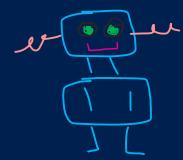


Developing and Deploying Intelligent Chat Bots: Cognitive Services



Main site: aka.ms/botedu
Chat room: aka.ms/botedu-discuss

Today's Agenda

- Cortana Intelligence Suite overview
- Cognitive Services overview
- Demos from Cognitive Services
- Intelligent Kiosk Demo
- Cognitive Services in Depth

Be ready for fun labs throughout ☺

Prerequisites for today

GitHub account
Microsoft account



Class site (for almost everything)

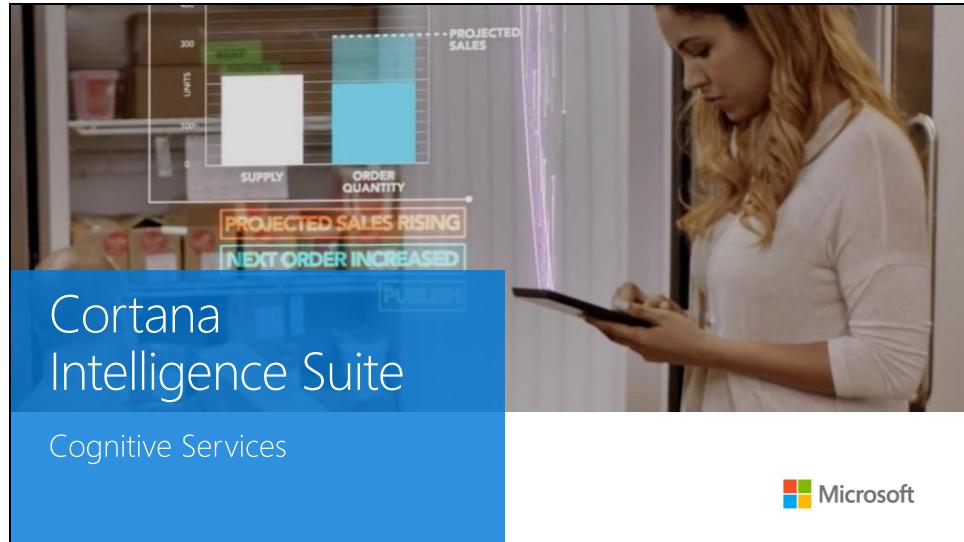
<https://aka.ms/botedu>

- This is a link to the wiki
- Click on "Code" for the rest

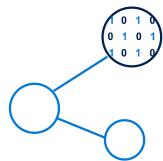
<https://aka.ms/botedu-discuss>

- This is our discussion forum – need GitHub id





Business is being transformed by three trends



Big Data



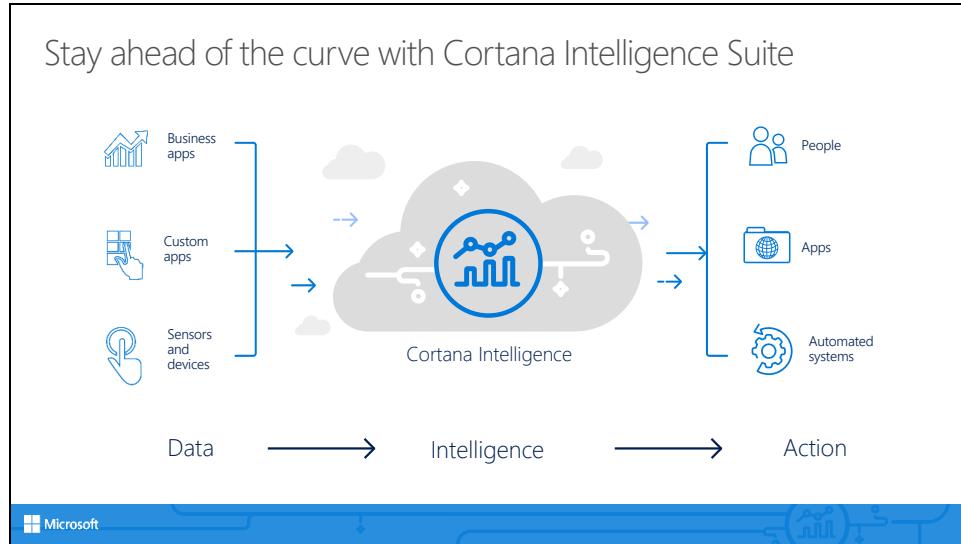
Cloud



Intelligence

Today, the world around us is rapidly changing, faster than at any other point in history. This is an indicator of the **fourth industrial revolution** that is emerging, largely driven by the rise of Big Data, the growth of the cloud, and a new era of intelligence capabilities.

Thanks to the exponential proliferation of small, inexpensive chips and processors, computers are as ubiquitous as the people who use them – from traditional computers, tablets and phones to sensors and wearables. Machines are everywhere, constantly creating, collecting and making sense of the data in our midst. With the magic of machine learning and the limitless computing power of the cloud, this data is giving rise to intelligence that is augmenting human capabilities in exciting new ways.

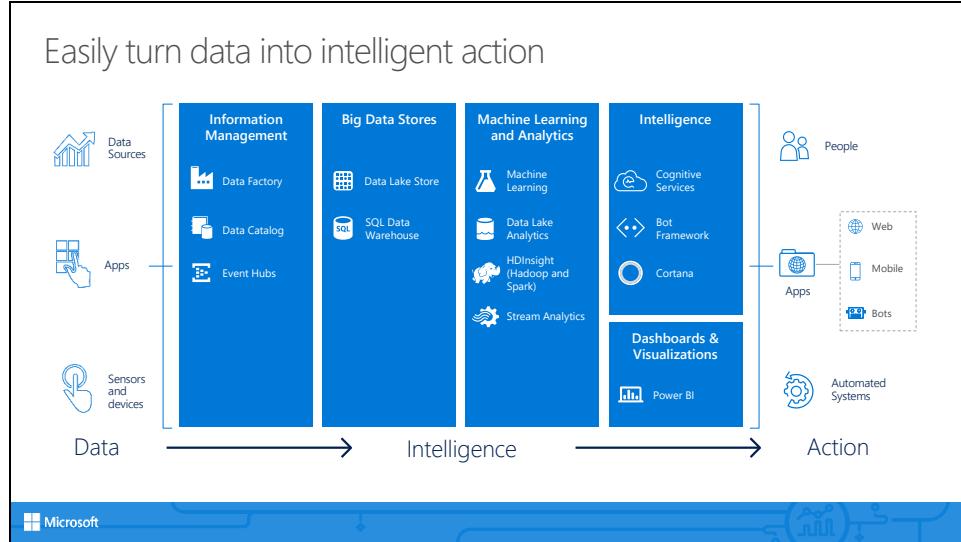


Cortana Intelligence is Microsoft's fully managed **intelligent, big data and advanced analytics** offering in the cloud, designed to help you **transform your data into intelligent action**.

