**CHAPTER 1**

**Describe Artificial Intelligence workloads and considerations**

*Artificial Intelligence (AI)* is computers thinking and acting in a way that simulates a human. AI is a technology that takes information from its environment and responds based on what it learns. The goal of AI is to create a machine that can mimic human behavior.

AI is more than learning—it is knowledge representation, reasoning, and abstract thinking. Machine learning (ML) is the subset of AI that takes the approach of teaching computers to learn for themselves, rather than teaching computers all that they need to know. ML is the foundation for modern AI. ML focuses on identifying and making sense of the patterns and structures in data.

ML is about machines’ reasoning and decision-making using software that learns from past experiences. ML allows computers to consistently perform repetitive and well-defined tasks that are difficult for humans to accomplish. Over the past few years, machine learning algorithms have proved that computers can learn tasks that are tremendously complicated for machines, demonstrating that ML can be employed in a wide range of scenarios and industries.

AI is now being embedded into the software you use today, sometimes without us realizing it. For example, Microsoft PowerPoint has a feature called Design Ideas that offers suggestions for themes and layouts for slides, and Microsoft Word offers suggestions to rewrite sentences to improve clarity.

This chapter is an overview of how Artificial Intelligence and machine learning can be used in different scenarios and industries ethically.

**Skills covered in this chapter:**

* [Skill 1.1: Identify features of common AI workloads](https://learning.oreilly.com/library/view/exam-ref-ai-900/9780137358076/ch01.xhtml#ch01lev1sec1)
* [Skill 1.2: Identify guiding principles for Responsible AI](https://learning.oreilly.com/library/view/exam-ref-ai-900/9780137358076/ch01.xhtml#ch01lev1sec2)

**Skill 1.1: Identify features of common AI workloads**

Artificial Intelligence is software that mimics human behaviors and capabilities. Today, software can use AI to automatically detect and predict actions that machines, and humans, should take.

Microsoft Azure provides a set of services for Artificial Intelligence and machine learning that you can utilize to create your own intelligent solutions. Microsoft Azure AI Fundamentals is a certification that requires you to have entry-level knowledge of AI and ML concepts and knowledge of the related Microsoft Azure services.

This skill covers how to:

* Describe Azure services for AI and ML
* Understand Azure Machine Learning
* Understand Azure Cognitive Services
* Describe the Azure Bot Service
* Identify common AI workloads

**Describe Azure services for AI and ML**

There is a wide and rapidly growing series of services in Azure for AI and ML. There are three services that are the focus of the AI-900 Fundamentals exam:

* **Cognitive Services**   A set of prebuilt services that you can easily use in your applications.
* **Azure Bot Service**   A service to help create and deploy chatbots and intelligence agents.
* **Azure Machine Learning**   A broad range of tools and services that allow you to create your own custom AI.

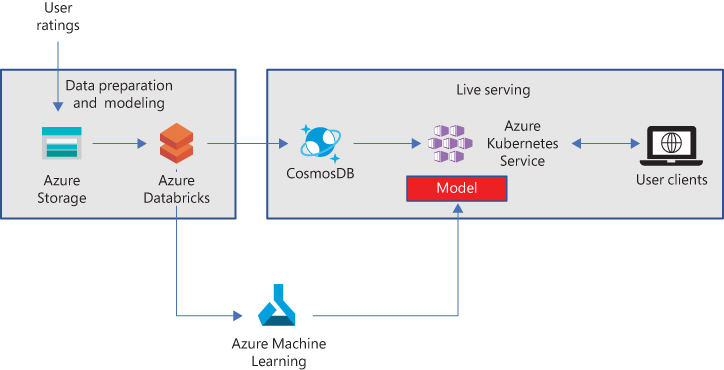
We will be exploring some of the features and capabilities of these three services in this book. However, these services do not work in isolation; they utilize many other Azure services to deliver solutions such as the following:

* Storage
* Compute
* Web Apps
* HD Insights
* Data Factory
* Cosmos DB
* Azure Functions
* Azure Kubernetes Service (AKS)

**Example ML architecture**

To explain how Azure services support Azure Machine Learning, consider the scenario of a company that wants to provide recommendations to its users. By providing personalized targeted recommendations, users will more likely purchase more of their products and user satisfaction will increase.

[Figure 1-1](https://learning.oreilly.com/library/view/exam-ref-ai-900/9780137358076/ch01.xhtml#ch01fig01) shows an example of an ML architecture to support recommendations.



