

The above diagram sums up the key concepts of Azure AI Fundamentals. This summary is what the entire 1st section is about. Along with this, you also need to have knowledge about the principles of responsible AI. Usually, questions based on definitions and assessing which principle is being followed are asked from this section. The following are the guiding principles:

1. **Fairness:** AI systems should treat all people fairly.
2. **Reliability and safety**: For example, consider an AI-based software system for an autonomous vehicle; or a machine learning model that diagnoses patient symptoms and recommends prescriptions. Unreliability in these kinds of system can result in a substantial risk to human life.
3. **Privacy and security:** AI systems should be secure and respect privacy.
4. **Inclusiveness**: AI systems should empower everyone and engage all parts of society, regardless of physical ability, gender, sexual orientation, ethnicity, or other factors.
5. **Transparency:** Users should be made fully aware of the purpose of the system, how it works, and what limitations may be expected.
6. **Accountability:**People should be accountable for AI systems and should ensure the solution meets ethical and legal standards that are clearly defined.

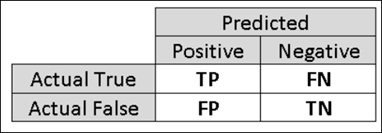
Apart from this, the process of anomaly detection should be understood well.

**2. Machine Learning**

* Difference between use cases of classification (labelled groups), regression (number prediction), Clustering( Unlabelled groups, based on similarity).
* Ability to distinguish between the following:

1. **Compute Instances**: Development workstations that data scientists can use to work with data and models.
2. **Compute Clusters**: Scalable clusters of virtual machines for on-demand processing of experiment code,
3. **Inference Clusters**: Deployment targets for predictive services that use your trained models.
4. **Attached Compute**: Links to existing Azure compute resources, such as Virtual Machines or Azure Databricks clusters.

* Automated Machine Learning — automated for supervised learning (Classification, Regression, Time-Series Forecasting)
* Best Model — Based on the transparency principle, this feature helps explain to the user why a particular model is chosen and how it complies with regulations and best practices.
* Metrics used for Classification (True Positive Rate / ROC Curve) and Regression (Mean Absolute Error, Root Mean Squared Error, Relative Squared Error, Relative Absolute Error, Coefficient of Determination)
* Confusion Matrix — How to read and identify True Positives, True Negatives, False Positives and False Negatives.



Confusion Matrix

* Model Evaluation — The Model evaluation module outputs a confusion matrix showing the number of true positives, false negatives, false positives, and true negatives, as well as ROC, Precision/Recall, and Lift curves.
* Feature Engineering — Using domain knowledge of the data to create features that help ML algorithms learn better. Scaling and normalization techniques are applied to facilitate feature engineering.
* Feature Selection — Selecting a subset of relevant, useful features to use in building an analytical model.

**3. Computer Vision Workloads**

* Difference between Image Classification, Object Detection, Semantic Segmentation, Image Analysis, Face Detection, Face Analysis, Face Recognition, Optical Character Recognition
* Custom Vision: service used to train only image classification and object detection using your own images.
* Face Service: service that enables you to build face detection and facial recognition solutions.
* Form Recognition: service used to extract information from scanned forms and invoices.

**4. Natural Language Processing (NLP) Workloads**

* Key phrase extraction — identifying the main talking points of the document(s).
* Entity recognition — returns a list of *entities* in the text that it recognizes. TAn entity is essentially an item of a particular type or a category; and in some cases, subtype. The type of entity may also be mentioned near it.
* Speech Recognition = Speech-to-text
* Speech Synthesis = Text-to-speech
* Utterances (what a user says), Entities (item to which an utterance refers) and Intents (purpose or goal)
* Language Understanding Intelligent Service (LUIS): used to train a language model that understands spoken or written commands.
* Translation of Speech is done by Speech and not Translator Text API

**5. Conversational AI Workloads**

* Different examples of Conversational AI — Eg: An automated-voice menu on the telephone.
* QnA Maker — It enables you to build a knowledge base. Questions and answers can be generated from an existing FAQ document or web page, imported from a pre-defined *chit-chat* data source or entered and edited manually.
* Azure Bot Service — It helps you to create, publish and manage bots.

***What is Conversational AI?***

The capability of a software "agent" to participate in a conversation.

**What is AI?**

AI is the creation of software that imitates human behaviours and capabilities. Key elements include:

Machine learning  
Anomaly detection  
Computer Vision  
Natural language processing  
Conversational AI

**What are the features and capabilities of Azure Machine Learning Service?**

**Automated machine learning**This feature enables non-experts to quickly create an effective machine learning model from data.

**Azure Machine Learning designer**  
A graphical interface enabling no-code development of machine learning solutions.

**Data and compute management**Cloud-based data storage and compute resources that professional data scientists can use to run data experiment code at scale.

**Pipelines**Data scientists, software engineers, and IT operations professionals can define pipelines to orchestrate model training, deployment, and management tasks.

**In which scenarios you use anomaly detection — a machine learning-based technique that analyzes data over time and identifies unusual changes?**

1. Monitor credit card transactions and detect unusual usage patterns that might indicate fraud.

2. An application that tracks activity in an automated production line and identifies failures.

3. A racing car telemetry system that uses sensors to proactively warn engineers about potential mechanical failures before they happen.

**Name the app based on Computer Vision?**

The **Seeing AI** app is a great example of the power of computer vision. Designed for the blind and low vision community, the Seeing AI app harnesses the power of AI to open up the visual world and describe nearby people, text and objects.

**What are the tasks that come under Computer Vision?**

**Image classification**Image classification involves training a machine learning model to classify images based on their contents. For example, in a traffic monitoring solution you might use an image classification model to classify images based on the type of vehicle they contain, such as taxis, buses, cyclists, and so on.

**Object detection**Object detection machine learning models are trained to classify individual objects within an image, and identify their location with a bounding box. For example, a traffic monitoring solution might use object detection to identify the location of different classes of vehicle.

**Semantic segmentation**Semantic segmentation is an advanced machine learning technique in which individual pixels in the image are classified according to the object to which they belong. For example, a traffic monitoring solution might overlay traffic images with "mask" layers to highlight different vehicles using specific colors.

**Image analysis**You can create solutions that combine machine learning models with advanced image analysis techniques to extract information from images, including "tags" that could help catalog the image or even descriptive captions that summarize the scene shown in the image.

**Face detection, analysis, and recognition**Face detection is a specialized form of object detection that locates human faces in an image. This can be combined with classification and facial geometry analysis techniques to infer details such as gender, age, and emotional state; and even recognize individuals based on their facial features.

**Optical character recognition (OCR)**Optical character recognition is a technique used to detect and read text in images. You can use OCR to read text in photographs (for example, road signs or store fronts) or to extract information from scanned documents such as letters, invoices, or forms.

**What are the Computer Vision services in Microsoft Azure?**

**Computer Vision**You can use this service to analyze images and video, and extract descriptions, tags, objects, and text.

**Custom Vision**Use this service to train custom image classification and object detection models using your own images.

**Face**The Face service enables you to build face detection and facial recognition solutions.

**Form Recognizer**Use this service to extract information from scanned forms and invoices.

**What are NLP services in Microsoft Azure?**

**Text Analytics**Use this service to analyze text documents and extract key phrases, detect entities (such as places, dates, and people), and evaluate sentiment (how positive or negative a document is).

**Translator Text**Use this service to translate text between more than 60 languages.

**Speech**Use this service to recognize and synthesize speech, and to translate spoken languages.

**Language Understanding Intelligent Service (LUIS**)  
Use this service to train a language model that can understand spoken or text-based commands.

**What are the Conversational AI services in Microsoft Azure?**

**QnA Maker**This cognitive service enables you to quickly build a knowledge base of questions and answers that can form the basis of a dialog between a human and an AI agent.

**Azure Bot Service**This service provides a platform for creating, publishing, and managing bots. Developers can use the Bot Framework to create a bot and manage it with Azure Bot Service - integrating back-end services like QnA Maker and LUIS, and connecting to channels for web chat, email, Microsoft Teams, and others.

**. What are the six guiding principles of responsible AI?**

**Fairness**AI systems should treat all people fairly. For example, suppose you create a machine learning model to support a loan approval application for a bank. The model should make predictions of whether or not the loan should be approved without incorporating any bias based on gender, ethnicity, or other factors that might result in an unfair advantage or disadvantage to specific groups of applicants.

