

Quiz [watsonx.ai gen AI Models L2]

You must receive a score of 75% or higher on the quiz to complete the course. [Back](#)

Next

Started on	Monday, December 30, 2024, 1:28 PM
State	Finished
Completed on	Monday, December 30, 2024, 1:47 PM
Time taken	19 mins 8 secs
Feedback	Congratulations, you passed the quiz!

Question **1**

Correct

Points out of 1.00

What tool is available to clients if they wanted to prompt tune or fine-tune the model?

- ☐ watsonx.ai Chat
- ☐ PromptLab
- ☐ watsonx.ai Freeform
- ☒ Tuning Studio ✓

Question **2**

Correct

Points out of 1.00

[Back](#)[Next](#)

What free of charge way sellers can team with IBM Consulting to allow clients to understand IBM's watsonx.ai Point of View (POV) better?

- ☐ Pilot Program
- ☐ Garage design session
- ☐ IBM consulting currently does not offer any service of this kind
- ☒ Half day strategy briefing session ✓

Question **3**

Correct

Points out of 1.00

What's the difference between smaller foundation models like Granite compared to a larger model like ChatGPT?

- ☐ IBM's smaller models are **INCAPABLE** of being used as a Chatbot due to the smaller size not allowing the model have chatting functionality
- ☐ The primary differentiator of IBM watsonx models is their ability to handle any task **WITHOUT** specialization or fine-tuning, making them a one-size-fits-all solution for clients
- ☐ IBM's smaller models **CANNOT** reach the same level of performance (accuracy) of larger models on specific use cases after fine-tuning due to the limitations in the models compute power
- ☒ Smaller models are trusted and target for specific business use cases, they perform (accuracy) on par and sometimes better than larger models after fine-tuning with lower cost and better efficiency ✓

Question **4**

Incorrect

Points out of 1.00

[Back](#)[Next](#)

What is IBM's differentiating approach to Foundation Models?

- ☒ To provide as many possible model variations as possible ✗
- ☐ To provide the largest models possible that is general-purpose
- ☐ Performant, Cost-effective, Open, and Trusted
- ☐ Large, Targeted, Cost-effective, and Performant

Question **5**

Correct

Points out of 1.00

What are the different capabilities of Mistral models supported in watsonx platform?

- ☐ Multilingual, Coding, RAG (Retrieval-Augmented Generation), Audio Processing
- ☐ Classification, Summarization, Generation, Image Processing
- ☐ Generative Capabilities, Computer Vision, Audio Processing, Coding
- ☒ Generative Capabilities, Multilingual, Coding, RAG (Retrieval-Augmented Generation) ✓

Question **6**

Correct

Points out of 1.00

[Back](#)[Next](#)

What is the main difference between the Llama models and Flan models?

- ☒ The Llama models are designed for a wide range of Natural Language capabilities (e.g. chat), whereas the Flan models are capable of handling a wide variety of instruction based tasks ✓
- ☐ The Llama models are designed for Computer Vision tasks such as Text-to-Image generation, and the Flan models are designed for Natural Language Processing
- ☐ The Llama models are designed for Natural Language Processing, whereas the Flan models are designed for Computer Vision tasks
- ☐ The Llama models are designed for handling a wide variety of instruction based tasks, whereas the Flan models are designed for Natural Language tasks in general

Question **7**

Incorrect

Points out of 1.00

[Back](#)[Next](#)

What is Hugging Face?

- ☒ Hugging Face is a data governance service, IBM has a ✗ partnership with Hugging Face providing OEM (Original Equipment Manufacturer) data governance products
- ☐ Hugging Face is a prompt tuning platform, where clients are able to prompt tune and test models
- ☐ Hugging Face is a prompt engineering platform, where clients are able to prompt engineer and test models
- ☐ Hugging Face is a machine learning platform, IBM has a partnership with Hugging Face to integrate open-source models into watsonx

Question 8

Correct

Points out of 1.00

[Back](#)[Next](#)

A client asks why does IBM have its own models and open-source models, how do you respond?

- ☐ IBM models are **ONLY** used for natural human language processing tasks, whereas open-source models are used for tasks such as code generation, image processing and audio processing
- ☐ IBM models **ALL** have a size below 25b parameters, whereas the open-source models are 70b parameters and above, allowing clients to have a wide variety of selection, so clients are **NOT** locked into any particular model and have the freedom to choose
- ☐ Open-source models are less trusted by clients due to lack of robust data governance and transparency in data curation and model training. Therefore IBM provides its own models for clients that priorities trustworthiness, and open-source models for clients that don't
- ☒ IBM's model strategy emphasizes providing clients with a range of options and flexibility. By partnering with Hugging Face, IBM offers access to a broad selection of state-of-the-art open-source models, which allows clients to choose models that best fit their specific use cases ✓

Question **9**

Correct

Points out of 1.00

[Back](#)

[Next](#)

What is the foundation model that IBM developed?

