* Strategic focus
* Exploration
* Responsible AI
* Resourcing
* Impact
* Continuous improvement

Augmentation versus automation

Use gen AI to enhance or **augment** your strategic thinking for:

* **Critical thinking and problem solving:**Gen AI can provide data and insights, but humans are still needed to interpret those insights and make informed decisions.
* **Creativity and innovation:**Gen AI can assist in generating ideas and exploring possibilities, but human ingenuity is still essential for pushing boundaries and developing truly innovative solutions.
* **Relationship building and collaboration:**Gen AI can facilitate communication and information sharing, but strong interpersonal skills are still crucial for building trust, fostering collaboration, and navigating complex human dynamics.
* **Strategic planning and vision:**Gen AI can help with forecasting and trend analysis, but human leadership is essential for setting a long-term vision, defining goals, and charting a course for the future.

Use gen AI to **automate** tasks that are:

* **Repetitive and rule-based:**data entry, information retrieval, content formatting, and basic code generation.
* **Time-consuming and resource-intensive:**research, data analysis, content summarization, and initial draft creation.

Even for task automation, humans-in-the-loop are a necessary component for the gen AI implementation and continuous improvement. Use people for:

* **Data selection and preparation:**ensuring that gen AI models are trained on high-quality, relevant data that is representative of the intended use cases.
* **Prompt design and refinement:**crafting prompts that elicit accurate and useful responses from gen AI models.
* **Output evaluation and refinement:**reviewing and editing gen AI-generated content to ensure accuracy, relevance, and alignment with brand guidelines.
* **Continuous monitoring and feedback:**providing feedback on gen AI performance and identifying areas for improvement.

**Embracing the potential of generative AI**

Don't be intimidated. Generative AI tools are evolving rapidly, but you don't need to be a technical expert to understand how to use them or how they can positively impact your business. Even though it's a relatively new technology, the principles behind it are easy to understand and apply across various business functions as you explored above.

Remember, the most successful gen AI implementations prioritize a clear vision and focus on impactful use cases. Additionally, tracking results are key elements for businesses to successfully implement gen AI and achieve a competitive advantage.

1.

What is the recommended approach for businesses to effectively implement gen AI?

Focus on a top-down strategy driven by executives.

**Combine a top-down strategic vision with bottom-up input from teams.**

Adopt the latest gen AI trends.

Focus on a bottom-up approach with no overarching plan.

Correct! This multi-directional approach ensures alignment with business goals while leveraging the insights and creativity of teams who will be using the technology.

check

2.

Which of the following is a benefit of using Google Cloud for gen AI development?

**It provides comprehensive services, including scalable infrastructure, enterprise-grade governance and security.**

It offers a limited set of pre-trained AI models for specific industry use cases.

It focuses primarily on individual developer tools without robust governance or scalability features.

It requires extensive upfront hardware investments for running complex AI workloads.

Correct! Google Cloud's AI platform encompasses the entire AI lifecycle, from data ingestion and preparation to model building, deployment, and monitoring, all supported by a scalable and secure infrastructure suitable for enterprise needs.

check

3.

Why is it important for mid-level managers and individual contributors to be involved in gen AI adoption?

They should avoid any involvement in gen AI experimentation.

**Their proximity to workflows allows them to identify impactful gen AI solutions.**

**Their focus should be solely on technical implementation details.**

They have limited understanding of daily operations and user needs.

Correct! Because they work closely with processes and users, they are well-positioned to identify areas where gen AI can make a real difference.

check

4.

What is a key advantage of using Google Cloud's gen AI ecosystem for businesses?

It assures businesses that they will achieve immediate and widespread success in all of their generative AI endeavors.

It eliminates the need for companies to cultivate internal gen AI knowledge and experimentation across their organizations.

**It allows businesses to leverage Google's AI advancements without starting from scratch.**

It forces each and every business to become an "AI-first" company.

Let’s recap the key differences between AI, ML, and gen AI from the video.

* **Artificial intelligence (AI):**The broad field of building machines that can perform tasks requiring human intelligence.
* **Machine learning (ML):** A subset of AI where machines learn from data.
* **Generative AI:**An application of AI that creates new content.

1.

How does consistency impact AI model training?

check

**Inconsistent formats and labeling can confuse the model and hinder learning**.

It reduces the need for data storage.

It increases the speed of data retrieval.

It ensures data is relevant to the task.

Correct! As the text illustrates, inconsistencies create confusion.

check

2.

What is the primary purpose of the data ingestion and preparation stage in the ML workflow?

Training machine learning models.

Monitoring model performance.

check

**Collecting, cleaning, and transforming raw data.**

Deploying trained models.

Correct! This accurately describes the data ingestion and preparation stage.

check

3.

What is a "model" in the context of machine learning?

A set of pre-defined rules for decision-making.

A visual representation of data.

A type of computer hardware.

check

**A complex mathematical structure that processes inputs to generate outputs.**

Correct. A model is the mathematical engine that learns from data.

check

4.

Arrange the ML lifecycle steps in the right order.

Model training > Data ingestion and preparation > Model deployment > Model management

Data ingestion and preparation > Model training > Model management > Model deployment

check

**Data ingestion and preparation > Model training > Model deployment > Model management**

Model training > Model deployment > Model management > Data ingestion and preparation

Correct! This accurately describes steps of the ML lifecycle in the correct order.

check

5.

What is the primary way that agents learn in reinforcement learning?

By analyzing large datasets of labeled examples.

By observing and imitating expert demonstrations.

By being explicitly programmed with the correct actions.

check

**By interacting with their environment and receiving feedback.**

Correct! Reinforcement learning agents learn by actively exploring their environment, receiving feedback (rewards or penalties), and adjusting their actions accordingly.

check

6.

Which of the following is an example of unstructured data?

check

**A collection of customer reviews in the form of free-text paragraphs.**

A database of product IDs, prices, and inventory levels.

A spreadsheet with customer names, addresses, and purchase history.

A table of employee names, salaries, and job titles.

Correct. Free-text reviews lack a predefined structure, making them unstructured data.

Deep learning

In previous lessons, you explored the connection between artificial intelligence and machine learning, as well as the core types of machine learning: supervised, unsupervised, and reinforcement learning.