**Reliability and safety**AI systems should perform reliably and safely. For example, consider an AI-based software system for an autonomous vehicle; or a machine learning model that diagnoses patient symptoms and recommends prescriptions. Unreliability in these kinds of system can result in substantial risk to human life.

**Privacy and security**AI systems should be secure and respect privacy. The machine learning models on which AI systems are based rely on large volumes of data, which may contain personal details that must be kept private. Even after the models are trained and the system is in production, it uses new data to make predictions or take action that may be subject to privacy or security concerns.

**Inclusiveness**AI systems should empower everyone and engage people. AI should bring benefits to all parts of society, regardless of physical ability, gender, sexual orientation, ethnicity, or other factors.

**Transparency**AI systems should be understandable. Users should be made fully aware of the purpose of the system, how it works, and what limitations may be expected.

**Accountability**People should be accountable for AI systems. Designers and developers of AI-based solution should work within a framework of governance and organizational principles that ensure the solution meets ethical and legal standards that are clearly defined.

# Describe fundamental principles of machine learning on Azure (30- 35%)

Practice questions based on these concepts

* Identify common machine learning types
* Describe core machine learning concepts
* Identify core tasks in creating a machine learning solution
* Describe the capabilities of no-code machine learning with Azure Machine Learning

**What is the process of machine learning regardless of the model?**

**Data Igestion**  
You need to get the data to train your model

**Data Pre processing**Identify the features that helps the model to predict and discarding others

**Data Cleaning**  
Fix any erros or remving the items which has erros

**Replacing Feature Values**  
find the replacement feature values if any missing. In this process you might use exisiting feature engineering to find the value

**Apply Algoritms**  
Apply alogorithms on this data for the processing until you are happy with the model pridictions

**Deploy Model**  
Fianlly you deploy your model into machine learning service so that applications can connect to it.

**How many kinds of Compute resources that data scientists can use to train their models?**

**Compute Instances**  
Development workstations that data scientists can use to work with data and models.

**Compute Clusters**:  
Scalable clusters of virtual machines for on-demand processing of experiment code.

**Inference Clusters**:  
Deployment targets for predictive services that use your trained models.

**Attached Compute**:  
Links to existing Azure compute resources, such as Virtual Machines or Azure Databricks clusters.

**40. What are the settings you need to create a compute instance?**

**Compute name**: enter a unique name

**Virtual Machine type**: CPU

**Virtual Machine size**: Standard\_DS2\_v2

**41. What are the settings you need to create a Compute Clusters?**

**Compute name**: enter a unique name

**Virtual Machine size**: Standard\_DS2\_v2

**Virtual Machine priority**: Dedicated

**Minimum number of nodes**: 2

**Maximum number of nodes**: 2**Idle seconds before scale down**: 120

**42. How do you make sure that to start the compute only when it is needed when creating computer clusters in the production environment?**

In a production environment, you'd typically set the **minimum number of nodes** value to 0 so that compute is only started when it is needed.

**43. How do you reduce the amount of time you spend waiting for the compute to start?**

To reduce the amount of time you spend waiting for it you've initialized it with two permanently running nodes.

**44. In the Machine Learning Studio, where do you register the data to train the model?**

Assets > Datasets

**45. How many ways you can import data for creating datasets?**

From Local files  
From datastore  
From web files  
From open datasets

**46. You have created a dataset and you want to see the Sample of the data. Where do you see in the Machine Learning Studio?**

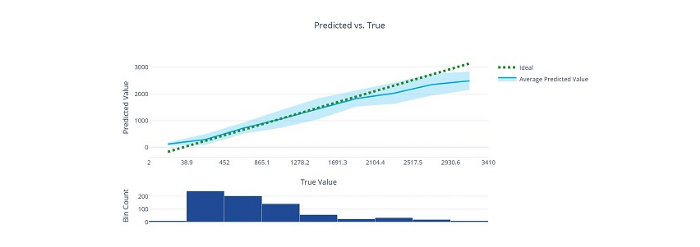
After the dataset has been created, open it and view the **Explore** page to see a sample of the data.

**47. Where do you run experiments in ML Studio?**

**Author > Automated ML**

pageCreate a new Automated ML run  
select dataset  
Configure run  
Task type and settings

**49. Which kind of model to produce the Predicted vs True chart?**



Regression

**49. An automobile dealership wants to use historic car sales data to train a machine learning model. The model should predict the price of a pre-owned car based on characteristics like its age, engine size, and mileage. What kind of machine learning model does the dealership need to create?**

Regression

**50. A bank wants to use historic loan repayment records to categorize loan applications as low-risk or high-risk based on characteristics like the loan amount, the income of the borrower, and the loan period. What kind of machine learning model does the bank need to create?**

Classification

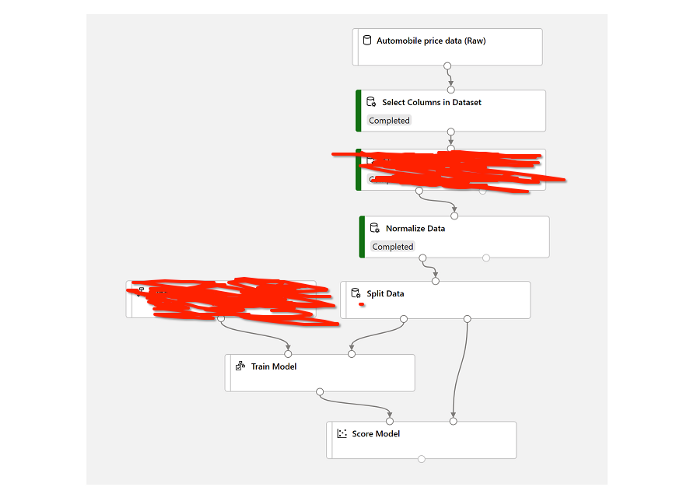
**50. Which of the following types of machine learning is an example of unsupervised machine learning?**

Clustering

**51. You are creating a model with the Azure Machine Learning designer. As a first step, you import the raw data. What are the next steps you need to do to prepare the data for the modeling?**



**52. You have created model with Azure Machine Learning designer using linear regression. What are the missing steps in the below diagram?**



Clean Missing Data  
Linear Regression

**53. What is Mean Absolute Error (MAE)?**

The average difference between predicted values and true values. This value is based on the same units as the label, in this case dollars. The lower this value is, the better the model is predicting.

**54. What is Root Mean Squared Error (RMSE)?**

The mean difference between predicted and true values is squared, and then the square root is calculated. The result is a metric based on the same unit as the label (dollars). When compared to the MAE (above), a larger difference indicates greater variance in the individual errors (for example, with some errors being very small, while others are large). If the MAE and RMSE are approximately the same, then all individual errors are of a similar magnitude.

**55. What is Relative Squared Error (RSE)?**

A relative metric between 0 and 1 based on the square of the differences between predicted and true values. The closer to 0 this metric is, the better the model is performing. Because this metric is relative, it can be used to compare models where the labels are in different units.

**56. What is Relative Absolute Error (RAE)?**

A relative metric between 0 and 1 based on the absolute differences between predicted and true values. The closer to 0 this metric is, the better the model is performing. Like RSE, this metric can be used to compare models where the labels are in different units.

**57. What is the Coefficient of Determination (R2)?**

This metric is more commonly referred to as R-Squared, and summarizes how much of the variance between predicted and true values is explained by the model. The closer to 1 this value is, the better the model is performing.

**58. You plan to use the Azure Machine Learning designer to create and publish a regression model. Which edition should you choose when creating an Azure Machine Learning workspace?**

Enterprise

**59. You are creating a training pipeline for a regression model, using a dataset that has multiple numeric columns in which the values are on different scales. You want to transform the numeric columns so that the values are all on a similar scale based relative to the minimum and maximum values in each column. Which module should you add to the pipeline?**

Normalize Data

**60. You use the Azure Machine Learning designer to create a training pipeline and an inference pipeline for a regression model. Now you plan to deploy the inference pipeline as a real-time service. What kind of compute target should you create to host the service?**

Inference Cluster

**61. \_\_\_\_\_\_\_ is a form of machine learning that is used to predict which category, or class, an item belongs to.**

Classification

**62. A health clinic might use the characteristics of a patient (such as age, weight, blood pressure, and so on) to predict whether the patient is at risk of diabetes. In this case, the characteristics of the patient are the features, and the label is a classification of either 0 or 1, representing non-diabetic or diabetic. What kind of model is this?**

Classification

**63. You are using the Azure Machine Learning designer to create a training pipeline for a binary classification model. You have added a dataset containing features and labels, a Two-Class Decision Forest module, and a Train Model module. You plan to use Score Model and Evaluate Model modules to test the trained model with a subset of the dataset that was not used for training. Which additional kind of module should you add?**

Split Data

**64. You use an Azure Machine Learning designer pipeline to train and test a binary classification model. You review the model’s performance metrics in an Evaluate Model module and note that it has an AUC score of 0.3. What can you conclude about the model?**

The model performs worse than random guessing.