It is a comprehensive suite that brings together technologies throughout Microsoft. It provides fast and flexible deployment, with a simple monthly subscription to reduce time and cost challenges.

Cortana Intelligence enables customers to benefit from Microsoft's investment in the intelligent cloud and advanced analytics, spanning our leading cloud platform with easy to use tools and services that integrate with existing infrastructure and enable enterprises to extend business solutions as their needs grow over time.

With Cortana Intelligence, we are taking years of research and innovation – spanning technology & infrastructure for advanced analytics, including capabilities such as machine learning, big data storage and processing in the cloud, intelligence capabilities like vision, face and speech recognition, and integration with Cortana, Microsoft's personal digital assistant, with the goal of helping enterprise customers make better, faster decisions to accelerate their speed of business.



Cortana Intelligence delivers an end-to-end platform with an integrated and comprehensive set of tools and services to help you build intelligent applications that let you easily take advantage of Advanced Analytics and intelligence capabilities.

First, Cortana Intelligence provides services to bring data in, so that you can analyze it. It provides information management capabilities like Azure Data Factory so that you can pull data from any source (relational DB like SQL or non-relational ones like your Hadoop cluster) in an automated and scheduled way, while performing the necessary data transforms (like setting certain data columns as dates vs. currency etc). Think ETL (Extract, Transform, Load) in the cloud. Event Hubs does the same for IoT type ingestion of data that streams in from lots of end points.

The data brought in then can be persisted in flexible big data storage services like Data Lake Store and Azure SQL Data Warehouse.

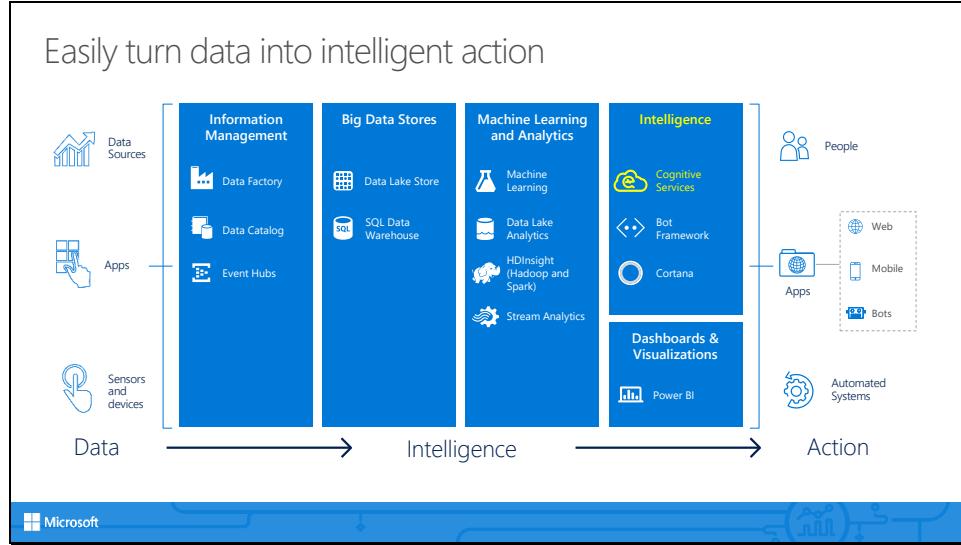
You can then use a wide range of analytics services from Machine Learning to Azure Data Lake Analytics to Azure HDInsight to Azure Stream Analytics to analyze the data stored in the big data storage. This means you can create analytics services and models specific to your business need (say real time demand forecasting).

The resultant analytics services and models created by taking these steps can then be surfaced as interactive dashboards and visualizations via Power BI.

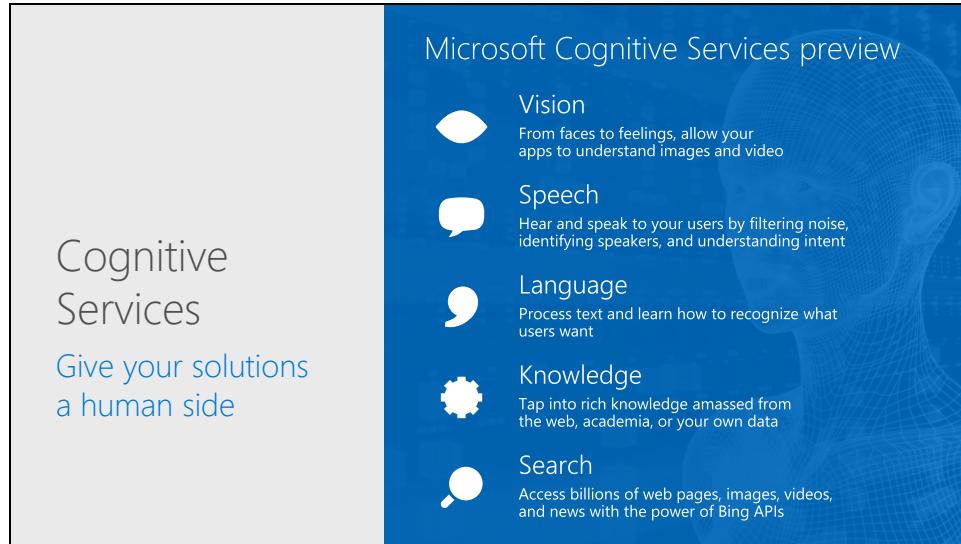
These same analytics services and models created can also be integrated into various different UI (web apps or mobile apps or rich client apps), or with Cortana, so end users can naturally interact with them via speech etc., and so that end users can get proactively be notified by Cortana if the analytics model finds a new anomaly (unusual growth in certain product purchases- in the case of real time demand forecasting example given above) or whatever deserves the attention of the business users. Similar integration can occur with Cognitive Services or Bot Framework based applications.

At a high level though, Cortana Intelligence capabilities are in three main areas: data, analytics and intelligence.

Slide 10



We're going to detail the Intelligence layer, more specifically the Microsoft Cognitive Services.



The slide features a white header section on the left and a blue header section on the right. The white section contains the text 'Cognitive Services' and 'Give your solutions a human side'. The blue section is titled 'Microsoft Cognitive Services preview' and lists five services: Vision, Speech, Language, Knowledge, and Search, each with a corresponding icon and a brief description.