Now, let's see how deep learning fits into this larger picture.

**Machine learning**

A broad field that encompasses many different techniques, one of which is deep learning (DL).

**Deep learning**

A powerful subset of machine learning, distinguished by its use of artificial neural networks. These networks enable the processing of highly complex patterns and the generation of sophisticated predictions.

Neural networks can leverage both labeled and unlabeled data, a strategy known as semi-supervised learning. They train on a blend of a small amount of labeled data and a large amount of unlabeled data. That way, they learn foundational concepts and generalize effectively to novel examples.

Generative AI uses the power of deep learning to create new content spanning text, images, audio, and beyond. Deep learning techniques, particularly those centered on neural networks, are the engine behind these generative models.

Foundation models

Remember foundation models from course one? Foundation models use deep learning. They are trained on massive datasets that allow them to learn complex patterns and perform a variety of tasks across different domains. They are incredibly powerful machine learning models trained on a massive scale, often using vast amounts of unlabeled data. This training allows them to develop a broad understanding of the world, capturing intricate patterns and relationships within the data they consume.

Think of a foundation model as a highly advanced learner. It's like a student who has read everything in an entire library, absorbing knowledge from countless books, articles, and websites. This deep and extensive learning allows foundation models to be adapted to a wide range of tasks.

**Large language models (LLMs)**​​​

One particularly exciting type of foundation model is the LLM. These models are specifically designed to understand and generate human language.

They can translate languages, write different kinds of creative content, and answer your questions in an informative way, even if they are open ended, challenging, or strange. This is likely the most common foundation model you've encountered, such as in popular generative AI chatbots like Gemini. They also help power many search engines you use today.

**Diffusion models**

Diffusion models are another type of foundational model. They excel in generating high-quality images, audio, and even video by iteratively refining noise (or unstructured/random data and patterns) into structured data. ,

deep learning provides the core technology, foundation models are the powerful architectures built on deep learning, and generative AI is the application of these models to create new and original content.

**Factors when choosing a model for your use case**

When picking a model for your gen AI use case, it's important to start to think about some criteria for different models. Here are some important factors to keep in mind.

**Modality**

When selecting a generative AI model, it's crucial to consider the modality of your input and output. Modality refers to the type of data the model can process and generate, such as text, images, video, or audio. If your application focuses on a single data type, like generating text-based articles or creating audio files, you'll want to choose a model optimized for that specific modality. For applications that require handling multiple data types, such as generating image captions (processing images and producing text) or creating video with accompanying audio, you'll need a multimodal model. These models can understand and synthesize information across different modalities.

**context window**

The context window refers to the amount of information a model can consider at one time when generating a response. A larger context window allows the model to "remember" more of the conversation or document, leading to more coherent and relevant outputs, especially for longer texts or complex tasks. However, larger context windows often come with increased computational costs. You need to balance the need for context with the practical limitations of your resources.

**Security** is paramount, especially when dealing with sensitive data. Consider the model's security features, including data encryption, access controls, and vulnerability management. Ensure the model complies with relevant security standards and regulations for your industry.

The **availability and reliability** of the model are crucial for production applications. Choose a model that is consistently available and performs reliably under load. Consider factors like uptime guarantees, redundancy, and disaster recovery mechanisms.

**cost**

Generative AI models can vary significantly in cost. Consider the pricing model, which might be based on usage, compute time, or other factors. Evaluate the cost-effectiveness of the model in relation to your budget and the expected value of your application. This is where selecting the right model for the right task is important. Be sure to match the model to the task; bigger isn't always better, and multi-modal capabilities aren't always necessary.

The **performance** of the model, including its accuracy, speed, and efficiency, is a critical factor. Evaluate the model's performance on relevant benchmarks and datasets. Consider the trade-offs between performance and cost.

Some models can be **fine-tuned or customized** for specific tasks or domains. If you have a specialized use case, consider models that offer fine-tuning capabilities. This often involves training the model further on a specific dataset related to your use case.

The **ease of integrating** the model into your existing systems and workflows is an important consideration. Look for models that offer well-documented APIs and SDKs.

**Google Cloud’s ML models**

Vertex AI streamlines the integration of advanced artificial intelligence capabilities into business applications, allowing for seamless discovery, deployment, and customization. These models empower businesses to leverage cutting-edge AI, providing the flexibility to work with many different models without the need for extensive in-house model development.

With Vertex AI you can access models developed by Google including **Gemini**, **Gemma**, **Imagen**, and **Veo**. You can also access proprietary third-party models, and openly available models.

**Gemini**

Gemini, a multimodal model, can understand and operate across diverse data formats, such as text, images, audio, and video. Gemini's multimodal design supports applications that require complex multimodal understanding, advanced conversational AI, content creation, and nuanced question answering.

**Gemma**

A family of lightweight, open models is built upon the research and technology behind Gemini. They offer developers a user-friendly and customizable solution for local deployments and specialized AI applications.

**Imagen**

A powerful text-to-image diffusion model, it excels at generating high-quality images from textual descriptions. This makes it invaluable for creative design, ecommerce visualization, and content creation.

**Veo**

A model capable of generating video content. It can produce videos based on textual descriptions or still images. Its functionality allows for the creation of moving images for applications such as film production, advertising, and online content.

**Foundation model limitations**

Foundation models, while groundbreaking, aren't without limitations. Recognizing these limitations is essential for the responsible and effective utilization of these powerful tools.

**Here is a detailed summary:**

| Feature | RAG | Fine-tuning | Grounding |
| --- | --- | --- | --- |
| Definition | Augments LLMs by retrieving relevant information from external knowledge bases and adding it to the prompt. | Further trains a pre-trained model on a new dataset to adapt it to a specific task or domain. | Connects an AI model's output to verifiable sources of information. |
| Process | Retrieve relevant information. → Add it to the prompt. → Generate a response. | Select a pre-trained model. → Gather data. → Train the model. → Evaluate and refine. | Provide access to data sources. → Use RAG or fine-tuning to connect the output. |
| Data sources | External knowledge bases (databases, documents, internet). | Task- or domain-specific datasets. | External knowledge bases or specific datasets. |
| Relationship to grounding | A specific technique for achieving grounding. | Improves a model's ability to be grounded in specific domains. | The overarching goal, achieved through techniques like RAG and fine-tuning. |

What is the primary role of humans in the loop (HITL) in machine learning?