**65. You use the Azure Machine Learning designer to create a training pipeline for a classification model. What must you do before deploying the model as a service?**

Create an inference pipeline from the training pipeline

**66. What is the Accuracy metric in the classification model?**

The ratio of correct predictions (true positives + true negatives) to the total number of predictions. In other words, what proportion of diabetes predictions did the model get right?

**67. What is called the F1 score metric in the classification model?**

An overall metric that essentially combines precision and recall.

**68. \_\_\_\_\_\_\_ is a form of machine learning that is used to group similar items into clusters based on their features?**

Clustering

**69. To train a clustering model, you need to apply a clustering algorithm to the data, using only the features that you have selected for clustering. You’ll train the model with a subset of the data, and use the rest to test the trained model. This is the complete pipeline for clustering what are the missing modules in the following pipeline?**



Normalize Data  
K-Means Clustering

**70. You are using an Azure Machine Learning designer pipeline to train and test a K-Means clustering model. You want your model to assign items to one of three clusters. Which configuration property of the K-Means Clustering module should you set to accomplish this?**

Set Number of Centroids to 3

**71. You use the Azure Machine Learning designer to create a training pipeline for a clustering model. Now you want to use the model in an inference pipeline. Which module should you use to infer cluster predictions from the model?**

Assign Data to Clusters

# Describe features of computer vision workloads on Azure (15–20%)

Practice questions based on these concepts

* Identify common types of computer vision solution
* Identify Azure tools and services for computer vision tasks

**72. What can Computer Vision cognitive service do?**

Interpret an image and suggest an appropriate caption.

Suggest relevant tags that could be used to index an image. Categorize an image.

Identify objects in an image.

Detect faces and people in an image.

Recognize celebrities and landmarks in an image.

Read text in an image.

**73. When using Computer Vision, what is the difference between Computer Vision and Cognitive Services?**

**Computer Vision**: A specific resource for the Computer Vision service. Use this resource type if you don’t intend to use any other cognitive services, or if you want to track utilization and costs for your Computer Vision resource separately.

**Cognitive Services**: A general cognitive services resource that includes Computer Vision along with many other cognitive services; such as Text Analytics, Translator Text, and others. Use this resource type if you plan to use multiple cognitive services and want to simplify administration and development.

**74. If the client wants to use the Computer Vision services what do they need?**

A **key** that is used to authenticate client applications.

An **endpoint** that provides the HTTP address at which your resource can be accessed.

**75. Can Computer Vision describe the images?**

Yes

**76. Computer Vision detects the objects in the image. Is this true?**

True

The object detection capability is similar to tagging, in that the service can identify common objects; but rather than tagging, or providing tags for the recognized objects only, this service can also return what is known as bounding box coordinates.

**77. Computer Vision detects the brands in the image. Is this true?**

True

This feature provides the ability to identify commercial brands. The service has an existing database of thousands of globally recognized logos from commercial brands of products.

**78. With Computer Vision you can categorize the people in the image. Is this true?**

True

**79. When categorizing an image, the Computer Vision service supports two specialized domain models. What are these?**

**Celebrities** — The service includes a model that has been trained to identify thousands of well-known celebrities from the worlds of sports, entertainment, and business.

**Landmarks** — The service can identify famous landmarks, such as the Taj Mahal and the Statue of Liberty.

**80. The Computer Vision service can use \_\_\_\_\_\_\_\_ capabilities to detect printed and handwritten text in images.**

optical character recognition (OCR)

**81. If you want to detect images that contain adult content or depict violent, gory scenes. Can Computer Vision service help in this scenario?**

Yes

Moderate content - detecting images that contain adult content or depict violent, gory scenes.

**82. You want to use the Computer Vision service to analyze images. You also want to use the Text Analytics service to analyze text. You want developers to require only one key and endpoint to access all of your services. What kind of resource should you create in your Azure subscription?**

Cognitive Services

**83. You want to use the Computer Vision service to identify the location of individual items in an image. Which of the following features should you retrieve?**

Objects

**84. You want to use the Computer Vision service to analyze images of locations and identify well-known buildings? What should you do?**

Retrieve the categories for the image, specifying the landmarks domain

**85. \_\_\_\_\_\_\_ is a machine learning technique in which the object being classified is an image, such as a photograph.**

Image classification

**86. What are the uses of Image classification?**

**Product identification** — performing visual searches for specific products in online searches or even, in-store using a mobile device.

**Disaster investigation** — evaluating key infrastructure for major disaster preparation efforts. For example, aerial surveillance images may show bridges and classify them as such. Anything classified as a bridge could then be marked for emergency preparation and investigation.

**Medical diagnosis** — evaluating images from X-ray or MRI devices could quickly classify specific issues found as cancerous tumors, or many other medical conditions related to medical imaging diagnosis.

**87. What are the resources available for Custom Vision in Azure?**

**Custom Vision**: A dedicated resource for the custom vision service, which can be either a training or a prediction resource.

**Cognitive Services**: A general cognitive services resource that includes Custom Vision along with many other cognitive services. You can use this type of resource for training, prediction, or both.

**88. The model training process is an iterative process in which the Custom Vision service repeatedly trains the model using some of the data, but holds some back to evaluate the model. What are the evaluation metrics?**

**Precision**: What percentage of the class predictions made by the model were correct? For example, if the model predicted that 10 images are oranges, of which eight were actually oranges, then the precision is 0.8 (80%).

**Recall**: What percentage of class predictions did the model correctly identify? For example, if there are 10 images of apples, and the model found 7 of them, then the recall is 0.7 (70%).

**Average Precision (AP)**: An overall metric that takes into account both precision and recall).

**89. Once you publish the model to your prediction resource. To use your model, what information that client application developers need?**

**Project ID**: The unique ID of the Custom Vision project you created to train the model.

**Model name**: The name you assigned to the model during publishing.

**Prediction endpoint**: The HTTP address of the endpoints for the prediction resource to which you published the model (**not** the training resource).

**Prediction key**: The authentication key for the prediction resource to which you published the model (**not** the training resource).

**90. You plan to use the Custom Vision service to train an image classification model. You want to create a resource that can only be used for model training, and not for prediction. Which kind of resource should you create in your Azure subscription?**

Custom Vision

**91. You train an image classification model that achieves less than satisfactory evaluation metrics. How might you improve it?**

Add more images to the training set.

**92. You have published an image classification model. What information must you provide to developers who want to use it?**

The project ID, the model name, and the key and endpoint for the prediction resource

**93. \_\_\_\_\_\_\_ is a form of machine learning-based computer vision in which a model is trained to recognize individual types of object in an image, and to identify their location in the image.**

Object detection

**94. What information object detection model returns?**

The class of each object identified in the image.

The probability score of the object classification (which you can interpret as the confidence of the predicted class being correct)

The coordinates of a bounding box for each object.

**95. What is the difference between Object detection and Image classification?**

Image classification is a machine learning based form of computer vision in which a model is trained to categorize images based on the primary subject matter they contain.

Object detection goes further than this to classify individual objects within the image, and to return the coordinates of a bounding box that indicates the object's location.

**96. What are the uses of object detection?**

Evaluating the safety of a building by looking for fire extinguishers or other emergency equipment.

Creating software for self-driving cars or vehicles with lane assist capabilities.

Medical imaging such as an MRI or x-rays that can detect known objects for medical diagnosis.

**97. What are the key considerations when tagging training images for object detection are ensuring that you have sufficient images of the objects?**

Preferably from multiple angles;   
Making sure that the bounding boxes are defined tightly around each object.