Microsoft Cognitive Services preview	
	Vision From faces to feelings, allow your apps to understand images and video
	Speech Hear and speak to your users by filtering noise, identifying speakers, and understanding intent
	Language Process text and learn how to recognize what users want
	Knowledge Tap into rich knowledge amassed from the web, academia, or your own data
	Search Access billions of web pages, images, videos, and news with the power of Bing APIs

What are Cognitive Services? Microsoft Cognitive Services are a new collection of intelligence and knowledge APIs that enable developers to ultimately build smarter apps.

So, what are Cognitive Services? Cognitive Services are a collection of artificial intelligence APIs, and we believe in *democratizing* artificial intelligence. So what that means is, regardless of your skill level -- whether you're a high school student running your first program or working in industry or in a giant enterprise -- that you should be able to use our APIs incredibly quickly in a matter of minutes.

And regardless of your platform -- whether you're on Android or iOS or Windows, or making a website -- all of our APIs are rest APIs, which means you can call them as long as you have an Internet connection. And so that's pretty huge because what we're doing is making it so that everyone can build these smarter, more context-aware applications.

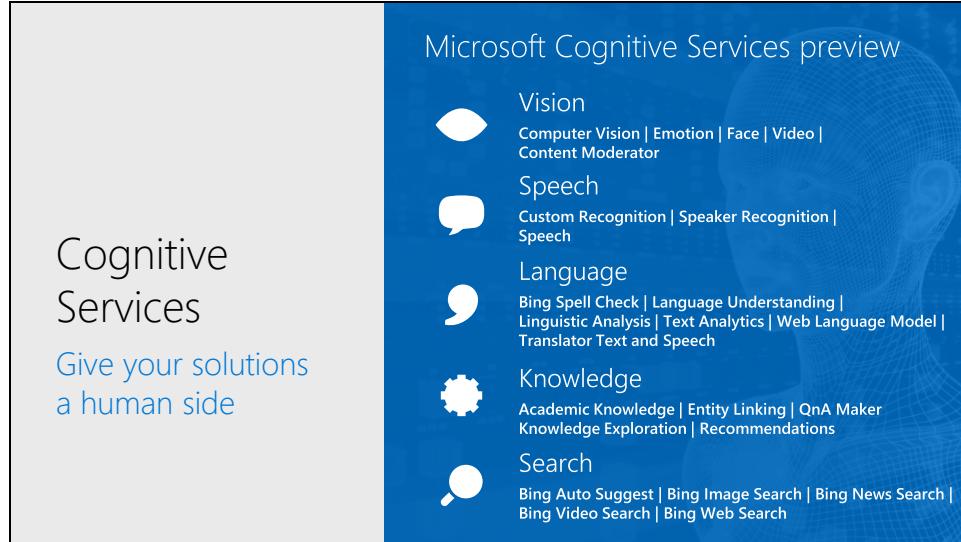
The technology used in our APIs is the same technology that powers our products today. And so, when you think of things like the Bing search APIs, it's the same technology from Bing.

Today I'm going to talk with you about the entire collection spanning vision, speech, language, knowledge, and search.

The other things that I want to point out is that you can get started for free with all of the APIs, but we do have pricing available for a number of them, which are in public preview on Azure.

The other piece is the developer resources. So, all of our documentation is on the website and actually in GitHub as well, so we do welcome community submissions. We have a set of SDKs that are also available on GitHub where we welcome pull requests and post everything on there. The SDKs vary from API to API, but they are all included in this one repository for people to see.

And then we have three different communities that we support. We have our MSDN forums, our Stack Overflow, and we have User Voice that we use for feedback requests.



The slide features a white header and a blue footer. The white header contains the text 'Cognitive Services' and 'Give your solutions a human side'. The blue footer is titled 'Microsoft Cognitive Services preview' and lists five service categories: Vision, Speech, Language, Knowledge, and Search, each with a corresponding icon and a list of sub-services.

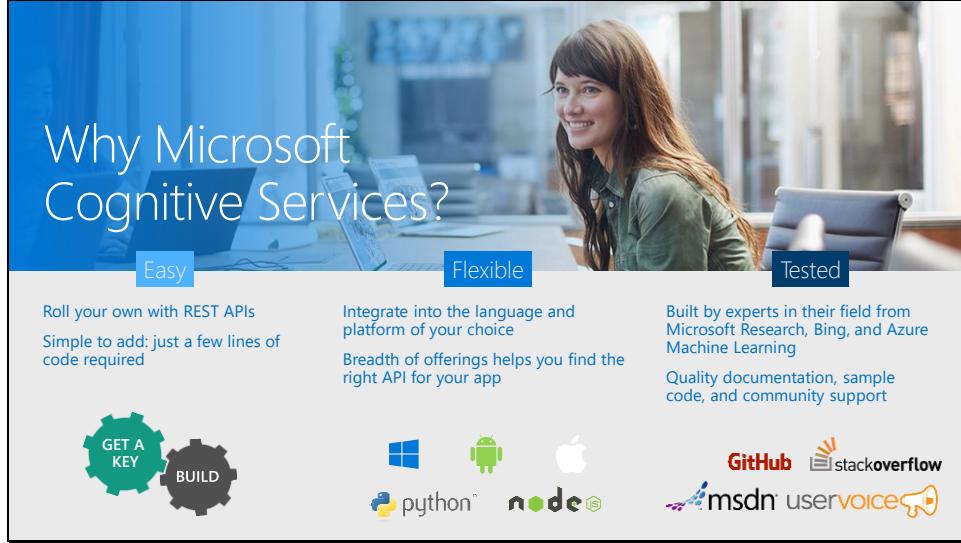
Microsoft Cognitive Services preview	
	Vision Computer Vision Emotion Face Video Content Moderator
	Speech Custom Recognition Speaker Recognition Speech
	Language Bing Spell Check Language Understanding Linguistic Analysis Text Analytics Web Language Model Translator Text and Speech
	Knowledge Academic Knowledge Entity Linking QnA Maker Knowledge Exploration Recommendations
	Search Bing Auto Suggest Bing Image Search Bing News Search Bing Video Search Bing Web Search

At Microsoft, we've been offering APIs for a very long time across the company. In delivering Microsoft Cognitive Services API, we started with 4 last year at /build (2015); added 7 more last December, and today we have 24 APIs in our collection.

Cognitive Services are available individually or as a part of the Cortana Intelligence Suite, formerly known as Cortana Analytics, which provides a comprehensive collection of services powered by cutting-edge research into machine learning, perception, analytics and social bots.

These APIs are powered by Microsoft Azure.

Developers and businesses can use this suite of services and tools to create apps that learn about our world and interact with people and customers in personalized, intelligent ways.



Why Microsoft Cognitive Services?