To automate all decision-making processes.

To replace AI algorithms entirely.

To eliminate the need for data collection and model training.

**To integrate human expertise into the ML process, especially for tasks requiring judgment or context**.

Correct. HITL brings human expertise to address limitations of AI and improve overall performance.

check

2.

Which type of foundation model is specifically designed to understand and generate human language?

Multimodal model

Diffusion model

**Large language model (LLM)**

Predictive model

Correct. LLMs are focused on human language processing and generation.

check

3.

Which techniques can be used to overcome the limitations of foundation model performance?

Preventing hallucinations by restricting the AI model's access to external knowledge sources.

**Grounding, prompt engineering, fine-tuning, and humans in the loop (HITL).**

Advanced algorithms and data structures.

Increased processing power and faster hardware.

Correct. These techniques offer a comprehensive approach to improving the performance and trustworthiness of generative AI models.

check

4.

Which type of foundation model would be most suitable for generating photorealistic images from textual descriptions?

Multimodal model

**Diffusion model**

Classification model

Large language model (LLM)

Correct! **Diffusion models excel at generating high-quality images from text**.

check

5.

What is the primary purpose of grounding in generative AI?

To improve the AI's ability to generate creative content.

To reduce the cost of deploying AI applications.

**To connect the AI's output to verifiable sources of information**.

To increase the speed and efficiency of AI model training.

Correct. Grounding aims to ensure the AI's responses are grounded in reality and not hallucinations.

check

6.

Why is fine-tuning and customization an important factor when choosing a model?

It reduces the cost of using the model.

**It allows the model to be adapted for specific tasks or domains.**

It increases the model's context window.

It ensures the model is always available.

Correct! Fine-tuning tailors the model to specialized needs.

**What does Secure AI mean?**

Secure AI is about preventing intentional harm being done to your applications. This is about protecting AI systems from malicious attacks and misuse. For all applications, including AI, you need to ensure security throughout the full lifecycle from development through deployment. This includes considering the data, infrastructure, and how and where applications are deployed.

To better understand how these security principles apply, let's reexamine the machine learning lifecycle from an earlier lesson.

**Gather your data**

The foundation of any robust AI system is secure data. Therefore, data must be protected and secured at all times. Access controls must be implemented to restrict who can access the data. They are also necessary to restrict who can add or input to the data. This is crucial to prevent data poisoning, a malicious attack where bad actors corrupt your AI model’s training data by introducing manipulated data. This can cause your model to learn incorrect patterns and make flawed (biased, inaccurate, or even harmful) predictions. It's akin to someone secretly swapping out vital ingredients in a recipe, leading to a disastrous final dish.

**Prepare your data**During data preparation, special attention should be paid to confidential or sensitive data that might be present in the training data. Where feasible, anonymization techniques should be applied to this data, mitigating potential privacy risks. Data also should be rigorously validated, using integrity checks and anomaly detection to prevent poisoning. Data should be securely processed with encryption both at rest and in use. Logging and real-time monitoring should also be performed.

**Train your model**Securing the model and the training process is paramount. This means safeguarding both the training data and model parameters from unauthorized access or modification. Model theft, where attackers steal proprietary or sensitive AI models, is a significant threat. Beyond gaining a competitive advantage, the stolen models could be used for malicious purposes, exploiting vulnerabilities or replicating sensitive functionality.

**Deploy and predict**Once your model is trained and ready for action, it's crucial to safeguard it within its operating environment. This means controlling access to the model, determining who can interact with it and how. If you're using a pre-built model, it's essential to verify its source and check for any potential vulnerabilities.

**Manage your model**Continuous monitoring and maintenance of the model's security are essential. Stay up to date on the latest security best practices in the field, specifically for your platform and model, and regularly update to patch vulnerabilities. Monitor model performance and outputs for anomalies or signs of tampering, and regularly review access permissions.

Think of it like installing a high-tech security system in your office. You wouldn't give everyone the master key, right? Similarly, you need to restrict access to your model and filter the information it receives and sends out.

One major risk here is adversarial attacks, where attackers try to trick the model with misleading information. Imagine someone trying to bypass your security system with a fake ID—that's essentially what an adversarial attack does to your AI. Therefore, it's critical to have safeguards in place to filter, sanitize, and check the data coming into your model, ensuring it's not being fed misleading information. Likewise, you need to monitor the information your model sends out. Think of it as screening outgoing messages to prevent any accidental leaks of sensitive data or the spread of harmful content.

Stay informed

Keeping AI applications secure demands constant vigilance and staying informed about the latest security threats and research.

**Applying the Secure AI Framework (SAIF)**​

Google has developed the Secure AI Framework, or SAIF, to establish security standards for building and deploying AI systems responsibly. This comprehensive approach to AI/ML model risk management addresses the key concerns of security professionals in the rapidly evolving landscape of AI. Following the security practices outlined in SAIF can help your organization find and stop threats, automatically strengthen its defenses, and manage the unique risks of each AI system. The framework is designed to integrate with your company’s existing security, ensuring that AI models are secure by default. For more information, explore [Google's Secure AI Framework (SAIF)](https://safety.google/cybersecurity-advancements/saif/).

**Google Cloud security tools**​

Platforms such as Google Cloud can facilitate secure development. In addition to the Secure AI Framework, Google Cloud offers a range of tools to ensure applications remain secure throughout their lifecycle. Google Cloud has built security into its core through a secure-by-design infrastructure, encompassing its global network and hardware, as well as robust encryption in transit and at rest. It provides customers with detailed control over access and usage of cloud resources through Identity and Access Management (IAM). Google Cloud Security Command Center provides a centralized view of your security posture across your entire Google Cloud environment, and Google Cloud also offers tools for monitoring various workloads.

Key takeaway

AI offers significant benefits, but it also introduces security risks like data poisoning, model theft, and prompt injection. To address these challenges, a secure foundation for AI applications is essential. Google Cloud's SAIF framework, combined with security tools, can help as you build and maintain secure AI systems.