**98. Which of the following results does an object detection model typically return for an image?**

A class label, probability, and bounding box for each object in the image

**99. You plan to use a set of images to train an object detection model, and then publish the model as a predictive service. You want to use a single Azure resource with the same key and endpoint for training and prediction. What kind of Azure resource should you create?**

Cognitive Services

**100. \_\_\_\_\_\_\_\_\_ is an area of artificial intelligence (AI) in which we use algorithms to locate and analyze human faces in images or video content.**

Face detection and analysis

**101. The facial landmarks can be used as features with which to train a machine learning model from which you can infer information about a person, such as their perceived age or perceived emotional state. Is this true?**

True

**102. What are the uses of face detection and analysis?**

**Security** — facial recognition can be used in building security applications, and increasingly it is used in smartphone operating systems for unlocking devices.

**Social media** — facial recognition can be used to automatically tag known friends in photographs.

Intelligent monitoring — for example, an automobile might include a system that monitors the driver’s face to determine if the driver is looking at the road, looking at a mobile device, or shows signs of tiredness.

Advertising — analysing faces in an image can help direct advertisements to an appropriate demographic audience.

**Missing persons** — using public cameras systems, facial recognition can be used to identify if a missing person is in the image frame.

**Identity validation** — useful at ports of entry kiosks where a person holds a special entry permit.

**103. What are the cognitive services that you can use to detect and analyse faces from Microsoft Azure?**

**Computer Vision**, which offers face detection and some basic face analysis, such as determining age.

**Video Indexer**, which you can use to detect and identify faces in a video.

**Face**, which offers pre-built algorithms that can detect, recognize, and analysed faces.

**104. What information client applications need to use face service?**

A **key** that is used to authenticate client applications.

An **endpoint** that provides the HTTP address at which your resource can be accessed.

**105. What are some of the tips that can help improve the accuracy of the detection in the images when using Face service?**

**Image format** — supported images are JPEG, PNG, GIF, and BMP

**File size** — 4 MB or smaller

**Face size range** — from 36 x 36 up to 4096 x 4096. Smaller or larger faces will not be detected

**There issues** — face detection can be impaired by extreme face angles, occlusion (objects blocking the face such as sunglasses or a hand). Best results are obtained when the faces are full-frontal or as near as possible to full-frontal

**106. You plan to use Face to detect human faces in an image. How does the service indicate the location of the faces it detects?**

A set of coordinates for each face, defining a rectangular bounding box around the face

**107. What is one aspect that may impair facial detection?**

Extreme angles

**108. You want to use Face to identify named individuals. What must you do?**

Use Face to create a group containing multiple images of each named individual, and train a model based on the group

**109. What are the uses of OCR?**

Note-

Taking digitizing forms, such as medical records or historical documents scanning printed or handwritten checks for bank deposits

**110. The basic foundation of processing printed text is \_\_\_\_\_\_\_?**

optical character recognition (OCR)

**111. \_\_\_\_\_\_\_\_\_ is an AI system not only reads the text characters but can use a semantic model to interpret the text is about.**

machine reading comprehension (MRC)

**112. What is OCR API?**

The OCR API is designed for quick extraction of small amounts of text in images. It operates synchronously to provide immediate results, and can recognize text in numerous languages.

**113. What is the information that OCR API returns?**

**Regions** in the image that contain text  
**Lines** of text in each region  
**Words** in each line of text

For each of these elements, the OCR API also returns bounding box coordinates that define a rectangle to indicate the location in the image where the region, line, or word appears.

**114. What is the Read API?**

The Read API uses the latest recognition models and is optimized for images that have a significant amount of text or has considerable visual noise.

**115. The Read API is a better option for scanned documents that have a lot of text. Is this true?**

True

**116. What is the information that Read API returns?**

**Pages** — One for each page of text, including information about the page size and orientation.  
**Lines** — The lines of text on a page.  
**Words** — The words in a line of text.

Each line and word includes bounding box coordinates indicating its position on the page.

**117. The OCR API works synchronously and the Read API works asynchronously. Is this correct?**

True

**118. Why the Read API works asynchronously?**

Because the Read API can work with larger documents

**119. You want to extract text from images and then use the Text Analytics service to analyze the text. You want developers to require only one key and endpoint to access all of your services. What kind of resource should you create in your Azure subscription?**

Cognitive Services

**120. You plan to use the Computer Vision service to read the text in a large PDF document. Which API should you use?**

The Read API

**121. The \_\_\_\_\_\_\_\_\_ in Azure provides intelligent form processing capabilities that you can use to automate the processing of data in documents such as forms, invoices, and receipts.**

Form Recognizer

**122. How many ways Form Recognizer supports automated document processing?**

2 ways

**A pre-built receipt model** that is provided out-of-the-box, and is trained to recognize and extract data from sales receipts.

**Custom models**, which enable you to extract what are known as key/value pairs and table data from forms. Custom models are trained using your own data, which helps to tailor this model to your specific forms. Starting with only five samples of your forms, you can train the custom model. After the first training exercise, you can evaluate the results and consider if you need to add more samples and re-train.

**123. Currently, the pre-built receipt model is designed to recognize common receipts, in English, that are common to the USA. Is this true?**

True

**124. What are the guidelines to get the best results when using a custom model?**

Images must be JPEG, PNG, BMP, PDF, or TIFF formats  
File size must be less than 20 MB  
Image size between 50 x 50 pixels and 10000 x 10000 pixel  
For PDF documents, no larger than 17 inches x 17 inches

**125. You plan to use the Form Recognizer pre-built receipt model. Which kind of Azure resource should you create?**

Form Recognizer

**126. You are using the Form Recognizer service to analyze receipts that you have scanned into JPG format images. What is the maximum file size of the JPG file you can submit to the pre-built receipt model?**

20 MB

# Describe features of Natural Language Processing (NLP) workloads on Azure (15–20%)

Practice questions based on these concepts

* Identify features of common NLP Workload Scenarios
* Identify Azure tools and services for NLP workloads

**127. What is Text Analytics?**

Text analytics is a process where an artificial intelligence (AI) algorithm, running on a computer, evaluates these same attributes in text, to determine specific insights.

**128. You need to use a service from Azure that determines the language of a document or text (for example, French or English). Which one should you use?**

**Text Analytics** cognitive service

**129. You need to use a service from Azure that performs sentiment analysis on text to determine a positive or negative sentiment. Which one should you use?**

**Text Analytics** cognitive service

**130. You need to use a service from Azure that extracts key phrases from the text that might indicate its main talking points. Which one should you use?**

**Text Analytics** cognitive service

**131. You need to use a service from Azure Identify and categorize entities in the text. Entities can be people, places, organizations, or even everyday items such as dates, times, quantities, and so on. Which one should you use?**

**Text Analytics** cognitive service

**132. You are planning to read text information only. Which resource you should provision?**

A **Text Analytics** resource - choose this resource type if you only plan to use the Text Analytics service, or if you want to manage access and billing for the resource separately from other services.

**133. You are planning to read text information and objects in the Image. Which resource you should provision?**

A **Cognitive Services** resource - choose this resource type if you plan to use the Text Analytics service in combination with other cognitive services, and you want to manage access and billing for these services together.

**134. The Text Analytics service has language detection capability and you can submit multiple documents at a time for analysis. Is this true?**

True

**135. You have submitted multiple documents to the Text Analytics service. What is the output for each document?**

\* The language name (for example "English")  
\* The ISO 6391 language code (for example, "en")  
\* A score indicating a level of confidence in the language detection.

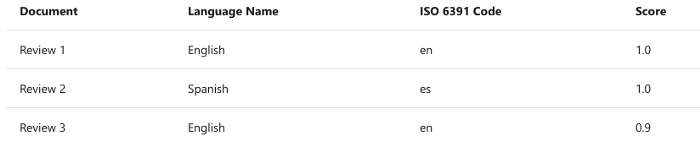
**136. Consider a scenario where you own and operate a restaurant where customers can complete surveys and provide feedback on the food, the service, staff, and so on. Suppose you have received the following reviews from customers:**

**Review 1**: “A fantastic place for lunch. The soup was delicious.”

**Review 2**: “Comida maravillosa y gran servicio.”

**Review 3**: “The croque monsieur avec frites was terrific. Bon appetit!”

You can use the Text Analytics service to detect the language for each of these reviews, and it might respond with the following results:



What does the information in the above table mean?