**FIGURE 1-1** Example ML architecture

**Understand Azure Machine Learning**

Azure Machine Learning is the foundation for Azure AI. In Azure Machine Learning, you can build and train AI models to make predictions and inferences.

Training a machine learning model requires lots of data and lots of computing resources. Azure provides many services for preparing data and then analyzing the data.

Azure ML is a platform for training, deploying, and managing machine learning models. We will cover machine learning in more detail in [Chapter 2](https://learning.oreilly.com/library/view/exam-ref-ai-900/9780137358076/ch02.xhtml#ch02).

**Machine learning model types**

Machine learning makes use of algorithms to identify patterns in data and take action. The types of machine learning models created from the outputs of the algorithms are as follows:

* **Anomaly Detection**   Finds unusual occurrences.
* **Classification**   Classifies images or predicts between several categories.
* **Clustering (including K-Means Clustering)**   Discovers structures.
* **Regression**   Predicts values.

***NOTE*   MODEL TYPES**

Understanding the differences between these model types is foundational knowledge for this exam.

For each model type, you will need to understand the following:

* How they work
* What they do
* What they can be used for
* The metrics they produce

We will address these points later in the book.

**Understand Azure Cognitive Services**

*Cognitive Services* is a suite of prebuilt AI services that developers can use to build AI solutions. Cognitive Services meets common AI requirements and allow you to add AI to your apps more quickly with less expertise.

Cognitive Services are machine learning models trained by Microsoft with massive volumes of data. While you can build your own custom AI models to perform the same analyses, Cognitive Services allow you to meet many AI requirements easily around processing images and analyzing text. However, Cognitive Services only address a subset of AI requirements. You can create your own machine learning models to meet more complex and specific needs.

Cognitive Services are available as a set of REST APIs for the following capabilities:

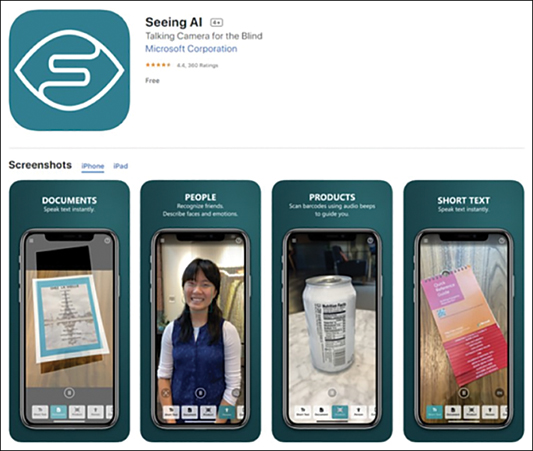
* Computer vision, including images and video
* Decision, including Anomaly Detector
* Language, including text analysis and Language Understanding
* Speech, including translation
* Intelligent search, including knowledge mining

Cognitive Services have a broad and growing range of features and capabilities. The Azure AI Fundamentals exam focuses on two of these capabilities:

* Image processing
* Natural Language Processing (NLP)

A great example of the use of Cognitive Services is the free Seeing AI app that uses these two capabilities. Designed for visually impaired people, this app turns the physical environment into an audible experience, locating faces, identifying objects, and reading documents.

[Figure 1-2](https://learning.oreilly.com/library/view/exam-ref-ai-900/9780137358076/ch01.xhtml#ch01fig02) shows the Seeing AI app in the App Store with sample screenshots and an overview of its features.



**FIGURE 1-2** The Seeing AI app

We will be looking at the computer vision and Natural Language Processing services of Azure Cognitive Services in [Chapters 3](https://learning.oreilly.com/library/view/exam-ref-ai-900/9780137358076/ch03.xhtml#ch03) and [4](https://learning.oreilly.com/library/view/exam-ref-ai-900/9780137358076/ch04.xhtml#ch04).

**Describe the Azure Bot Service**

The *Azure Bot Service* is a cloud-based platform for developing, deploying, and managing bots. Azure Bot Services provide the capabilities known as conversational AI.

Conversational AI is where the computer simulates a conversation with a user or customer. Conversational AI has extended beyond simple chatbots to intelligence agents and virtual assistants like Cortana.

There are two conversational AI services included in the Azure AI Fundamentals exam:

* **QnA Maker**   A tool to build a bot using existing support and other documentation.
* **Azure Bot Service**   Tools to build, test, deploy, and manage bots.

Both QnA Maker and the Azure Bot Service leverage the Language Understanding (LUIS) service in Cognitive Services.