Responsible AI

Responsible AI means ensuring your AI applications avoid intentional and unintentional harm. It’s important to consider ethical development and positive outcomes for the software you develop. The same holds true, and perhaps even more so, for AI applications.

Building on security

The foundation of responsible AI is security. Secure applications protect both your company and your users. Think of it like building a house: if the foundation is weak, the entire structure is compromised, no matter how beautiful the design. Just as a strong foundation ensures a stable and safe house, robust security forms the essential foundation for building truly responsible AI.

*Select the + buttons below to expand each section and learn more.*

**Transparency is key**

**[Users need to know](https://storage.googleapis.com/cloud-training/cls-html5-courses/C-GENDEC-B/v1.0.0/C-GENDEC-B%20-%20Gen%20AI%20Unlock%20Foundational%20Concepts/index.html)**

Transparency is paramount for responsible applications. Users need to understand how their information is being used and how the AI system works. This transparency should extend to all aspects of the AI's operation, including data handling, decision-making processes, and potential biases.

**Privacy in the age of AI**

**[Users need to know](https://storage.googleapis.com/cloud-training/cls-html5-courses/C-GENDEC-B/v1.0.0/C-GENDEC-B%20-%20Gen%20AI%20Unlock%20Foundational%20Concepts/index.html)**[​](https://storage.googleapis.com/cloud-training/cls-html5-courses/C-GENDEC-B/v1.0.0/C-GENDEC-B%20-%20Gen%20AI%20Unlock%20Foundational%20Concepts/index.html)

Protecting privacy often involves anonymizing or pseudonymizing data, ensuring individuals can't be easily identified. AI models can sometimes inadvertently leak sensitive information from their training data, so it's crucial to implement safeguards to prevent this.

**Data quality, bias, and fairness**

**[Ethical AI requires high quality data](https://storage.googleapis.com/cloud-training/cls-html5-courses/C-GENDEC-B/v1.0.0/C-GENDEC-B%20-%20Gen%20AI%20Unlock%20Foundational%20Concepts/index.html)**

**[Understanding and mitigating bias](https://storage.googleapis.com/cloud-training/cls-html5-courses/C-GENDEC-B/v1.0.0/C-GENDEC-B%20-%20Gen%20AI%20Unlock%20Foundational%20Concepts/index.html)**[​](https://storage.googleapis.com/cloud-training/cls-html5-courses/C-GENDEC-B/v1.0.0/C-GENDEC-B%20-%20Gen%20AI%20Unlock%20Foundational%20Concepts/index.html)

AI systems are not independent of the world they are built in. They can inherit and amplify existing societal biases. This can lead to unfair outcomes, such as a resume-screening tool that favors a certain demographic of candidates due to historical biases in hiring data. It's like training a dog with biased commands; the dog will learn and replicate those biases. To counter this, fairness must be a core principle in AI development.

**Accountability and explainability**

**[Fairness requires accountability](https://storage.googleapis.com/cloud-training/cls-html5-courses/C-GENDEC-B/v1.0.0/C-GENDEC-B%20-%20Gen%20AI%20Unlock%20Foundational%20Concepts/index.html)**[​](https://storage.googleapis.com/cloud-training/cls-html5-courses/C-GENDEC-B/v1.0.0/C-GENDEC-B%20-%20Gen%20AI%20Unlock%20Foundational%20Concepts/index.html)

Fairness requires accountability. We need to know who is responsible for the AI's actions and understand how it makes decisions. This is where explainability comes in. Explainable AI makes the decision-making processes of AI models transparent and understandable. This is crucial for building trust, debugging errors, and uncovering hidden biases. Think of it like a judge explaining their verdict; without a clear explanation, it's hard to trust the decision. Tools like Google Cloud’s Vertex Explainable AI can help understand model outputs and identify potential biases. Understanding how your application uses and interprets the AI's output is crucial for ensuring responsible use.

Legal implications

Beyond considerations like fairness and bias, AI development is increasingly governed by legal frameworks. Key areas include data privacy, non-discrimination, intellectual property, and product liability.

* Laws mandate responsible data handling, bias mitigation, and transparency in algorithmic decision-making.
* Organizations must also adhere to the specific rules and limitations of the AI models they employ, ensuring compliance with licensing agreements and legal standards.
* The legal landscape is rapidly evolving, requiring organizations to stay informed and seek legal counsel to ensure compliance.
* Legal compliance is not just a regulatory hurdle. Navigating these legal implications is crucial for building trustworthy AI.

What is the primary goal of ethical AI development?

To comply with all legal and regulatory requirements.

**To ensure AI systems are used responsibly and do not cause harm.**

To promote transparency and accountability in AI systems.

To maximize AI capabilities regardless of societal impact.

Correct! This is the core purpose of ethical AI development.

check

2.

Which of the following is a key aspect of securing the "model training" phase?

Minimizing the amount of training data used.

Ensuring the model's output is visually appealing.

Optimizing the model for faster processing.

**Safeguarding training data and model parameters from unauthorized access.**

Correct! Protecting sensitive data and model configurations is crucial for security in the training phase.

check

3.

What is a potential consequence of using inaccurate or incomplete data in AI training?

It results in an inefficient user interface.

**It introduces biased outcomes and unfair results.**

It results in data storage inefficiency.

It eliminates the need for fine-tuning and further model training.

Correct! Inaccurate or incomplete data can introduce or amplify biases in AI models, leading to unfair or discriminatory outcomes. This is a critical ethical consideration in AI development.

check

4.

What is the primary goal of the Secure AI Framework (SAIF)?

To focus solely on preventing external attacks on AI systems.

**To establish security standards for building and deploying AI responsibly, addressing the unique challenges and threats in the AI landscape.**

To restrict AI development and deployment due to concerns about potential security risks.

To promote innovation and accelerate AI development without considering security risks.

Correct. SAIF aims to provide a comprehensive framework for secure AI development and deployment.

check

5.

You're developing an AI-powered loan application assessment system. Which steps would you take to ensure your system is developed responsibly? Select two.

Prioritize speed and efficiency over fairness and transparency, as these factors are more important for business success.