**Review 1**: It detected English with 1.0 confidence  
**Review 2**: It detected Spanish with 1.0 confidence  
**Review 3**: The language detection service will focus on the **predominant** language in the text. The service uses an algorithm to determine the predominant language, such as length of phrases or total amount of text for the language compared to other languages in the text.

The predominant language will be the value returned, along with the language code. The confidence score may be less than 1 as a result of the mixed language text.

**137. When the text in the document is ambiguous or mixed language content. What is the output of the Text Analytics service?**

An ambiguous content example would be a case where the document contains limited text, or only punctuation. For example, using the service to analyze the text ":-)", results in a value of **unknown** for the language name and the language identifier, and a score of **NaN** (which is used to indicate not a number).

**138. What does the confidence score of NaN Text Analytics service output mean?**

Ambiguous or mixed language content

**139. What is the Sentiment Analysis?**

The **Text Analytics service** can evaluate text and return sentiment scores and labels for each sentence. This capability is useful for detecting positive and negative sentiment in social media, customer reviews, discussion forums and more.

**140. What are the score ranges of Sentiment Analysis from the Text Analytics service?**

Using the pre-built machine learning classification model, the service evaluates the text and returns a sentiment score in the range of 0 to 1, with values closer to 1 being a positive sentiment. Scores that are close to the middle of the range (0.5) are considered neutral or indeterminate.

**141. What does the sentiment analysis score of 0.5 mean?**

Indeterminate sentiment

A score of 0.5 might indicate that the sentiment of the text is indeterminate, and could result from text that does not have sufficient context to discern a sentiment or insufficient phrasing. For example, a list of words in a sentence that has no structure, could result in an indeterminate score.

**142. You are using the Text Analytics service for sentiment analysis. You have used the wrong language code. For example, A language code (such as “en” for English, or “fr” for French) is used to inform the service which language the text is in. What score does the service return?**

The service will return a score of precisely 0.5.

**143. What is Keyphrase extraction?**

Key phrase extraction is the concept of evaluating the text of a document, or documents, and then identifying the main talking points of the document(s).

**144. You are running a restaurant and have collected thousands of reviews through a number of surveys. You don’t have time to go through each review but, you want to know the most talking points. What feature of Text Analytics would help here?**

**Key phrase extraction:** you can use the key phrases to identify important elements of the review.

**145. What is Entity Recognition?**

You can provide the Text Analytics service with unstructured text and it will return a list of entities in the text that it recognizes. The service can also provide links to more information about that entity on the web. An entity is essentially an item of a particular type or a category;

and in some cases, subtype

**146. You want to use the Text Analytics service to determine the key talking points in a text document. Which feature of the service should you use?**

Key phrase extraction

**147. You use the Text Analytics service to perform sentiment analysis on a document, and a score of 0.99 is returned. What does this score indicate about the document sentiment?**

The document is positive.

**148. When might you see NaN returned for a score in Language Detection?**

When the language is ambiguous

**149. What is Speech recognition?**

The ability to detect and interpret spoken input.

Speech recognition is concerned with taking the spoken word and converting it into data that can be processed - often by transcribing it into a text representation. The spoken words can be in the form of a recorded voice in an audio file, or live audio from a microphone.

**150. What is Speech synthesis?**

The ability to generate spoken output.

Speech synthesis is in many respects the reverse of speech recognition. It is concerned with vocalizing data, usually by converting text to speech

**151. What are the models you use to accomplish Speech recognition?**

**An acoustic model** that converts the audio signal into phonemes (representations of specific sounds).

**A language model** that maps phonemes to words, usually using a statistical algorithm that predicts the most probable sequence of words based on the phonemes.

**152. What are some of the use cases for speech recognition?**

\* Providing closed captions for recorded or live videos  
\* Creating a transcript of a phone call or meeting  
\* Automated note dictation  
\* Determining intended user input for further processing

**153. What are some of the use cases for speech synthesis?**

\* Generating spoken responses to user input.  
\* Creating voice menus for telephone systems.  
\* Reading email or text messages aloud in hands-free scenarios.  
\* Broadcasting announcements in public locations, such as railway stations or airports.

**154. What are the required elements for the speech synthesis?**

The text to be spoken.  
The voice to be used to vocalize the speech.

To synthesize speech, the system typically tokenizes the text to break it down into individual words, and assigns phonetic sounds to each word. It then breaks the phonetic transcription into prosodic units (such as phrases, clauses, or sentences) to create phonemes that will be converted to audio format. These phonemes are then synthesized as audio by applying a voice, which will determine parameters such as pitch and timbre; and generating an audio wave form that can be output to a speaker or written to a file.

**155. What are the services for speech recognition and speech synthesis from Azure?**

The **Speech-to-Text** API  
The **Text-to-Speech** API

**156. You want to use a service from Azure for just translating user spoken output to text. Which resource you should be provisioned in the Azure subscription?**

A **Speech** resource - choose this resource type if you only plan to use the Speech service, or if you want to manage access and billing for the resource separately from other services.

**157. You can use the speech-to-text API to perform real-time or batch transcription of audio into a text format. What does it mean?**

Real-time speech-to-text allows you to transcribe text in audio streams. You can use real-time transcription for presentations, demos, or any other scenario where a person is speaking.

Not all speech-to-text scenarios are real time. You may have audio recordings stored on a file share, a remote server, or even on Azure storage. You can point to audio files with a shared access signature (SAS) URI and asynchronously receive transcription results.

**158. You have a person speaking right now and you want to transcribe that into written output. Which transcription should you use?**

Real-time transcription

**159. You have thousands of stored audio files and you want to transcribe that into written output. Which transcription should you use?**

Batch transcription

**160. Why Batch transcription is asynchronous?**

Batch transcription should be run in an asynchronous manner because the batch jobs are scheduled on a best-effort basis. Normally a job will start executing within minutes of the request but there is no estimate for when a job changes into the running state.

**161. You plan to build an application that uses the Speech service to transcribe audio recordings of phone calls into text and then submits the transcribed text to the Text Analytics service to extract key phrases. You want to manage access and billing for the application services in a single Azure resource. Which type of Azure resource should you create?**

Cognitive Services

**162. You want to use the Speech service to build an application that reads incoming email message subjects aloud. Which API should you use?**

Text-to-Speech

**163. What is Text Translation?**

Text translation can be used to translate documents from one language to another, translate email communications that come from foreign governments, and even provide the ability to translate web pages on the Internet. Many times you will see a Translate option for posts on social media sites, or the Bing search engine can offer to translate entire web pages that are turned in search results.

**164. What is Speech Translation?**

Speech translation is used to translate between spoken languages, sometimes directly (speech-to-speech translation) and sometimes by translating to an intermediary text format (speech-to-text translation).

**165. What is the service from Microsoft Azure for Text Translation?**

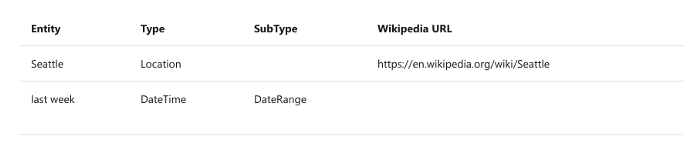
The **Translator Text** service, which supports text-to-text translation.

**166. What is the service from Microsoft Azure for Speech Translation?**

The **Speech** service, which enables speech-to-text and speech-to-speech translation.

**167. What is the output if you use the Text Analytics service to detect entities in the following restaurant review extract:**

**“I ate at the restaurant in Seattle last week.”**



**168. What are the services you should provision in your Azure subscription if you want to manage access and billing for each service individually?**

There are dedicated **Translator Text** and **Speech** resource types

**169. The Text Translator service supports text-to-text translation of more than 60 languages. is this correct?**

True

**170. Using the Text Translate service you can specify one from a language with multiple to languages, enabling you to simultaneously translate a source document to multiple languages. Is this true?**

True

**171. How do you handle brand names which are the same in all languages when using Text Translate service?**

**Selective translation**. You can tag content so that it isn't translated.

**172. When using the Text Translate you can control profanity translation by either marking the translated text as profane or by omitting it in the results. Is this correct?**

True

**Profanity filtering**. Without any configuration, the service will translate the input text, without filtering out profanity. Profanity levels are typically culture-specific but you can control profanity translation by either marking the translated text as profane or by omitting it in the results.