Easy Flexible Tested

Roll your own with REST APIs
Simple to add: just a few lines of code required

Integrate into the language and platform of your choice
Breadth of offerings helps you find the right API for your app

Built by experts in their field from Microsoft Research, Bing, and Azure Machine Learning
Quality documentation, sample code, and community support

GET A KEY BUILD

python node.js Windows Android Apple

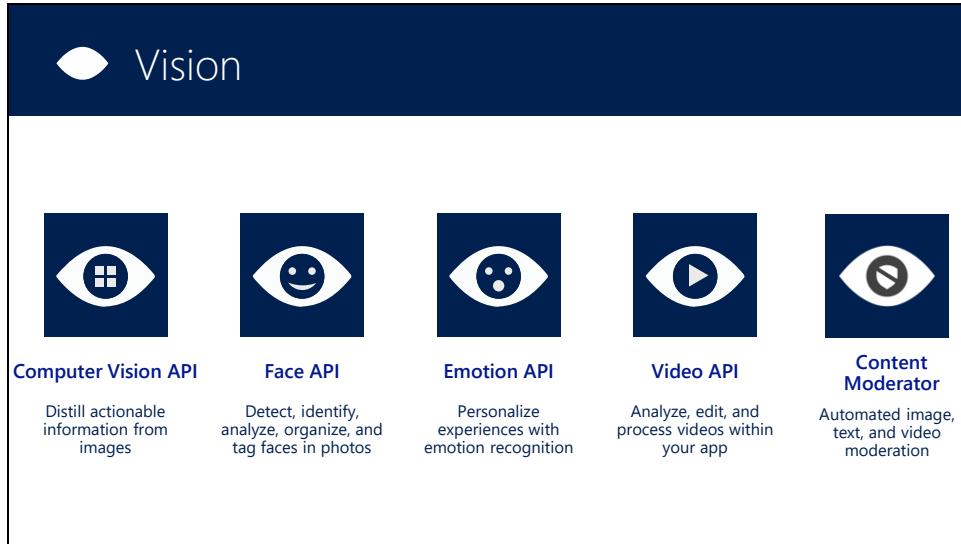
GitHub stackoverflow msdn uservoice

Why choose these APIs? They work, and it's easy.

Easy: The APIs are easy to implement because of the simple REST calls. Being REST APIs, there's a common way to implement and you can get started with all of them for free simply by going to one place, one website, www.microsoft.com/cognitive. (You don't have to hunt around to different places.)

Flexible: We've got a breadth of intelligence and knowledge APIs so developers will be able to find what intelligence feature they need; and importantly, they all work on whatever language, framework, or platform developers choose. So, devs can integrate into their apps—iOS, Android, Windows—using their own tools they know and love (such as python or node.js, etc.).

Tested: Tap into an ever-growing collection of powerful AI algorithms developed by experts. Developers can trust the quality and expertise built into each API by experts in their field from Microsoft's Research organization, Bing, and Azure machine learning and these capabilities are used across many Microsoft first party products such as Cortana, Bing and Skype.



Vision

Computer Vision API: as a free trial on the website microsoft.com/cognitive. There are also SDKs and Samples available on GitHub or through NuGet, Maven, and Cocoapods for select platforms to make development easier. It's important to note here that it's not client side running code, but light wrappers around the REST calls to make integration easy.

A photo app would use this as a way to tag user photos and make it easier for users to search through their collections. An assistive app would use this as a way to describe the surroundings to visually-impaired users. Works really well on both indoor or outdoor images; it can recognize common household objects, and it can describe outdoor scenes. However, we did not train on aerial images (say from drones), or on many close ups (so pictures where we zoomed in extremely on the subject won't do well). We also do really well recognizing celebrities (as long as most of the face is visible, and they were facing the camera).

Face API: Some potential uses for this technology include facial login, photo tagging, and home monitoring. Or attribute detection to know age, gender, facial hair, etc.

Emotion API: is available in the Azure marketplace, as a free trial on the website microsoft.com/cognitive. See Computer Vision description.

Build an app that responds to moods. Using facial expressions, this cloud-based API can detect happiness, neutrality, sadness, contempt, anger, disgust, fear, and surprise. The AI understands these emotions based on universal facial expressions, and it functions cross-culturally, so your app will work around the world. Some use cases would be an advertising company wants to test user response to an ad, a tv studio wants to track responses to a pilot.

Video API: as a free trial on the website microsoft.com/cognitive. See Computer Vision description.

It brings Microsoft state of the art video processing algorithms to developers. With Video API, developers can analyze and automatically edit videos, including stabilize videos, create motion thumbnails, track faces, and detect motion. Use cases: For Stabilization: If you have multiple action videos, you can use the stabilization algorithm to make them less shaky and easier to watch. You can also use the stabilization algorithm as a first step in performing other video APIs. For Face Tracking: You can track faces in a video to do A/B testing in a retail setting. You can combine Video API Face Tracking with capabilities in Face API to search through surveillance, crime, or media footage to look for certain person. For Motion Detection: Instead of having to watch long clips of surveillance footage, the API will let you know what time motion occurred and its duration. For Video Thumbnail: Take a long video, such as a keynote presentation, and automatically create a short preview clip of the talk. For Face Tracking: Works best for frontal faces. Currently cannot detect small faces, side or partial faces. For Motion Detection: Detects motion on a stationary background (e.g. fixed camera). Current limitations of the algorithms include night vision videos, semi-transparent objects, and small objects. For Video Thumbnail: Take a long video, such as a keynote presentation, and automatically create a short preview clip of the talk.

Content Moderator: With content moderator, perform automated image, text and video moderation.

Image Moderation

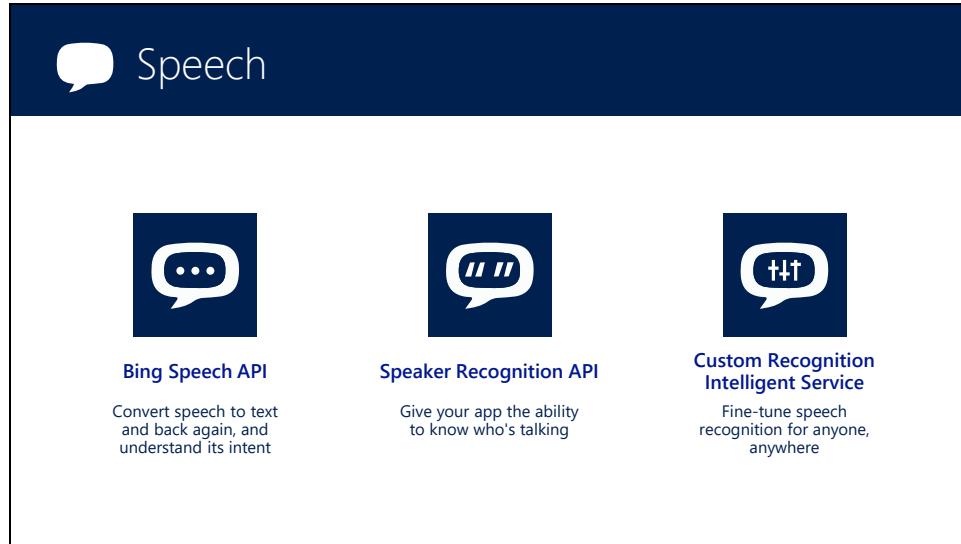
Automatically evaluate images for offensive and unwanted content across differing file types—including altered images. Content Moderator includes optical character recognition (OCR) and face detection to improve the chances of detecting unwanted or exploitative content and implements fuzzy matching against custom blacklists and whitelists.