Use a complex "black box" model that makes highly accurate predictions, even if its decision-making process is difficult to understand.

**Train the model on a dataset that includes a diverse range of applicants, ensuring representation across different demographics and socioeconomic backgrounds**.

Correct! This helps mitigate bias and promotes fairness.

**Regularly audit the model's performance to identify and mitigate any biases that may emerge over time**.

Correct! Ongoing monitoring and evaluation are crucial for maintaining fairness and accountability.

Deploy the model without any human oversight, trusting its ability to make fair and unbiased decisions.

**Day 3 Question**

1.

What are the two key elements that distinguish AI agents from standalone AI models?

Natural language processing and machine learning.

Automation and personalization.

**Reasoning loop and tools.**

Data access and user interface.

Correct! AI agents possess a reasoning loop that allows them to perceive their environment, analyze information, make decisions, and take actions. AI agents can also utilize tools to interact with their environment and achieve their goals.

check

2.

A news organization wants to develop an AI agent that delivers personalized news to each user. The agent should be able to:

• Learn the user's interests and reading habits.

• Filter and prioritize news articles based on relevance.

• Summarize key information from multiple sources.

• Recommend related articles and diverse perspectives.

• Adapt to the user's feedback and evolving preferences.

How would the "agents" layer contribute to the functionality of this personalized news reader application?

The agents layer would provide the underlying AI models for natural language processing and understanding.

**The agents layer would define the specific tasks the AI performs, such as filtering articles, summarizing information, and making recommendations**.

The agents layer would design the user interface of the news reader application.

The agents layer would provide the infrastructure for storing and processing the news data.

Correct! The agents layer defines the AI's actions and goals within the application.

check

3.

A travel agency wants to use an AI agent to help customers plan their vacations. The agent should be able to:

• Gather customer preferences (budget, destination interests, travel dates).

• Search for flights, accommodations, and activities.

• Create personalized itineraries with options and recommendations.

• Book flights and hotels based on customer choices.

• Provide ongoing support throughout the trip.

How would the "agents" layer and the "applications" layer work together to create this AI powered travel planning experience?

**The agents layer would define the AI's capabilities (searching, booking, recommending), while the applications layer would provide the user-facing tool (website or app) to interact with the agent.**

The agents layer would provide the user interface for the travel planning application.

The applications layer would provide the computational resources for the AI agent to operate.

The applications layer would determine the specific tasks the AI agent can perform, such as searching for flights or booking hotels.

Correct! The agents layer defines the AI's actions, and the applications layer provides the way for users to interact with those actions.

check

4.

A game developer wants to create more realistic and engaging non-player characters (NPCs) in their game. They envision NPCs that can:

• Engage in dynamic conversations with the player.

• React to the player's actions and choices.

• Adapt their behavior based on the game's environment and storyline.

• Exhibit unique personalities and backstories.

Which layer of the gen AI landscape would be MOST crucial in defining the behaviors and capabilities of these AI powered NPCs?

Infrastructure

**Agents**

Models

Platforms

Correct! Agents define the specific tasks and actions that the AI can perform, such as engaging in conversations, reacting to events, and adapting their behavior.

check

5.

A data science team is training a new generative AI model on a massive dataset of images. They need access to powerful hardware and software resources to handle the computationally intensive training process.

Which layer of the gen AI landscape would provide the necessary computational power and storage for this data science team?

Applications

Platforms

Agents

**Infrastructure**

Correct! The infrastructure layer provides the hardware and software foundation for the entire gen AI landscape.

1.

What are GPUs and TPUs in the context of AI infrastructure?

Software applications for managing AI models.

**Specialized processors designed for parallel processing in AI tasks.**

Storage devices for storing large AI datasets.

Networking protocols for connecting AI systems.

Correct! GPUs and TPUs are crucial for accelerating the training and execution of AI models, especially in computationally intensive tasks.

check

2.

What is the purpose of AutoML in Vertex AI?

**To allow users to build and train AI models with minimal technical expertise.**

To provide pre-trained models for various use cases.

To manage and monitor the performance of deployed AI models.

To automate the entire ML workflow, from data preparation to deployment.

Correct! **AutoML simplifies the model-building process, making AI accessible to a wider range of users, regardless of their coding skills**.

check

3.

Which of the following statements best describes Model Garden on Vertex AI?

A platform for managing and monitoring AI infrastructure.

A service designed for finding and deploying pre-trained AI models.

A platform for developing AI models from scratch and experimenting further.

**A service for discovering, customizing, and deploying existing AI models.**

Correct! Model Garden offers a wide variety of pre-trained models from Google and its partners, allowing you to readily leverage existing AI capabilities.

check

4.

Why is high-performance storage important for generative AI?

To automate the deployment and management of AI models.

To ensure the responsible use of AI systems.

**To store and efficiently access the massive datasets used in AI training.**

To provide a user-friendly interface for interacting with AI models.

Correct! High-performance storage is essential for handling the vast amounts of data required for training and running generative AI models effectively.

check

5.

What is the role of MLOps in Vertex AI?

To learn how to best achieve a goal based on inputs and tools.

To allow users to interact with AI capabilities.

To provide the physical hardware for AI systems.

**To automate, standardize, and manage the ML project lifecycle.**

Correct: MLOps tools in Vertex AI help streamline the ML workflow, enabling better collaboration and continuous improvement in AI projects.

What is Gemini Nano designed for?

Running complex simulations on supercomputers

Training large language models in data centers

**Deploying AI on edge devices and smartphones**

Managing and monitoring AI infrastructure

Correct! Gemini Nano is a compact and efficient AI model specifically optimized for resource-constrained devices, enabling on-device AI capabilities.

check

2.

Who is responsible for building and deploying custom AI agents and integrating AI capabilities into applications?

AI practitioners

Business leaders

**Developers**

Network engineers

Correct! Developers play a crucial role in creating and implementing AI solutions tailored to specific business needs and application requirements.

Your marketing team needs to quickly generate engaging product descriptions for an upcoming ecommerce sale. Which approach is the most efficient way to achieve this?

**Use a gen AI powered application like Gemini for Workspace to draft the descriptions.**

Hire a team of copywriters to write each description manually.