**173. \_\_\_\_\_\_\_\_ used to transcribe speech from an audio source to text format?**

**Speech-to-text**

**174. \_\_\_\_\_\_\_\_ used to generate spoken audio from a text source?**

**Text-to-speech**

**175. \_\_\_\_\_\_\_\_ used to translate speech in one language to text or speech in another?**

**Speech Translation**

**176. You are developing an application that must take English input from a microphone and generate a real-time text-based transcription in Hindi. Which service should you use?**

Speech

**177. You need to use the Translator Text service to translate email messages from Spanish into both English and French? What is the most efficient way to accomplish this goal?**

Make a single call to the service; specifying a "from" language of "es", a "to" language of "en", and another "to" language of "fr".

**178. On Microsoft Azure, language understanding is supported through the \_\_\_\_\_\_\_\_\_\_\_?**

**Language Understanding Intelligent Service**

**179. To work with Language Understanding, you need to take into account three core concepts. What are these concepts?**

utterances, entities, and intents.

**180. What are Utterances?**

An utterance is an example of something a user might say, and which your application must interpret. For example, when using a home automation system, a user might use the following utterances:"Switch the fan on.""Turn on the light."

**181. What are the Entities?**

An entity is an item to which an utterance refers. For example, **fan** and **light** in the following utterances:"Switch the **fan** on.""Turn on the **light**."

**182. What are Intents?**

An intent represents the purpose, or goal, expressed in a user's utterance. For example, for both of the previously considered utterances, the intent is to turn a device on; so in your Language Understanding application, you might define a **TurnOn** intent that is related to these utterances.

**183. What is None intent?**

In a Language Understanding application, the **None** intent is created but left empty on purpose. The None intent is a required intent and can't be deleted or renamed. Fill it with utterances that are outside of your domain.

**184. Creating a language understanding application with Language Understanding consists of two main tasks. What are these tasks?**

First you must define entities, intents, and utterances with which to train the language model - referred to as authoring the model.  
   
Then you must publish the model so that client applications can use it for intent and entity prediction based on user input.

**185. How many types of entities and what are those?**

There are four types of entities:

**Machine-Learned**: Entities that are learned by your model during training from context in the sample utterances you provide.

**List**: Entities that are defined as a hierarchy of lists and sublists. For example, a **device** list might include sublists for **light** and **fan**. For each list entry, you can specify synonyms, such as **lamp** for **light**.

**RegEx**: Entities that are defined as a regular expression that describes a pattern — for example, you might define a pattern like **[0–9]{3}-[0–9]{3}-[0–9]{4}** for telephone numbers of the form **555–123–4567**.**Pattern.any**: Entities that are used with patterns to define complex entities that may be hard to extract from sample utterances.

**186. You need to provision an Azure resource that will be used to author a new Language Understanding application. What kind of resource should you create?**

Language Understanding

**187. You are authoring a Language Understanding application to support an international clock. You want users to be able to ask for the current time in a specified city, for example, “What is the time in London?”. What should you do?**

Define a "city" entity and a "GetTime" intent with utterances that indicate the city intent.

**188. You have published your Language Understanding application. What information does a client application developer need to get predictions from it?**

The endpoint and key for the application's prediction resource

# Describe features of conversational AI workloads on Azure (15–20%)

Practice questions based on these concepts

* Identify common use cases for conversational AI
* Identify Azure services for conversational AI

**189. Name one example of Conversational AI?**

chat interface

**190. What do you need to implement a conversation AI-based chatbot?**

A **knowledge base** of question and answer pairs -- usually with some built-in natural language processing model to enable questions that can be phrased in multiple ways to be understood with the same semantic meaning.

**bot service** that provides an interface to the knowledge base through one or more channels.

**191. What is the Azure service to create and publish a knowledge base with built-in natural language processing capabilities?**

QnA Maker

**192. What is the Azure service that provides a framework for developing, publishing, and managing bots on Azure?**

Azure Bot Service.

**193. You can write code to create and manage knowledge bases using the QnA Maker REST API or SDK. Is this true?**

True in most scenarios it is easier to use the QnA Maker portal.

**194. To create a knowledge base, you must first provision a QnA Maker resource in your Azure subscription. Is this true?**

True

**195. After provisioning a QnA Maker resource, you can use the QnA Maker portal to create a knowledge base that consists of question-and-answer pairs. What are the ways to get this knowledge base?**

\* Generated from an existing FAQ document or web page.  
\* Imported from a pre-defined chit-chat data source.  
\* Entered and edited manually.

**196. Most of the time the knowledge base is created by FAQs. Is this true?**

False

A knowledge base is created using a combination of all of these techniques; starting with a base dataset of questions and answers from an existing FAQ document, adding common conversational exchanges from a chit-chat source, and extending the knowledge base with additional manual entries.

**197. There are so many alternatives to asking a question how do you solve this problem while creating a knowledge base?**

Questions in the knowledge base can be assigned alternative phrasing to help consolidate questions with the same meaning. For example, you might include a question like:

What is your head office location?

You can anticipate different ways this question could be asked by adding an alternative phrasing such as:

Where is your head office located?

**198. How to train the knowledge base?**

After creating a set of question-and-answer pairs, you must train your knowledge base. This process analyzes your literal questions and answers and applies a built-in natural language processing model to match appropriate answers to questions, even when they are not phrased exactly as specified in your question definitions.

**199. How to test the knowledge base?**

After training, you can use the **built-in test interface in the QnA Maker portal** to test your knowledge base by submitting questions and reviewing the answers that are returned.

**200. When to publish the knowledge base?**

When you're satisfied with your trained knowledge base, you can publish it so that client applications can use it over its REST interface.

**201. What does client applications need to access the published knowledge base?**

\* The knowledge base ID  
\* The knowledge base endpoint  
\* The knowledge base authorization key

**202. You have created and published a knowledge base. You want to deliver it to users through a custom bot. What should you do to accomplish this?**

You can create a custom bot by using the **Microsoft Bot Framework SDK** to write code that controls conversation flow and integrates with your QnA Maker knowledge base.

**203. How many ways you can create bots for your knowledge base?**

1. Custom bot by **Microsoft Bot Framework SDK**2. Automatic bot creation functionality of QnA Maker

**204. What is the Automatic bot creation functionality of QnA Maker?**

The automatic bot creation functionality of QnA Maker enables you create a bot for your published knowledge base and publish it as an Azure Bot Service application with just a few clicks.

**205. Can you extend and configure the bot?**

Yes

After creating your bot, you can manage it in the Azure portal, where you can:\* Extend the bot's functionality by adding custom code.  
\* Test the bot in an interactive test interface.  
\* Configure logging, analytics, and integration with other services.

**206. When your bot is ready you can connect to only one channel at one time. Is this true?**

False

When your bot is ready to be delivered to users, you can connect it to multiple channels; making it possible for users to interact with it through web chat, email, Microsoft Teams, and other common communication media.

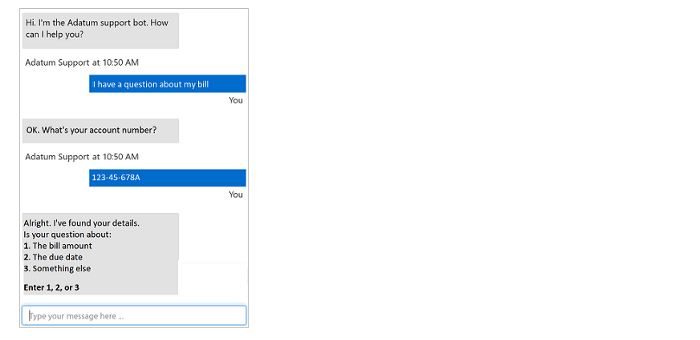
**207. Your organization has an existing frequently asked questions (FAQ) document. You need to create a QnA Maker knowledge base that includes the questions and answers from the FAQ with the least possible effort. What should you do?**

Import the existing FAQ document into a new knowledge base.

**208. You need to deliver a support bot for internal use in your organization. Some users want to be able to submit questions to the bot using Microsoft Teams, others want to use a web chat interface on an internal web site. What should you do?**

Create a knowledge base. Then create a bot for the knowledge base and connect the Web Chat and Microsoft Teams channels for your bot

**209. Bots are designed to interact with users in a conversational manner, as shown in this example of a chat interface. What kind of Azure resource should we use to accomplish this?**



Azure Bot Service.