Text Moderation

Detect profanity in more than 100 languages and match text against your custom lists automatically. Content Moderator also checks for malware and phishing URLs plus personally identifiable information (PII).

Video Moderation

Enable the proactive detection of adult content in videos with moderation handled in the cloud by Azure Media Services.



Speech APIs provide state-of-the-art algorithms to process spoken language powered by Bing. This also includes speech synthesis for a subset of languages supported by speech recognition. With these APIs developers can easily include the ability to add speech driven actions into their applications. In certain cases, the APIs allow for real-time interaction with the user as well. Additional capabilities include voice recognition and speaker identification as well as providing partial transcription, meaning that for supported languages the developer can get partial results before the user has finished speaking. The initial release supports 7 languages.

Bing Speech API: At the time of publication, this data was not available. Please email [Rebecca Duffy, redduffy@microsoft.com](mailto:redduffy@microsoft.com), if you would like more information.
We also have SDKs available for Speech.

CRIS: The Custom Recognition Intelligent Service (CRIS) enables you to create customized language models and acoustic models tailored to your application and your users. By uploading speech and/or text data to CRIS that reflects your application and your users, you can create custom models that can be used in conjunction with Microsoft's existing state-of-the-art speech models. To customize the acoustic model to a particular domain, a collection of speech data is required. This collection consists of a set of audio files of speech data, and a text file of transcriptions of each audio file. The audio data should be representative of the scenario in which you would like to use the recognizer. If you were building an app to search MSDN by voice, it's likely that terms like "object-oriented" or "namespace" or "dot net" will appear more frequently than in typical voice applications. Customizing the language model will enable the system to learn this. CRIS uses acoustic or language model adaptation to enable the speech recognizer to learn the characteristics of the customer's data while still getting the benefits of all the data and expertise that went into creating the base models that power the Speech API.

For acoustic model adaptation, the technology is described in this paper:
<http://research.microsoft.com/pubs/194346/0007893.pdf>

Works well when the data is uploaded to CRIS is truly representative of the user population and the expected usage of the application. For language model adaptation, it works best when the data uploaded reflects what

people would actually say. Uploading simply a list of new terms is better than nothing but will not be as effective. For acoustic model adaptation, if you want to adapt to elderly speech, you need to upload the speech from many different elderly users, not just one or two. If you upload just one person's voice, CRIS will learn to do a great job on that voice but will not necessarily learn to generalize to other elderly voices. Similarly, if you want to adapt to a new environment, like a factory, you should upload speech data from many speakers in the factory, not just one.

Speaker Recognition API: Microsoft's state-of-the-art cloud-based speaker recognition algorithms to recognize a human's voice in audio streams. It comprises two components: speaker verification and speaker identification. Speaker Verification can automatically verify and authenticate users from their voice or speech. It is tightly related to authentication scenarios and is often associated with a pass phrase. Hence, we opt for text-dependent approach, which means speakers need to choose a specific pass phrase to use during both enrollment and verification phases. Speaker Identification can automatically identify the person speaking in an audio file given a group of prospective speakers. The input audio is paired against the provided group of speakers, and in case there is a match found, the speaker's identity is returned. It is text-independent, which means that there are no restrictions on what the speaker says during the enrollment and recognition phases. A use case is biometric authentication using voice.

<https://blogs.technet.microsoft.com/machinelearning/2015/12/14/now-available-speaker-video-apis-from-microsoft-project-oxford/>

We also have SDKs available for Speaker Reco

Bing Spell Check API
Detect and correct spelling mistakes within your app

Web Language Model API
Leverage the power of language models trained on web-scale data

Linguistic Analysis API
Easily parse complex text with language analysis

Language Understanding Intelligent Service
Teach your apps to understand commands from your users

Text Analytics API
Detect sentiment, key phrases, topics, and language from your text

Translator
Easily perform speech and text translation

Bing Spell Check API: Microsoft's state-of-the-art cloud-based spelling algorithms to detect a wide variety of spelling errors and provide accurate suggestions. Using Bing Spell Check, your mobile and PC apps will be powered with state-of-the-art spelling capabilities. Our service is trained on a massive corpus of data gleaned from billions of web pages. There is no need to train your own models. The speller model is updated regularly and incorporates new terms and brands almost as quickly as they appear on the web. This API is available through Microsoft Cognitive Services for customers with low-volume and high-latency jobs. For high-volume and low-latency we have an internal API which may be more suitable.

Use cases: 1) Improve the quality of a website's product search 2) provide spell correction for a keyboard app 3) provide spell correction for text fields in an app or web page 4) detect errors in UI text and user data. See https://blogs.msdn.microsoft.com/msr_er/2010/12/15/building-a-better-speller-bing-and-microsoft-research-offer-prizes-for-best-search-engine-spelling-alteration-services/ The speller is exceptional at common spelling errors with low edit-distance (such as febuary->February) but a lot of other spellers are good at that as well. We Do a very good job with word breaking, proper names in context (try "director stephen spielberg") and fictional character names, just a few examples. Areas that are a challenge are capitalization (we don't know what to do with "March" for example, even with context) and consistency (there are times when we will flag a word only intermittently based on the context).

Web Language Model API: Web Language API indexes the web and Bing queries to allow users to calculate the probabilities of natural language expressions and estimate a list of most likely words to follow an existing sequence of words. Use this API to insert spaces into a string of words without spaces, like a hashtag or URL. Use this API to rerank machine translation/speech recognition candidate sentences based on probability of the sentence.

Use this API for academic research. <http://research.microsoft.com/apps/pubs/default.aspx?id=130762>

We also have SDKs available for WebLM

Linguistic Analysis API: The Linguistic Analysis API helps you gain insights from text. Given a natural language parse, it's easy to identify the concepts and entities (noun phrases), actions (verbs and verb phrases), descriptive words, and more. The processed text can provide useful features for classification tasks such as sentiment analysis.

We also have SDKs available for Linguistic Analysis.