Build a custom model from scratch to generate product descriptions.

Fine-tune a pre-trained language model from Model Garden on existing product data.

Correct! Gen AI applications are designed for fast content generation and leveraging existing applications is often more affordable than building custom models or hiring a large team.

check

4.

What is the primary advantage of edge computing for AI applications?

Increased data storage capacity

Reduced need for data preprocessing

**Real-time responsiveness and reduced latency**

Lower development costs

Correct! Edge computing brings AI processing closer to the data source, enabling faster response times, which is crucial for applications like self-driving cars and drones.

**Day 4 Questions**

**Using examples in your prompts**

When it comes to prompting techniques, there are a few different approaches.

* Zero-shot prompting is like asking a foundation model to complete a task with no prior examples, relying solely on its existing knowledge.
* One-shot prompting involves showing the foundation model just one example, allowing it to learn and apply that knowledge to similar situations.
* Few-shot prompting, on the other hand, provides the foundation model with multiple examples to learn from, which helps it better understand the task and improve its performance.

1.

You're leading a virtual team meeting in Google Meet with colleagues located in different time zones. Some attendees are struggling to understand the discussion due to language barriers. How can Gemini assist in making this Google Meet video conference more accessible for all participants?

Gemini can provide translated summaries of the spoken conversation in the meeting chat.

**Gemini can generate live translated captions with speaker identification, making it easier for everyone to follow along, regardless of language or accent differences.**

Gemini can automatically adjust the volume of each speaker to ensure everyone is heard equally.

Gemini can automatically summarize the discussion for those who missed the meeting.

Correct! Gemini can provide accurate live translated captions in Google Meet, improving accessibility and understanding for all participants.

check

2.

You're designing a language learning app and want to incorporate an AI tutor that can provide personalized feedback to students practicing their conversational skills. Which role-prompt would be most effective in guiding the AI's interactions with students?

"You are a language model built to help students learn a new language."

"Provide personalized feedback to students practicing conversational skills."

"You are a robot programmed to assess language proficiency."

**"You are a helpful and encouraging tutor who provides constructive feedback on language learning exercises."**

Correct! It emphasizes the AI's role as a helpful and encouraging tutor, providing constructive feedback to students as they practice their conversation skills.

check

3.

What is the main advantage of using role prompting?

It improves the AI model's ability to understand complex questions.

It enables complex tasks like planning a business trip itinerary.

It allows the AI model to learn from a few examples.

**It influences the style, tone, and focus of the AI's responses by assigning it a specific persona.**

Correct! The main advantage of role prompting is that it allows you to guide the AI model's responses by assigning it a specific persona, influencing the style, tone, and focus of its output.

4.

You're drafting a proposal for a new client and want to ensure your writing is clear, concise, and persuasive. How can Gemini in Google Docs assist you in this task?

Gemini in Docs can analyze your proposal's content and automatically generate relevant visuals like charts and graphs to enhance its impact.

**Gemini in Docs can provide suggestions for improving your writing style, grammar, and tone, and even generate different phrasing options to enhance your message.**

Gemini in Docs can automatically assign tasks to team members based on their roles on the client project.

Gemini in Docs can predict the client's response to your proposal based on their past interactions.

Correct! Gemini can assist with refining your writing in Google Docs, ensuring your proposal is clear and persuasive.

check

5.

Which prompting technique relies entirely on the foundation model's pre-existing knowledge, without any provided examples?

**Zero-shot prompting**

Multi-shot prompting

One-shot prompting

Few-shot prompting

Correct! Zero-shot prompting requires the model to perform a task with no prior examples, relying solely on its internal knowledge.

1.

How does grounding improve the reliability of a Gem in Gemini?

It allows the Gem to access and process information from any website on the internet.

**It connects the Gem's output to specific and verifiable sources of information.**

It enables the Gem to learn from user interactions and improve its responses over time.

It prevents the Gem from generating any responses that are inaccurate.

Correct! Grounding ensures that the Gem's responses are based on reliable sources, making it more trustworthy.

check

2.

You're a product manager preparing for a meeting to discuss a new product launch. You have market research reports, competitor analysis, and customer feedback surveys stored in your Google Drive. How can NotebookLM help you prepare for this meeting?

NotebookLM can automatically send calendar invites to all meeting participants with a suggested agenda.

**NotebookLM can summarize the key findings from your documents, identify potential challenges, and provide talking points for discussion, helping you lead a productive meeting.**

NotebookLM can automatically generate a presentation with slides and talking points for your meeting.

NotebookLM can predict the questions your team will ask based on their personalities and prepare answers in advance.

Correct! NotebookLM can analyze your documents and provide a synthesized overview, making it easier to prepare and lead a meeting.

check

3.

What is the primary advantage of using Saved info in Gemini compared to reusing prompts by copying and pasting them from a document?

**Saved info ensures that specific information, like your role or company details, is consistently applied across all your Gemini interactions.**

Saved info allows you to store prompts with richer formatting, such as bolding and bullet points.

Saved info is automatically shared with all your colleagues, promoting collaboration.

Saved info is where you save your Gemini passwords.

Correct! Saved info provides a persistent context for your Gemini conversations, avoiding repetition and potential inconsistencies.

check

4.

What is the primary function of Retrieval-Augmented Generation (RAG) in AI models?

To replace human intervention in the AI model's output generation process.

To restrict the AI model's output to a predefined set of responses.

To enhance the AI model's ability to memorize vast amounts of information.

**To enable the AI model to access and utilize external knowledge sources for generating outputs**.

Correct! RAG is all about connecting AI models to a wider world of information beyond their initial training data. This helps them give more accurate and up-to-date answers.

check

5.

What is a key advantage of using NotebookLM Enterprise for team projects?

**You can manage notebook access using predefined identity and access management (IAM) roles in NotebookLM Enterprise.**

Google has access to all your data in NotebookLM Enterprise.

Google has control over your data in NotebookLM Enterprise.

NotebookLM Enterprise is the same as the NotebookLM personal product.

Correct! NotebookLM Enterprise is designed for teams and organizations, so it allows access to notebooks to be controlled using predefined IAM roles. This ensures that sensitive data is only accessible to authorized individuals.

check

6.