# Conclusion

The AI fundamentals exam is multiple-choice, multiple answers, text-based, drag-and-drop, fill in the blanks exam. These sample questions definitely help you prepare for the certification. I would recommend you go through the documentation first and then refer to this afterward or right before the exam.

**Utterances**

An utterance is an example of something a user might say, and which your application must interpret. For example, when using a home automation system, a user might use the following utterances:

"*Switch the fan on.*"

"*Turn on the light.*"

**Entities**

An entity is an item to which an utterance refers. For example, **fan** and **light** in the following utterances:

"*Switch the****fan****on.*"

"*Turn on the****light****.*"

You can think of the **fan** and **light** entities as being specific instances of a general **device** entity.

**Intents**

An intent represents the purpose, or goal, expressed in a user's utterance. For example, for both of the previously considered utterances, the intent is to turn a device on; so in your Conversational Language Understanding application, you might define a **TurnOn** intent that is related to these utterances.

A Language Understanding application defines a model consisting of intents and entities. Utterances are used to train the model to identify the most likely intent and the entities to which it should be applied based on a given input. The home assistant application we've been considering might include multiple intents, like the following examples:

| **Intent** | **Related Utterances** | **Entities** |
| --- | --- | --- |
| Greeting | "Hello" |  |
|  | "Hi" |  |
|  | "Hey" |  |
|  | "Good morning" |  |
| TurnOn | "Switch the fan on" | fan (device) |
|  | "Turn the light on" | light (device) |
|  | "Turn on the light" | light (device) |
| TurnOff | "Switch the fan off" | fan (device) |
|  | "Turn the light off" | light (device) |
|  | "Turn off the light" | light (device) |
| CheckWeather | "What is the weather for today?" | today (datetime) |
|  | "Give me the weather forecast" |  |
|  | "What is the forecast for Paris?" | Paris (location) |
|  | "What will the weather be like in Seattle tomorrow?" | Seattle (location), tomorrow (datetime) |
| None | "What is the meaning of life?" |  |
|  | "Is this thing on?" |  |

In this table there are numerous utterances used for each of the intents. The intent should be a concise way of grouping the utterance tasks. Of special interest is the ***None*** intent. You should consider always using the None intent to help handle utterances that do not map any of the utterances you have entered. The None intent is considered a fallback, and is typically used to provide a generic response to users when their requests don't match any other intent.

# Getting started with Conversational Language Understanding

Creating an application with Conversational Language Understanding consists of two main tasks. First you must define entities, intents, and utterances with which to train the language model - referred to as authoring the model. Then you must publish the model so that client applications can use it for intent and entity prediction based on user input.

**Best practice is to use the Language portal for authoring and to use the SDK for runtime predictions.**

The endpoint and key for the application's prediction resource

Correct. Client applications must connect to the endpoint of the prediction resource, specifying an associated authentication key.

Conversations typically take the form of messages exchanged in turns; and one of the most common kinds of conversational exchange is a question followed by an answer. This pattern forms the basis for many user support bots, and can often be based on existing FAQ documentation. To implement this kind of solution, you need:

* A *knowledge base* of question and answer pairs - usually with some built-in natural language processing model to enable questions that can be phrased in multiple ways to be understood with the same semantic meaning.
* A *bot service* that provides an interface to the knowledge base through one or more channels.

### Define questions and answers

After provisioning a Language service resource, you can use the Language Studio's custom question answering feature to create a knowledge base that consists of question-and-answer pairs. These questions and answers can be:

* Generated from an existing FAQ document or web page.
* Entered and edited manually.

In many cases, a knowledge base is created using a combination of all of these techniques; starting with a base dataset of questions and answers from an existing FAQ document and extending the knowledge base with additional manual entries.

Questions in the knowledge base can be assigned alternative phrasing to help consolidate questions with the same meaning. For example, you might include a question like:

What is your head office location?

You can anticipate different ways this question could be asked by adding an alternative phrasing such as:

Where is your head office located?

When you're satisfied with your knowledge base, deploy it. Then you can use it over its REST interface. To access the knowledge base, client applications require:

* The knowledge base ID
* The knowledge base endpoint
* The knowledge base authorization key

While these techniques can be used to great effect, programming them can be complex. In Microsoft Azure, the **Language** cognitive service can help simplify application development by using pre-trained models that can:

* Determine the language of a document or text (for example, French or English).
* Perform sentiment analysis on text to determine a positive or negative sentiment.
* Extract key phrases from text that might indicate its main talking points.
* Identify and categorize entities in the text. Entities can be people, places, organizations, or even everyday items such as dates, times, quantities, and so on.

To use the Language service in an application, you must provision an appropriate resource in your Azure subscription. You can choose to provision either of the following types of resource:

* A **Language** resource - choose this resource type if you only plan to use natural language processing services, or if you want to manage access and billing for the resource separately from other services.
* A **Cognitive Services** resource - choose this resource type if you plan to use the Language service in combination with other cognitive services, and you want to manage access and billing for these services together.

## Language detection

Use the language detection capability of the Language service to identify the language in which text is written. You can submit multiple documents at a time for analysis. For each document submitted to it, the service will detect:

* The language name (for example "English").
* The ISO 6391 language code (for example, "en").
* A score indicating a level of confidence in the language detection.

For example, consider a scenario where you own and operate a restaurant where customers can complete surveys and provide feedback on the food, the service, staff, and so on. Suppose you have received the following reviews from customers:

**Review 1**: "A fantastic place for lunch. The soup was delicious."

**Review 2**: "Comida maravillosa y gran servicio."

**Review 3**: "The croque monsieur avec frites was terrific. Bon appetit!"

You can use the text analytics capabilities in the Language service to detect the language for each of these reviews; and it might respond with the following results:

| **Document** | **Language Name** | **ISO 6391 Code** | **Score** |
| --- | --- | --- | --- |
| Review 1 | English | en | 1.0 |
| Review 2 | Spanish | es | 1.0 |
| Review 3 | English | en | 0.9 |

Notice that the language detected for review 3 is English, despite the text containing a mix of English and French. The language detection service will focus on the ***predominant*** language in the text. The service uses an algorithm to determine the predominant language, such as length of phrases or total amount of text for the language compared to other languages in the text. The predominant language will be the value returned, along with the language code. The confidence score may be less than 1 as a result of the mixed language text.

### Ambiguous or mixed language content

There may be text that is ambiguous in nature, or that has mixed language content. These situations can present a challenge to the service. An ambiguous content example would be a case where the document contains limited text, or only punctuation. For example, using the service to analyze the text ":-)", results in a value of **unknown** for the language name and the language identifier, and a score of **NaN** (which is used to indicate not a number).

## Sentiment analysis

The text analytics capabilities in the Language service can evaluate text and return sentiment scores and labels for each sentence. This capability is useful for detecting positive and negative sentiment in social media, customer reviews, discussion forums and more.

Using the pre-built machine learning classification model, the service evaluates the text and returns a sentiment score in the range of 0 to 1, with values closer to 1 being a positive sentiment. Scores that are close to the middle of the range (0.5) are considered neutral or indeterminate.

For example, the following two restaurant reviews could be analyzed for sentiment:

"We had dinner at this restaurant last night and the first thing I noticed was how courteous the staff was. We were greeted in a friendly manner and taken to our table right away. The table was clean, the chairs were comfortable, and the food was amazing."

and

"Our dining experience at this restaurant was one of the worst I've ever had. The service was slow, and the food was awful. I'll never eat at this establishment again."

The sentiment score for the first review might be around 0.9, indicating a positive sentiment; while the score for the second review might be closer to 0.1, indicating a negative sentiment.

### Indeterminate sentiment

A score of 0.5 might indicate that the sentiment of the text is indeterminate, and could result from text that does not have sufficient context to discern a sentiment or insufficient phrasing. For example, a list of words in a sentence that has no structure, could result in an indeterminate score. Another example where a score may be 0.5 is in the case where the wrong language code was used. A language code (such as "en" for English, or "fr" for French) is used to inform the service which language the text is in. If you pass text in French but tell the service the language code is **en** for English, the service will return a score of precisely 0.5.

## Key phrase extraction

Key phrase extraction is the concept of evaluating the text of a document, or documents, and then identifying the main talking points of the document(s). Consider the restaurant scenario discussed previously. Depending on the volume of surveys that you have collected, it can take a long time to read through the reviews. Instead, you can use the key phrase extraction capabilities of the Language service to summarize the main points.