LUIS: Language Understanding Intelligent Service (LUIS) allows developers to build a model that understands natural language and commands tailored to their application. Example: You can say "turn down the thermostat in the living room," send it to a LUIS model, and instead of just returning the text that represents what was said, LUIS will return: the action is "turn down," the location is "living room," and the target is "thermostat." LUIS allows developers to iteratively build on these models and take speech or text input and return a structured representation of what the person said. Not only that but by build LUIS will help developers create and train smart conversational bot (Intercom or Slack) with a single button. LUIS will also offer action fulfillment capabilities by simple integration with webhooks. LUIS works pretty well it comes to intents. For the entities, the learning curve is slower especially when the number of entities increases. LUIS only supports 20 intents & 10 entities yet by build each entities can have up to 10 children.

Text Analytics API: Detect sentiment, key phrases, topics, and language from your text.

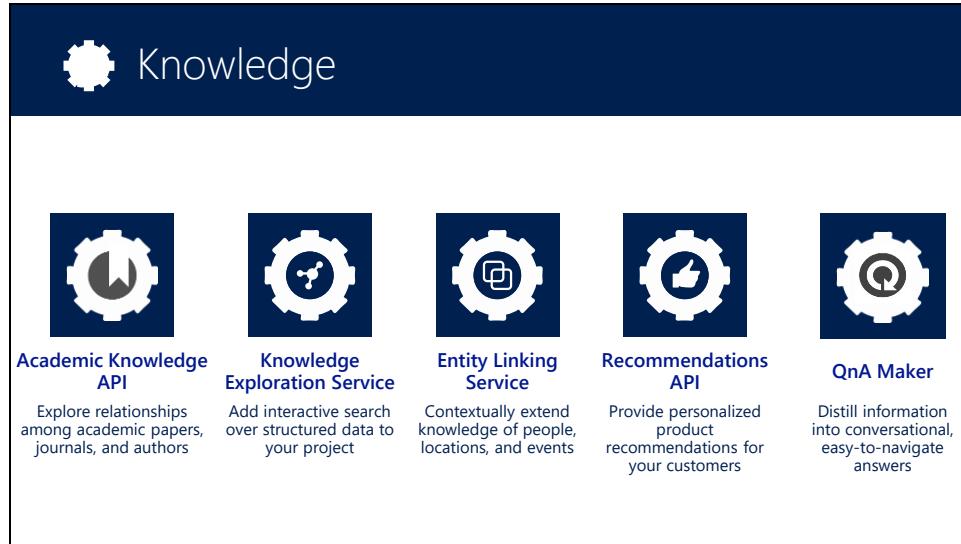
Sentiment analysis : The API returns a numeric score between 0 and 1. Scores close to 1 indicate positive sentiment and scores close to 0 indicate negative sentiment. Sentiment score is generated using classification techniques. The input features of the classifier include n-grams, features generated from part-of-speech tags, and word embeddings. English, French, Spanish and Portuguese text are supported.

Key phrase extraction : The API returns a list of strings denoting the key talking points in the input text. We employ techniques from Microsoft Office's sophisticated Natural Language Processing toolkit. English, German, Spanish, and Japanese text are supported.

Topic detection : This is a newly released API that returns the detected topics for a list of submitted text records. A topic is identified with a key phrase, which can be one or more related words. This API requires a minimum of 100 text records to be submitted, but is designed to detect topics across hundreds to thousands of records. Note that this API charges 1 transaction per text record submitted. The API is designed to work well for short, human-written text such as reviews and user feedback.

Language detection : The API returns the detected language and a numeric score between 0 and 1. Scores close to 1 indicate 100% certainty that the identified language is true. A total of 120 languages are supported.

Microsoft Translator: Add speech translation, for any of the 9 supported languages, and text translation, for any of the 60 supported languages, to your app. Grow your potential user base by localizing your app and its content with clear translations.



Academic Knowledge API: The Academic Knowledge API enable developers to interpret user queries for academic intent and retrieve rich entity information about research papers, authors, journals, conferences, and universities from the Microsoft Academic Graph (MAG). 1. Developers can use this API to build search features such as the knowledge-based query auto-suggest and search results at <http://academic.microsoft.com>. 2. Universities can retrieve analytics data about their researchers' publications, topics, and venues. 3. Conference organizers can analyze the citation patterns of their conference papers. 4. Data scientists and computer science researchers can develop new ranking and analysis approaches over a large heterogeneous network.

A publication about our approach can be found here:

<http://research.microsoft.com/apps/pubs/default.aspx?id=246609> The underlying data graph (Microsoft Academic Graph) is constructed over discovered academic papers on the web. The data is largely unstructured, of variable quality, and ambiguous. We are constantly working to improve this data quality and to correctly aggregate the multiple instances of authors, papers, etc. into a single representative entity in the graph.

Knowledge Exploration Service API: Enable interactive search experience over structured data via natural language. Indexes customer's structured data (BYOD) with support for prefix completion. Generates annotated interpretations/completions of natural language queries. Publishers/libraries can use this to create interactive search over their publications similar to academic.microsoft.com. Merchants can create interactive search experience to help users find and discover products. Data owners can create systems that answer natural language user questions. It works best when the structured data is clean and the natural language structure is simple and predictable. Otherwise, customers will have to invest a bit of work to generate the structured data and author the grammar. We are working on future components to simplify both authoring aspects.

Entity Linking Service API: Given a specific paragraph of text within a document, the Entity Linking will recognize and identify each separate entity based on the context and linking the entity to wikipedia. Use cases: A news agency would use this to analysis their news article to create relations between articles, a news agency would use this to generate tags for article and make recommendation for reader, a company would use this to track the PR

articles mentioned it and product comments to track customer feedback. We also have SDKs available for Entity Linking.

Recommendations API: With Recommendations API, provide personalized product recommendations for your customer and improve sales in your store.

Frequently Bought Together (FBT) recommendations : Learn from your previous transactions. When a customer visits a particular item, the recommendations engine suggests additional items that are likely to be purchased together in the same transaction.