We will look at bots in more detail in [Chapter 5](https://learning.oreilly.com/library/view/exam-ref-ai-900/9780137358076/ch05.xhtml#ch05).

**Identify common AI workloads**

There are many use cases for AI. Here we will look at some common AI workloads, describing their features and providing some examples for their use.

Image ***EXAM TIP***

You will be asked in the exam to pick the correct AI workload from an example given, or to identify an example for a specified AI workload. You should make sure you understand the use cases for each workload.

**Prediction**

*Prediction*, or forecasting, is where the computer identifies patterns in your historical data, and through machine learning associates the patterns with outcomes. You can then use the prediction model to predict the outcome for new data.

Types of predictions include the following:

* **Binary prediction**   There are two possible outcomes for a question—yes/no or true/false.
* **Multiple outcome prediction**   The question can be answered from a list of two or more outcomes.
* **Numerical prediction**   The question is answered with a continuous number, not explicitly limited to a specific range of values.

Prediction models can be used in many varied scenarios. For example, prediction models are found in financial services when evaluating credit applications or evaluating risk. You can use a prediction model to forecast customer churn rate, or to forecast the number of support calls that will be received, or if an opportunity will be converted to a sale.

Following are some questions that could be answered using prediction:

* Will this customer be approved for a $10,000 credit limit?
* Will this flight be delayed?
* Will this flight be on-time, slightly delayed, or very late?
* How many minutes will this flight be delayed by?

**Anomaly detection**

*Anomaly detection* analyzes data over time and identifies unusual changes, often for real-time data streams.

Anomaly detection, also known as outlier detection, can find dips and spikes that may indicate a potential issue. Such issues are hard to spot when analyzing aggregate data, as the data points are hidden in the vast volume of data.

Anomaly detection can identify trend changes. Typically, the anomaly will indicate problems such as a sticking valve, payment fraud, a change in the level vibration on a bearing, or errors in text.

Anomaly detection enables pre-emptive action to be taken before a problem becomes critical or adversely affects business operations.

***NOTE*   ANOMALY DETECTION**

Anomaly detection does not predict when a failure will occur; for this, you should use a prediction model.

There are several algorithms that can be used for anomaly detection. The Azure Anomaly Detector service selects the best algorithm based on the data, making Anomaly Detector easy to use with a single API call. Anomaly Detector can also run in a Docker container, so it can be deployed at the edge on devices themselves.

Following are some examples of the use of anomaly detection:

* **Monitoring IoT devices**   Checking the telemetry from devices in real-time to find anomalies.
* **Fault detection in electricity systems**   Identifying spikes and dips in the electrical supply.
* **Computer network traffic access attacks**   Detecting unusual network activity both inside the network and on the network perimeter.
* **Financial system fraud**   Identifying potential fraudulent payments from the patterns of payments.
* **Hospital infection**   Recognizing high mortality rates from a particular infection than from other causes of death.
* **Crowd surveillance**   Identifying changing crowd behaviors in complex situations.

**Computer vision**

*Computer vision* is the processing of still images and video streams. Computer vision can interpret the image and provide detail and understanding about the image in computer-readable form.

Computer vision can determine if the image contains a specific object (object detection) and can extract details from the image, such as colors or text.

Computer vision can:

* Describe an image
* Categorize an image
* Tag an image
* Detect objects
* Detect faces
* Identify brands and products
* Identify famous people
* Identify landmarks
* Extract text

There are many potential uses for computer vision:

* **Reading text and barcodes**   Reading and identifying text and barcodes every day is not an easy job for a human.
* **Product assembly**   Manufacturers can ensure that assembly of products and components are strictly adhering to standards. For example, pharmaceutical manufacturers can inspect bottles in 360 degrees to ensure correct packaging. They can examine critical features of packaged bottles like cap seal, position, and labels.
* **Monitoring the length of the queue in retail stores**   Retailers can determine if more checkout staff are required.
* **Detecting abnormalities in health scans**   Computer vision can sift the scans, freeing up scarce highly skilled diagnostic staff to analyze the more complex scanned images.

**Natural Language Processing**

*Natural Language Processing (NLP)* is the analysis of text to extract information in a form that can be used by a computer.

Natural Language Processing interprets spoken and written text. NLP can analyze the text to determine the language used, determine the sentiment expressed, extract key phrases, and identify key entities and actions.