You're a freelance writer who specializes in creating different types of content, including blog posts, social media captions, and website copy. You find yourself frequently switching between different writing styles and tones depending on the project. You want to streamline your workflow and ensure consistency in your writing across various formats. Which feature of Gemini best suits your needs?

Gemini Saved info

Gemini Deep Research

Gemini for Google Cloud

**Gems**

Correct! This allows you to quickly switch between different writing modes and maintain consistency within each content type.

1.

A team of developers wants to improve their coding efficiency and collaboration. Which Gemini for Google Cloud tool would be most beneficial for them?

Gemini Cloud Assist

Gemini in Databases

**Gemini Code Assist**

Gemini in Security Command Center

Correct! Gemini Code Assist acts as an AI pair programmer, helping developers write better code collaboratively.

check

2.

Which of the following statements about the Gemini for Google Cloud enterprise security model is true?

Customer data is shared with third-party developers.

**Customer data is encrypted and access controls are in place**.

Google controls access and permissions for customer data.

Customer data is included in publicly accessible datasets.

Correct! This is fundamental to Gemini's security model, ensuring data confidentiality and protection.

1.

Which of the following scenarios best illustrates the concept of a multi-agent system in the context of generative AI?

**A system where multiple specialized AI agents collaborate to achieve a complex goal, such as booking a trip.**

A single AI agent capable of performing multiple tasks, such as writing emails and summarizing documents.

An AI agent that uses multiple tools, like a database and a web search API, to complete a task.

An AI agent that can be customized with different parameters to adapt to different tasks.

Correct. It highlights the core idea of multi-agent systems: collaboration between specialized agents.

check

2.

Which parameter in Google AI Studio would you adjust to control the randomness and creativity of Gemini's output?

Output length

Safety settings

Token count

**Temperature**

Correct! Temperature influences the randomness and unpredictability of the AI's response.

check

3.

What is the primary distinction between Google AI Studio and Vertex AI Studio in terms of their target users?

Google AI Studio is for building custom models, while Vertex AI Studio is for using pre-trained models.

Google AI Studio is for accessing image generation models, while Vertex AI Studio is for text-based models.

Google AI Studio requires coding experience, while Vertex AI Studio offers a no-code interface for interacting with models.

**Google AI Studio is designed for beginners and experimentation, while Vertex AI Studio caters to professionals and larger-scale projects.**

Correct. Google AI Studio prioritizes ease of use and accessibility, while Vertex AI Studio offers more advanced features and scalability for professionals.

close

4.

Imagine an AI agent tasked with scheduling a team meeting. The agent needs to access the team members' calendars, propose suitable time slots, and send out invitations. Which combination of agent tools would be most essential for this task?

Data stores and functions

**Extensions and plugins**

closeFunctions and databases

Plugins and data stores

Incorrect. While functions can perform calculations, and databases can store structured data like user information or meeting details, neither provides the necessary tools for interacting with external calendar systems or sending invitations.

check

5.

Which of the following statements best describes the function of a "reasoning loop" within a generative AI agent?

It establishes a one-time set of instructions for the agent to follow without deviation.

It enables the agent to interact with users in a conversational manner, responding to prompts and questions.

**It guides the agent through an iterative process of observation, internal reasoning, and decision-making to solve tasks.**

It restricts the agent's actions to a predefined set of responses, ensuring predictable behavior.

Correct. The reasoning loop is the core process driving the agent's problem-solving abilities.

1.

What is the primary purpose of Google Agentspace within an organization?

To provide a customer-facing interface for interacting with AI-powered chatbots.

**To provide a centralized platform for organizing and accessing AI agents that can utilize company data to assist employees.**

To enhance the organization's cybersecurity defenses by detecting and preventing data breaches.

To create a centralized repository for storing and managing all company data.

Correct. Agentspace enables the creation and deployment of AI agents designed to help employees with various tasks.

check

2.

Which core capability of Vertex AI Search helps mitigate the issue of hallucinations in generative AI, ensuring that search results are grounded in reliable information?

Granular access controls

**Grounding**

Advanced analytics

Multimodal search

Correct. Grounding responses in reliable sources like knowledge graphs and factual data helps validate the information and reduce the likelihood of hallucinations.

check

3.

According to the provided information, how does a large language model (LLM) interact with data stores in a retrieval-augmented generation (RAG) workflow?

The LLM uses data stores to permanently expand its base knowledge and parameters.

The LLM exclusively relies on data stores for generating responses, disregarding its internal knowledge.

**The LLM queries data stores to retrieve information relevant to a user's request, enriching its response.**

The LLM passively receives information from data stores, waiting for updates to be pushed to it.

Correct. In RAG, the LLM actively interacts with data stores, retrieving relevant information to enhance its understanding of the user's query and generate more informed responses.

check

4.

What are the key steps in planning for impact when implementing generative AI solutions?

Focus on technology adoption and assume positive impact will follow.

Prioritize AI capabilities without considering ethical implications.

Collect and analyze data with a focus on short-term goals in the beginning.

**Define key metrics, collect and analyze data, and iterate and improve.**

Correct! This is the most comprehensive approach to ensuring your generative AI initiatives are successful. Defining clear metrics helps you track progress, collecting and analyzing data provides insights into what's working and what's not, and iterating and improving allows you to optimize your solutions over time.

check

5.

Which of the following options best identifies the key components of an AI agent?

A complex algorithm, a user interface, and a data storage system.

**A reasoning loop, tools, and a foundational model.**

A detailed user manual, customer service chatbots, and advanced marketing techniques.

API connections, cloud functions, and a sophisticated database.

Correct. Generative AI agents consist of a foundational model, tools, and a reasoning loop. The reasoning loop enables the agent to take in information, reason, and make decisions. Tools allow the agent to interact with the outside world. And the foundational model powers the agent.

check

6.

What is a key advantage of using a hybrid approach (combining deterministic and generative AI) when building conversational agents with Google Cloud's Customer Engagement Suite?

It allows agents to handle only simple, predictable customer queries.

**It enables strict control over agent behavior while leveraging the flexibility of generative AI.**

It eliminates the need for human agents in customer interactions.

It restricts agent responses to predefined scripts, ensuring consistency.