You might receive a review such as:

"We had dinner here for a birthday celebration and had a fantastic experience. We were greeted by a friendly hostess and taken to our table right away. The ambiance was relaxed, the food was amazing, and service was terrific. If you like great food and attentive service, you should try this place."

Key phrase extraction can provide some context to this review by extracting the following phrases:

* attentive service
* great food
* birthday celebration
* fantastic experience
* table
* friendly hostess
* dinner
* ambiance
* place

Not only can you use sentiment analysis to determine that this review is positive, you can use the key phrases to identify important elements of the review.

## Entity recognition

You can provide the Language service with unstructured text and it will return a list of entities in the text that it recognizes. The service can also provide links to more information about that entity on the web. An entity is essentially an item of a particular type or a category; and in some cases, subtype, such as those as shown in the following table.

To enable this kind of interaction, the AI system must support two capabilities:

* **Speech recognition** - the ability to detect and interpret spoken input.
* **Speech synthesis** - the ability to generate spoken output.

## Speech recognition

Speech recognition is concerned with taking the spoken word and converting it into data that can be processed - often by transcribing it into a text representation. The spoken words can be in the form of a recorded voice in an audio file, or live audio from a microphone. Speech patterns are analyzed in the audio to determine recognizable patterns that are mapped to words. To accomplish this feat, the software typically uses multiple types of models, including:

* An acoustic model that converts the audio signal into phonemes (representations of specific sounds).
* A language model that maps phonemes to words, usually using a statistical algorithm that predicts the most probable sequence of words based on the phonemes.

The recognized words are typically converted to text, which you can use for various purposes, such as.

* Providing closed captions for recorded or live videos
* Creating a transcript of a phone call or meeting
* Automated note dictation
* Determining intended user input for further processing

## Speech synthesis

Speech synthesis is in many respects the reverse of speech recognition. It is concerned with vocalizing data, usually by converting text to speech. A speech synthesis solution typically requires the following information:

* The text to be spoken.
* The voice to be used to vocalize the speech.

To synthesize speech, the system typically tokenizes the text to break it down into individual words, and assigns phonetic sounds to each word. It then breaks the phonetic transcription into prosodic units (such as phrases, clauses, or sentences) to create phonemes that will be converted to audio format. These phonemes are then synthesized as audio by applying a voice, which will determine parameters such as pitch and timbre; and generating an audio wave form that can be output to a speaker or written to a file.

You can use the output of speech synthesis for many purposes, including:

* Generating spoken responses to user input.
* Creating voice menus for telephone systems.
* Reading email or text messages aloud in hands-free scenarios.
* Broadcasting announcements in public locations, such as railway stations or airports.

## Azure resources for the Speech service

To use the Speech service in an application, you must create an appropriate resource in your Azure subscription. You can choose to create either of the following types of resource:

* A **Speech** resource - choose this resource type if you only plan to use the Speech service, or if you want to manage access and billing for the resource separately from other services.
* A **Cognitive Services** resource - choose this resource type if you plan to use the Speech service in combination with other cognitive services, and you want to manage access and billing for these services together.

## The speech-to-text API

You can use the speech-to-text API to perform real-time or batch transcription of audio into a text format. The audio source for transcription can be a real-time audio stream from a microphone or an audio file.

The model that is used by the speech-to-text API, is based on the Universal Language Model that was trained by Microsoft. The data for the model is Microsoft-owned and deployed to Microsoft Azure. The model is optimized for two scenarios, conversational and dictation. You can also create and train your own custom models including acoustics, language, and pronunciation if the pre-built models from Microsoft do not provide what you need.

### Real-time transcription

Real-time speech-to-text allows you to transcribe text in audio streams. You can use real-time transcription for presentations, demos, or any other scenario where a person is speaking.

In order for real-time transcription to work, your application will need to be listening for incoming audio from a microphone, or other audio input source such as an audio file. Your application code streams the audio to the service, which returns the transcribed text.

### Batch transcription

Not all speech-to-text scenarios are real time. You may have audio recordings stored on a file share, a remote server, or even on Azure storage. You can point to audio files with a shared access signature (SAS) URI and asynchronously receive transcription results.

Batch transcription should be run in an asynchronous manner because the batch jobs are scheduled on a best-effort basis. Normally a job will start executing within minutes of the request but there is no estimate for when a job changes into the running state.

## The text-to-speech API

The text-to-speech API enables you to convert text input to audible speech, which can either be played directly through a computer speaker or written to an audio file.

### Speech synthesis voices

When you use the text-to-speech API, you can specify the voice to be used to vocalize the text. This capability offers you the flexibility to personalize your speech synthesis solution and give it a specific character.

The service includes multiple pre-defined voices with support for multiple languages and regional pronunciation, including standard voices as well as neural voices that leverage neural networks to overcome common limitations in speech synthesis with regard to intonation, resulting in a more natural sounding voice. You can also develop custom voices and use them with the text-to-speech API

## Supported Languages

Both the speech-to-text and text-to-speech APIs support a variety of languages. Use the links below to find details about the supported languages:

* [Speech-to-text languages](https://docs.microsoft.com/en-us/azure/cognitive-services/speech-service/language-support#speech-to-text).
* [Text-to-speech languages](https://docs.microsoft.com/en-us/azure/cognitive-services/speech-service/language-support#text-to-speech).

# Get started with translation in Azure

Completed100 XP

* 3 minutes

Microsoft Azure provides cognitive services that support translation. Specifically, you can use the following services:

* The **Translator** service, which supports text-to-text translation.
* The **Speech** service, which enables speech-to-text and speech-to-speech translation.

## Azure resources for Translator and Speech

Before you can use the Translator or Speech services, you must provision appropriate resources in your Azure subscription.

There are dedicated **Translator** and **Speech** resource types for these services, which you can use if you want to manage access and billing for each service individually.

Alternatively, you can create a **Cognitive Services** resource that provides access to both services through a single Azure resource, consolidating billing and enabling applications to access both services through a single endpoint and authentication key.

## Text translation with the Translator service

The Translator service is easy to integrate in your applications, websites, tools, and solutions. The service uses a Neural Machine Translation (NMT) model for translation, which analyzes the semantic context of the text and renders a more accurate and complete translation as a result.

### Translator service language support

The Translator service supports text-to-text translation between [more than 60 languages](https://docs.microsoft.com/en-us/azure/cognitive-services/translator/languages). When using the service, you must specify the language you are translating ***from*** and the language you are translating ***to*** using ISO 639-1 language codes, such as en for English, fr for French, and zh for Chinese. Alternatively, you can specify cultural variants of languages by extending the language code with the appropriate 3166-1 cultural code - for example, en-US for US English, en-GB for British English, or fr-CA for Canadian French.

When using the Translator service, you can specify one ***from*** language with multiple ***to*** languages, enabling you to simultaneously translate a source document into multiple languages.

### Optional Configurations

The Translator API offers some optional configuration to help you fine-tune the results that are returned, including:

* **Profanity filtering**. Without any configuration, the service will translate the input text, without filtering out profanity. Profanity levels are typically culture-specific but you can control profanity translation by either marking the translated text as profane or by omitting it in the results.
* **Selective translation**. You can tag content so that it isn't translated. For example, you may want to tag code, a brand name, or a word/phrase that doesn't make sense when localized.

## Speech translation with the Speech service

The Speech service includes the following application programming interfaces (APIs):

* **Speech-to-text** - used to transcribe speech from an audio source to text format.
* **Text-to-speech** - used to generate spoken audio from a text source.
* **Speech Translation** - used to translate speech in one language to text or speech in another.

You can use the **Speech Translation** API to translate spoken audio from a streaming source, such as a microphone or audio file, and return the translation as text or an audio stream. This enables scenarios such as real-time closed captioning for a speech or simultaneous two-way translation of a spoken conversation.

### Speech service language support

As with the Translator service, you can specify one source language and one or more target languages to which the source should be translated. You can translate speech into [over 60 languages](https://docs.microsoft.com/en-us/azure/cognitive-services/speech-service/language-support#speech-translation).

The source language must be specified using the extended language and culture code format, such as es-US for American Spanish. This requirement helps ensure that the source is understood properly, allowing for localized pronunciation and linguistic idioms.

The target languages must be specified using a two-character language code, such as en for English or de for German.