- ☐ PaLM
- ☐ Llama
- ☐ T5
- ☒ Granite ✓

Question 10

Incorrect

Points out of 1.00

[Back](#)[Next](#)

A client is thinking of prompt tuning one larger model to handle all of their use cases, and the client believes that this will improve performance and decrease costs. What is the **CORRECT** response? (Answers below are comparing larger models with ≥ 80 billion parameters and smaller models with < 80 billion parameters)

- ☐ Prompt tuning a large language model will be costly due to the models large parameter count, but this will only be a one time cost. The model will perform well on all use cases and would in the long term save money for the client
- ☒ A larger model can handle multiple use cases well out-of-the-box without any additional prompt tuning. Clients will achieve excellent results for every specific task out-of-the-box ✗
- ☐ Prompt tuning a large language model will be costly due to the models large parameter count, and the models performance may **NOT** be optimal for all use cases. IBM's strategy is to provide multiple smaller, more targeted models that perform extremely well on its specific use case, which are easier to train and prompt tune to generate better responses
- ☐ Prompt tuning a large language model will **NOT** be costly compared to smaller models and will perform well on all use cases; it is the best route for the client to choose if they are looking for a single, all-encompassing solution that eliminates the need for multiple specialized models

Question **11**

Correct

Points out of 1.00

[Back](#)[Next](#)

Are all models supported for the Custom Foundation Models feature to watsonx.ai?

- ☐ No, currently clients are only able to import variations of the Llama model architecture into the watsonx platform
- ☐ Yes, clients are able to import any model from HuggingFace and custom made models into the watsonx platform
- ☐ Yes, clients are able to import any model from HuggingFace into the watsonx platform, but not custom made models
- ☒ No, currently only variations of foundational models tested and released in watsonx.ai are supported ✓

Question **12**

Correct

Points out of 1.00

[Back](#)[Next](#)

A client wants to use a natural language that is currently **NOT** supported in watsonx.ai models, and want to bring a pre-existing open source model into watsonx.ai, what is the **CORRECT** response?

- ☐ The watsonx platform only supports languages that are pre-integrated into its models, so the client would need to use a different platform that supports their specific language needs
- ☐ The granite models have add-ons for more language support that clients are able to choose
- ☒ The client is also able to bring their own models to the watsonx platform using the Custom Foundation Models feature ✓
- ☐ No, clients are only able to use the languages supported by the Granite models

Question **13**

Correct

Points out of 1.00

[Back](#)[Next](#)

What is chat alignment?

- ☐ Technique used to improve the speed of language models by aligning the processing power of multiple GPUs (Graphic Processing Units)
- ☐ A method of ranking chat platforms based on their user engagement metrics
- ☒ Process of encoding human values and goals into large language models to make them as helpful, safe, and reliable as possible ✓
- ☐ Process of aligning words from one language to another to improve translation tasks

Question **14**

Correct

Points out of 1.00

What is generative AI used for?

- ☒ To generate new content, such as reports, emails, translation, and more ✓
- ☐ To be used **ONLY** for summarization tasks, and classification tasks
- ☐ To be used for highly specialized medical diagnosis
- ☐ To make ethical and moral decisions for enterprises

Question **15**

Correct

Points out of 1.00

[Back](#)[Next](#)

Which watsonx.ai feature allows clients to experiment with prompt engineering against foundation models?

- ☒ PromptLab ✓
- ☐ Tuning Studio
- ☐ watsonx API (Application Programming Interface) / SDK (Software Development Kit)
- ☐ InstructLab

Question **16**

Incorrect

Points out of 1.00

A client asks what the difference is between Traditional AI models and Foundation Models, how do you respond?

- ☐ Traditional AI models require supervised training on labeled data, Foundation Models use self-supervised training on a large pool of unlabeled data and do **NOT** require labeled data
- ☒ Traditional AI models require a large amount of labeled data, whereas Foundation Models do **NOT** require any labeled data ✗
- ☐ Open-source models are less trusted by clients due to lack of robust data governance and transparency in data curation and model training
- ☐ Traditional AI models are only used for classification tasks, whereas Foundation Models are only used for summarization tasks

Question **17**

Correct

Points out of 1.00

[Back](#)[Next](#)

What is the purpose of fine-tuning foundation models?

- ☐ To reduce the size and complexity of the model, allowing for faster inference
- ☐ To make the models more generalizable and adaptable
- ☐ To decrease the training time of the models
- ☒ To specialize the models on specific tasks or use cases ✓

Question **18**

Incorrect

Points out of 1.00

To ensure that models train on trusted data, what are the two essential stages that data must go through before training?

- ☒ Data quality assessment, Data language detection ✗
- ☐ Data deduplication, Data preprocessing
- ☐ Data filtering, Data deduplication
- ☐ Data acquisition, Data preprocessing

Question **19**

Correct

Points out of 1.00

[Back](#)[Next](#)

What are the different ways for clients to increase performance (accuracy) of models on a specific use case?

- ☒ Fine-tuning, Prompt Engineering, Prompt Tuning ✓
- ☐ Clients **CANNOT** increase the performance of models on specific use cases since the model weights are unchangeable
- ☐ Increasing computational resources
- ☐ Clients can increase performance by providing the model with generic data

Question **20**

Correct

Points out of 1.00

What is the difference between Prompt Tuning and Fine-tuning?

- ☐ Fine-tuning does **NOT** update model weights and focuses on prompt design. Prompt tuning also does **NOT** update weights but uses input/output examples
- ☐ Fine-tuning does **NOT** update the model weights but improves results using input/output examples. Prompt tuning updates the model weights with a labeled dataset.
- ☐ Fine-tuning does **NOT** update the model weights but uses input/output examples. Prompt tuning also does **NOT** update the weights but focuses on designing efficient prompts
- ☒ Fine-tuning updates the model weights with a labeled dataset to improve results. Prompt tuning does **NOT** update the weights, but improves results by providing examples of input/output pairs. ✓