NLP can be used in many scenarios, such as the following:

* **Decision support**   Assisting users in making decisions from unstructured and rapidly changing data.
* **Bots and intelligent agents**   Understanding the user’s question without having been trained in the exact words or phrasing so that a response can be formulated.
* **Translate commands into actions**   Understanding the intent in the user’s request and performing the requested action.
* **SPAM detection**   Classifying emails based on the text in the subject and body.
* **Monitoring news**   Extracting relevant news from different feeds that are relevant to the user.
* **Grammar checkers**   Flagging words or phrases and attempting to offer improvements.

**Knowledge mining**

*Knowledge mining* is the process of extracting key insights from structured and unstructured data sources.

Knowledge mining uses a combination of AI services, including Azure Cognitive Search, to extract meaning and relationships from large amounts of information. This information can be held in structured and unstructured data sources, documents, and databases. Knowledge mining uncovers hidden insights in your data.

Microsoft provides a Knowledge Mining Solution Accelerator to help you ingest different data and document source, enrich, and index the data, and provides a user interface to explore the results.

**Conversational AI**

*Conversational AI* is the process of building AI agents to take part in conversations with humans. Conversational AI is commonly experienced by humans as chatbots on websites and other systems.

AI agents (bots) engage in conversations (dialogs) with human users. Bots use natural language processing to make sense of human input, identify the actions the human wants to perform, and identify the entity on which the actions are to be performed. Bots can prompt the human for the information required to complete a transaction.

There are three common types of bot that you may encounter:

* Webchat
* Telephone voice menus (IVR)
* Personal Digital Assistants

You can use bots in many scenarios, including the following:

* **Customer support**   Handling common questions and inquiries and providing a method for escalating to a human agent.
* **FAQs**   Providing an interactive question and answer service to users over the web or in an app.
* **Online ordering**   Supporting customer ordering and answering common questions about the products or delivery.
* **Travel reservation and booking**   Assisting the customer in finding flights and accommodations that match their requirements and making the booking on their behalf.
* **Healthcare triage**   Guiding the user when triaging a patient to check for symptoms.

**Skill 1.2: Identify guiding principles for Responsible AI**

Responsible AI is the provision of AI-based solutions to difficult problems without any unintended negative consequences.

This section covers the six principles of Responsible AI. By following these principles, you can ensure that your AI-enhanced solutions will put people first.

Image ***EXAM TIP***

Make sure you can describe each principle of Responsible AI in a single sentence.

This skill covers how to:

* Describe the Fairness principle
* Describe the Reliability & Safety principle
* Describe the Privacy & Security principle
* Describe the Inclusiveness principle
* Describe the Transparency principle
* Describe the Accountability principle
* Understand Responsible AI for Bots
* Understand Microsoft’s AI for Good program

***NOTE*   PRINCIPLES FOR RESPONSIBLE AI**

These principles should be taken into account when creating solutions that use AI. Not every principle will apply to every requirement, but you should check your solution against each principle to see if it applies.

**Describe the Fairness principle**

The Fairness principle of Responsible AI is concerned with treating all people fairly and reducing unfairness.

A Responsible AI-based solution must operate without giving any unfair advantage to, withholding opportunities from, or allocating resources to a specific group of people. There should be no bias regarding a person’s gender, or any other characteristic.

AI systems can reinforce existing stereotypes and underrepresentation. If correctly addressed, AI systems can reduce unfairness.

An AI model should be interpreted to quantify the extent of how data influences the model’s prediction to help eliminate bias.

The Fairness principle means:

* Eliminating bias for gender, age, or ethnicity
* Removing unfair advantages
* Preventing unfair allocation of resources
* Preventing the withholding of information
* Mitigating bias at each stage of developing and deploying AI systems

Following are some examples where the Fairness principle can have a significant impact:

* Bank loans/credit decisions
* Hiring staff
* Criminal justice system

**Describe the Reliability & Safety principle**

Reliability & Safety requires the rigorous testing of an AI-based system’s functionality and deployment to ensure that it works as expected and eliminates potential risk to human life.

Features of Reliability & Safety are:

* Rigorous testing
* Works as expected
* Eliminates threat of harm to human life

Areas where Reliability & Safety must be applied are as follows:

* Autonomous vehicles
* Healthcare diagnosis

**Describe the Privacy & Security principle**

Privacy & Security requires that an AI-based system should be secure and respect privacy. AI-based systems typically operate on high volumes of data, including personal data that should not be disclosed.

The reliance on data used in training the model, and new data used for predictions, is subject to privacy rules.

AI systems that run on a user’s device should not leak the user’s data. One way to achieve this is to run the AI processing on the device and not transfer the personal data to a cloud service.