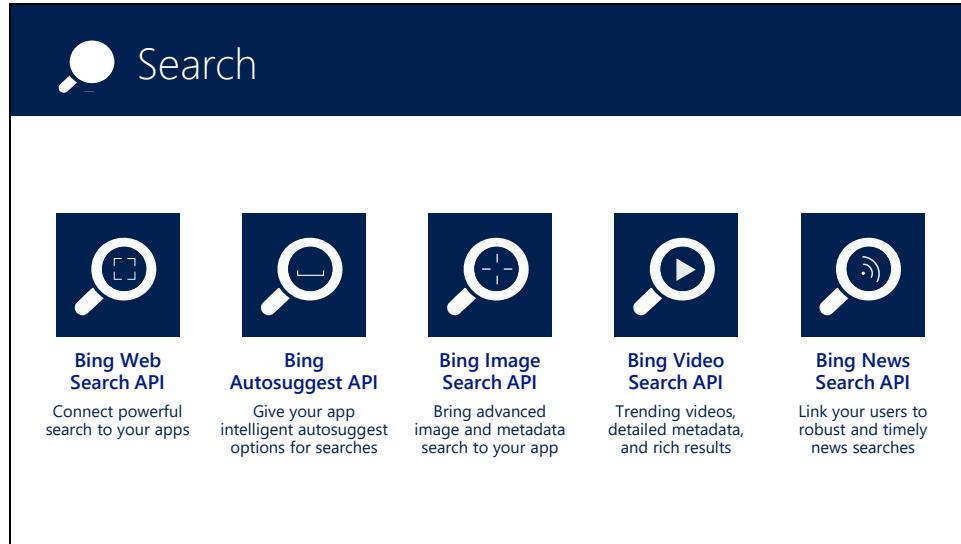
Item to item recommendations : This is the "Customers who liked this product also liked these other products" scenario. Increase the discoverability of items in your catalog by showing relevant products to your customers.

Personalized user recommendations : Using a customer's prior activity, personalize their experience by recommending items that they might be interested in. For example, using a customer's viewing history for movies, it's possible to recommend additional movies and shows of interest.

QnA Maker: With QnA Maker, extract all possible pairs of questions and answers from user provided content – FAQ URLs, documents and editorial content

Test, train and publish: Edit, remove or add pair before testing and training the knowledge base and publishing your knowledge base as an API endpoint

Integrates with other APIs and solutions : Use QnA Maker with Cognitive Services such as LUIS & create something as elegantly simple as a chat bot that answers FAQs, or as complex as an interactive virtual guide.



The Search APIs provide access to the search technology that powers Bing.com and a long list of 1st-party (Office, Cortana, Xbox, Edge) and 3rd-party (AOL, Apple, Amazon, Yahoo etc.) partners. In total, we have 4 Search APIs for web, image, video and news search. In addition we provide access to our Autosuggest (type-ahead) and Spell Check services.

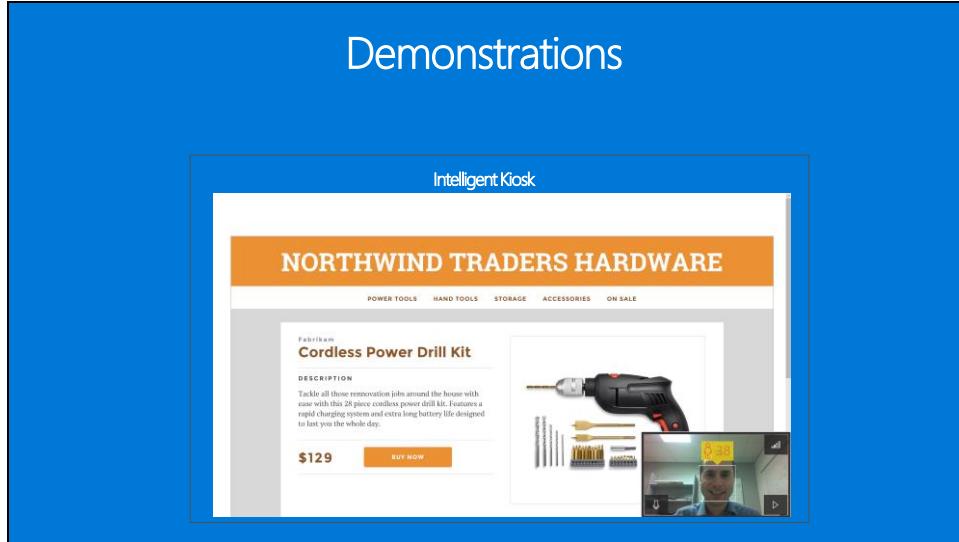
Bing Web Search API: is the main Search API. With one programmatic call, the user can get back relevant webpage, image, video, and news results. It's equivalent to searching in the web vertical in bing.com. The Search APIs provide the power of the Bing Search engine to developers. You get access to many of the latest and greatest capabilities the bing.com site has to offer, in the form of an API. It's the same architecture stack so it's a powerful way to bring the knowledge and intelligence of the search engine to your own experience. The API is built directly on top of the entire Bing stack- you get the quality, relevance, performance and continuous improvements that the rest of the site gets. Similar to Bing.com- tail queries, or obscure terms may have limited results. We're always working to improve the edge cases though and API users will benefit from that.

Bing Autosuggest API: At the time of publication, this data was not available. Please email Rebecca Duffy, reduffy@microsoft.com, if you would like more information.

Bing Image Search API: At the time of publication, this data was not available. Please email Rebecca Duffy, reduffy@microsoft.com, if you would like more information.

Bing Video Search API: At the time of publication, this data was not available. Please email Rebecca Duffy, reduffy@microsoft.com, if you would like more information.

Bing News Search API: At the time of publication, this data was not available. Please email Rebecca Duffy, reduffy@microsoft.com, if you would like more information.



You can also find the following interactive demonstrations in the Cognitive Services Site :

Vision

[Computer Vision](#)
[Content Moderator](#)
[Emotion](#)
[Face](#)
[Video](#)

Speech

[Bing Speech](#)
[Custom Recognition](#)
[Speaker Recognition](#)

Language

[Bing Spell Check](#)
[Language Understanding](#)
[Linguistic Analysis](#)
[Text Analytics](#)
[Translator](#)
[WeBLM](#)

Knowledge

[Academic](#)
[Entity Linking](#)
[Knowledge Exploration](#)
[QnA Maker](#)
[Recommendations](#)

Search

[Bing Autosuggest](#)
[Bing Image Search](#)
[Bing News Search](#)

[Bing Video Search](#)