Correct. The hybrid approach combines the strengths of both methods, providing control and flexibility.

**Day1:Quiz**

**check**

1.

A CEO is hesitant to invest in generative AI because they believe it's just a technology for building chatbots. Which of the following statements would BEST demonstrate that generative AI can offer their business much more?

Scheduling meetings and managing calendars to help employees improve their time management.

Counting the number of website visitors and their locations.

**Creating photorealistic images of new product prototypes based on text descriptions.**

Providing generic responses to all customer inquiries, regardless of the specific issue.

Correct! This example highlights the content creation capabilities of generative AI beyond chatbots.

close

2.

Which of the following statements accurately describe Gemini? Select two.infoNote: To get credit for a multiple-select question, you must select all of the correct options and none of the incorrect ones.

closeGemini is a type of artificial intelligence that can create new content, including text, images, music, and even code.

Incorrect. This is a general definition of generative AI, not specific to Gemini.

Gemini is a type of generative AI specifically designed for generating images.

Gemini is an AI assistant that can help you be more creative and productive.

**Gemini is a chatbot that can answer questions and generate creative content.**

**Gemini is a generative AI model (or family of models) developed by Google**.

Correct! This is the most accurate and concise definition of Gemini. It highlights its core nature as a generative AI model.

check

3.

Which of the following BEST defines generative AI?

A type of artificial intelligence that analyzes existing data to identify patterns and make predictions.

**A type of artificial intelligence that can create new content, including images, text and music.**

A specific application, like a chatbot, that utilizes AI technology.

A type of artificial intelligence that focuses on automating repetitive tasks and improving efficiency.

Correct! Generative AI goes beyond analysis and can generate new content across multiple modalities.

check

4.

Which of the following describes a multimodal gen AI application?

Using Imagen to create an image for a website.

**Using gen AI to analyze customer sentiment in video testimonials and survey data.**

Using Gemini in Gmail to write an email.

Using NotebookLM Business to summarize a financial report.

Correct! This is a multimodal application because it combines video analysis (visual and audio) with text analysis (survey data) to understand customer sentiment.

1.

Which of the following is NOT a key feature of foundation models?

Trained on diverse data.

Flexible to support various use cases.

Adaptable to new domains and tasks.

**Specialized to specific tasks.**

Correct! Foundation models are designed to be more general purpose and adaptable to various tasks.

check

2.

What is the primary difference between foundation models and traditional AI models?

Foundation models cannot be adapted to new tasks, while traditional models can.

Foundation models are trained on specific data for a single task, while traditional models are trained on diverse data for various tasks.

**Foundation models are trained on massive amounts of diverse data for various tasks, while traditional models are trained on specific data for a single task.**

Foundation models are only trained on text data, while traditional models use images and code.

Correct! This is the key difference between foundation models and traditional AI models.

check

3.

Which of the following best defines a foundation model?

Traditional machine learning algorithms that rely on explicitly defined rules.

Small, very specialized AI models trained on narrow datasets in order to perform specific tasks.

**Large AI models trained on a vast quantity of data, capable of adapting to a variety of tasks**.

Hardware infrastructure used to train and deploy AI models.

Correct! Foundation models, including LLMs, are trained on massive datasets and can be adapted to various applications. They differ from traditional AI models that are designed for specific tasks.

check

4.

How do foundation models and prompt engineering work together to create value in generative AI?

Prompt engineering trains foundation models on specific tasks, allowing them to generate highly specialized content and insights.

**Foundation models offer a vast knowledge base, and prompt engineering guides the model to use this knowledge in responses.**

Foundation models provide the computing power for generative AI, while prompt engineering directs that power to complete specific tasks.

Foundation models ensure the ethical use of generative AI, while prompt engineering focuses on improving the quality and creativity of outputs.

Correct! Foundation models are trained on massive amounts of text and code, giving them a broad understanding of the world and how language works. Prompt engineering involves crafting specific instructions and questions to guide the model's use of this knowledge to generate the desired output.

check

5.

What is the purpose of a prompt in the context of foundation models?

**To provide input to the model and trigger an output**.

To evaluate the model's performance.

To fine-tune the model for a specific task.

To train the model on new data.

Correct! This is the main function of a prompt. It's a specific instruction or question given to the model to generate a response.

1.

What is the recommended approach for businesses to effectively implement gen AI?

Focus on a top-down strategy driven by executives.

**Combine a top-down strategic vision with bottom-up input from teams.**

Adopt the latest gen AI trends.

Focus on a bottom-up approach with no overarching plan.

Correct! This multi-directional approach ensures alignment with business goals while leveraging the insights and creativity of teams who will be using the technology.

check

2.

Which of the following is a benefit of using Google Cloud for gen AI development?

**It provides comprehensive services, including scalable infrastructure, enterprise-grade governance and security.**

It offers a limited set of pre-trained AI models for specific industry use cases.

It focuses primarily on individual developer tools without robust governance or scalability features.

It requires extensive upfront hardware investments for running complex AI workloads.

Correct! Google Cloud's AI platform encompasses the entire AI lifecycle, from data ingestion and preparation to model building, deployment, and monitoring, all supported by a scalable and secure infrastructure suitable for enterprise needs.

check

3.

Why is it important for mid-level managers and individual contributors to be involved in gen AI adoption?

They should avoid any involvement in gen AI experimentation.

**Their proximity to workflows allows them to identify impactful gen AI solutions.**

Their focus should be solely on technical implementation details.

They have limited understanding of daily operations and user needs.

Correct! Because they work closely with processes and users, they are well-positioned to identify areas where gen AI can make a real difference.

check

4.

What is a key advantage of using Google Cloud's gen AI ecosystem for businesses?

It assures businesses that they will achieve immediate and widespread success in all of their generative AI endeavors.

It eliminates the need for companies to cultivate internal gen AI knowledge and experimentation across their organizations.

**It allows businesses to leverage Google's AI advancements without starting from scratch.**

It forces each and every business to become an "AI-first" company.

Correct! Google Cloud's gen AI ecosystem provides businesses with access to state-of-the-art AI models, tools, and frameworks, allowing them to quickly and easily implement AI solutions without having to develop everything from scratch.