To be responsible, you should:

* Respect privacy.
* Be secure.
* Avoid disclosing personal data.

Concerns covered by Privacy & Security are as follows:

* **Data origin**   Where has the data come from—is it user or public data?
* **Data use**   Validating that the data you are using has not been corrupted or interfered with.

**Describe the Inclusiveness principle**

Inclusiveness requires AI-based solutions to empower everyone and supply benefits to all parts of society, regardless of gender, physical ability, ethnicity, sexual orientation, or any other factors.

Inclusiveness means:

* Empowering everyone
* Engaging all communities in the world
* Intentionally designing for the inclusivity principle

**Describe the Transparency principle**

Transparency is the principle that AI-based solutions should be understandable. Users should be aware of the purpose of an AI-based system, how it operates, its scope, and its limitations.

Transparency is essentially about gaining the trust of users.

Transparency means:

* Defining the purpose of the use of AI in your solution
* Defining the scope of AI in your solution
* Stating the limitations of AI in your solution

AI-based solutions should be understandable. You must be open about how and why you are using AI to users and other stakeholders.

Transparency also means that people can understand the behavior of an AI system; the outputs of an AI algorithm should be able to be interpreted. For example, in some industries, you must be able to explain to regulators how the AI algorithm has generated its results.

**Describe the Accountability principle**

Accountability requires the people involved in designing and developing AI-based solutions to operate within a clear governance framework due to the impact AI can have on the world.

Accountability requires the people involved in designing and developing AI-based solutions to follow clearly defined ethical policies and legal standards.

Accountability requires that you have:

* Governance framework
* Ethical policies
* Legal standards

Civil liberties are an area where accountability is a crucial factor when determining if and how to use AI—for example, the use of facial recognition.

***NEED MORE REVIEW?*   RESPONSIBLE AI**

For more information on Responsible AI, see <https://www.microsoft.com/ai/responsible-ai>.

**Understand Responsible AI for Bots**

To illustrate what Responsible AI means, let’s look at building a chatbot.

When building a chatbot, you need to consider many of the principles outlined previously. The Transparency principle means that:

* A customer should know they are interacting with a bot.
* The purpose of the bot should be clear.
* The limitations of the bot should be stated.

You can meet this requirement by having a clear welcome message, stating that a bot is responding. It should be possible to seamlessly transfer to a human agent.

Bots operate best when they have a clear purpose. Bots that attempt to handle every possible scenario often perform poorly. The scope of a bot should be reduced to a clearly defined purpose.

***NEED MORE REVIEW?*   RESPONSIBLE CONVERSATIONAL AI**

Watch this demo of Responsible Conversational AI at <https://aidemos.microsoft.com/responsible-conversational-ai/building-a-trustworthy-bot>.

**Understand Microsoft’s AI for Good program**

AI for Good is a Microsoft program that puts AI technology, cloud software, and other resources into the hands of those working to create a more sustainable and accessible world.

AI for Good is enabling advances in healthcare, environmental protection, humanitarian action, cultural heritage, and other areas to make a better world for everyone.

***NEED MORE REVIEW?*   AI FOR GOOD**

For more information on AI for Good, see <https://www.microsoft.com/ai/ai-for-good>.

**Chapter summary**

In this chapter, you learned some of the general concepts related to Artificial Intelligence. You learned about the features of common AI workloads, and you learned about the principles of Responsible AI. Here are the key concepts from this chapter:

* Artificial Intelligence is a technology that mimics the human brain and uses machines to complete complex tasks that humans find difficult to do.
* Machine learning uses algorithms to discover patterns and structures in existing data, building a model that can be used to take actions (like classify and predict) on unseen data.
* Cognitive Services is a set of prebuilt AI models, trained by Microsoft, that you can use without being a data scientist to add intelligence into your applications.
* Azure Machine Learning is a set of tools and services that you can use to create custom AI models using your own data.
* Azure Bot Service allows you to build and deploy chatbots and intelligence agents.
* Regression is the type of machine learning that predicts values from historical data.
* Classification is the type of machine learning that classifies images or predicts between several distinct categories.
* Clustering discovers structures in data, identifying groups based on similarities in the data.
* Anomaly detection finds unusual occurrences or events in time-series data. Anomaly detection can identify outliers in the data.
* Prediction is the model that predicts outcomes for new data. Predictions are based on historical data. You can have predictions based on regression, classification, and clustering algorithm. Prediction is also known as forecasting.
* Computer vision is the analysis of images and video to extract information that can be used by computers. Computer vision can interpret the contents of the image to classify the image, detect objects in the image, and analyze and describe the image.
* Natural Language Processing is the analysis of speech and text to extract the meaning and intent of words in a way that can be used by computers. Natural Language Processing can interpret text. Natural Language Processing can analyze the text to determine the language used, determine the sentiment expressed, extract key phrases, and identify key entities and actions.
* Conversational AI is used to create applications where AI agents engage humans in conversations (dialogs). Conversational AI is commonly experienced by humans as chatbots on websites.
* Knowledge mining uses a combination of AI services to extract meaning and relationships from large amounts of information. This information can be held in structured and unstructured data sources, documents, and databases. Knowledge mining uncovers hidden insights in your data.
* Responsible AI is the use of AI in solutions without having unintended negative impacts.
* Fairness is the principle that AI-based systems should treat all people fairly and reduce bias.
* Reliability & Safety is the principle that requires rigorous testing of AI-based systems to eliminate harm to human life.
* Privacy & Security is the principle that AI-based systems should be secure and respect privacy of personal data.
* Inclusiveness is the principle that AI-based solutions should empower everyone and supply benefits to all parts of society, regardless of any characteristics or factors.
* Transparency is the principle that AI-based solutions should be understandable, or interpretable. Transparency requires that AI-based systems have a defined purpose and scope and are clear on the limitations of the AI in the solution.
* Accountability is the principle that everyone at every stage in the development of AI-based systems is accountable for the impact that system may have.

**Thought experiment**

In this thought experiment, demonstrate your skills and knowledge of the topics covered in this chapter. You can find the answers in the section that follows.

You work for Contoso Medical Group (CMG), and your management is interested in using AI in your applications and operations. CMG manages and monitors drug trials, evaluating the efficacy of the treatments.

The CMG IT department is resource-constrained, and they do not have data scientists or skilled AI developers available.

Having timely and accurate responses from patients improves the accuracy of the analysis performed. CMG has created an app to capture and track a patient’s daily symptoms. CMG has recently added the capability of the app to take pictures to capture skin conditions. CMG is unable to analyze the images due to the volume of images being captured. CMG is concerned about the amount of data storage for these images, as well as controlling access to the images.

CMG receives a lot of patient history and prescription records that are keyed into CMG’s computer systems. These paper records are important information used to track a patient’s response to drugs and treatments.

The support department is unable to handle the many inquiries CMG receives. Customers are receiving inconsistent responses depending on whom they speak to and how they are accessing customer support, whether by phone, web, or mobile app.

Your manager has come to you asking for solutions that address these issues. Whatever solution you offer must consider that the medical data in this application is covered under HIPAA, and your manager wants CMG to retain all control of the data. Your manager also wants to carefully control costs.

You have decided that CMG can use AI, but there are several issues that you need to resolve before proceeding.

Answer the following questions:

1. Which AI workload should you use for the customer support department?
2. Which principle of Responsible AI should you employ to gain the trust of users in your bot?
3. Which AI workload should you use to analyze the images for skin conditions?
4. How can you address the storage requirements for the images?
5. Which principle of Responsible AI protects a patient’s personal information?
6. Which AI workload could identify adverse reactions to a drug treatment?
7. Which principle of Responsible AI requires rigorous testing of your AI-based app?

**Thought experiment answers**

This section contains the solutions to the thought experiment. Each answer explains why the answer choice is correct.

1. Conversational AI will allow simple inquiries to be handled by an automated bot. You can create a chatbot for your website, create an assistant for customer service agents, and even enable a bot in a mobile app.
2. Transparency is the principle that AI-based solutions should be understandable. Users should be aware of the purpose of the AI-based system, how it operates, and its scope and limitations. A chatbot should clearly tell the user that it is a bot and what it can and cannot do.
3. Computer vision allows you to analyze images. You can train computer vision with existing images to classify images to determine the type of skin complaint.
4. You can run AI in the app on the mobile device. The image will not need to be stored and will not leave the device. You will instead store the results of the classification along with other symptom data and discard the image.
5. Privacy & Security is concerned with keeping AI-based solutions secure and preventing personal data from being disclosed.
6. Anomaly detection can detect adverse reactions. It can detect where there is a change in trends and can detect unusual readings.
7. Reliability & Safety requires the rigorous testing of an AI-based system’s functionality and deployment to ensure that it works as expected and to eliminate potential risk to human life.