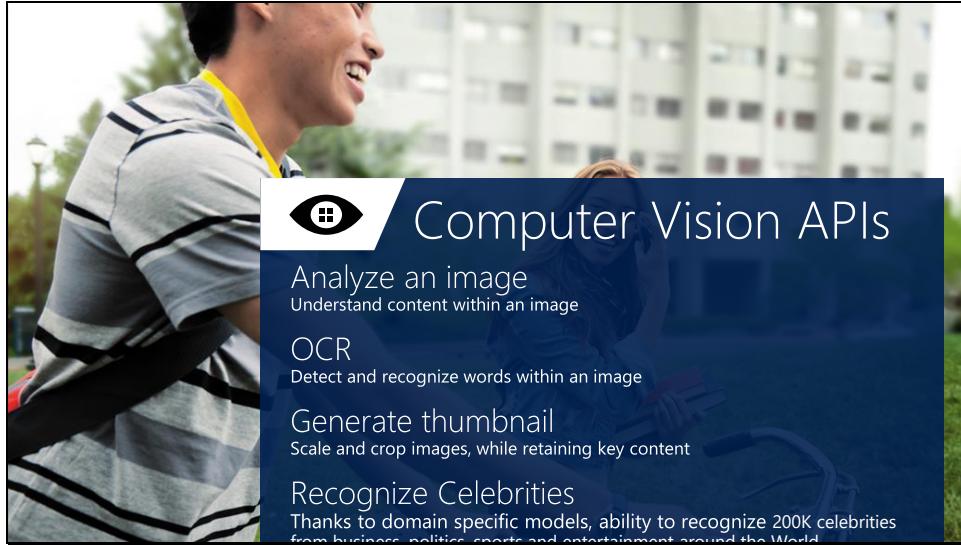
[Bing Web Search](#)

Vision



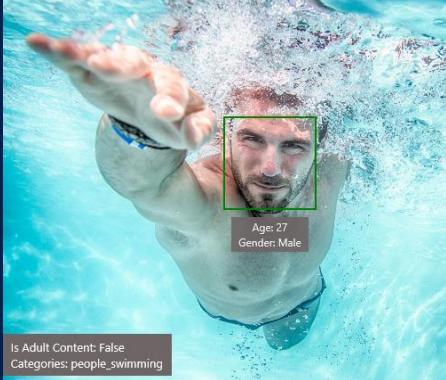
From faces to feelings, allow your
apps to understand images and video

Computer Vision | Emotion | Face |
Video | Content Moderator



Do you need an API that gives you actionable information about images used in your app? The Computer Vision API gives you the tools to understand the contents of any image. Create tags identifying objects, beings, or actions present in the image, and then craft coherent sentences to describe it. Whether you want to execute better image search, or you want to create an assistive app for the visually impaired, the Computer Vision API helps get the job done.

Analyze image



Type of image

Clip Art Type	0 Non-clipart
Line Drawing Type	0 Non-Line Drawing
Black & White Image	False

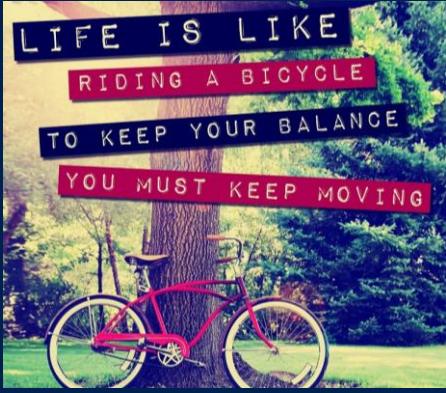
Content of image

Categories	[{"name": "people_swimming", "score": 0.099609375}]
Adult Content	False
Adult Score	0.18533889949321747
Faces	[{"age": 27, "gender": "Male", "faceRectangle": {"left": 472, "top": 258, "width": 199, "height": 199}]]

Image colors

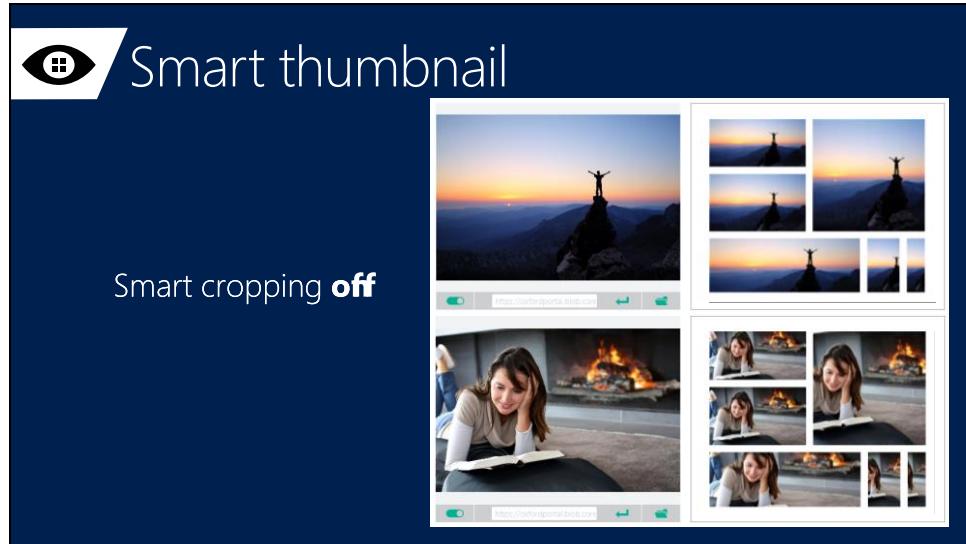
Dominant Color Background	White
Dominant Color Foreground	Grey
Dominant Colors	White
Accent Color	

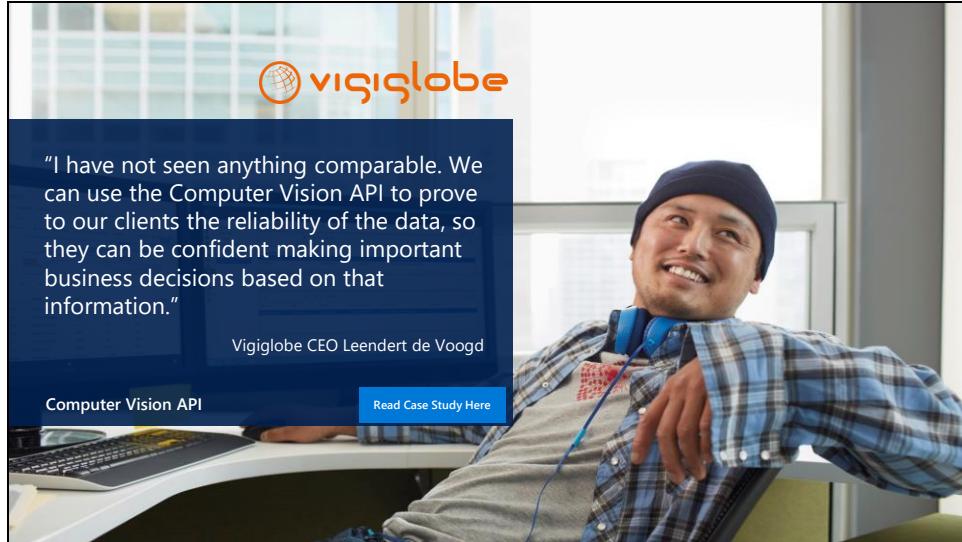
OCR



Good at
Scanned documents
Photos with text
Fine grained location information

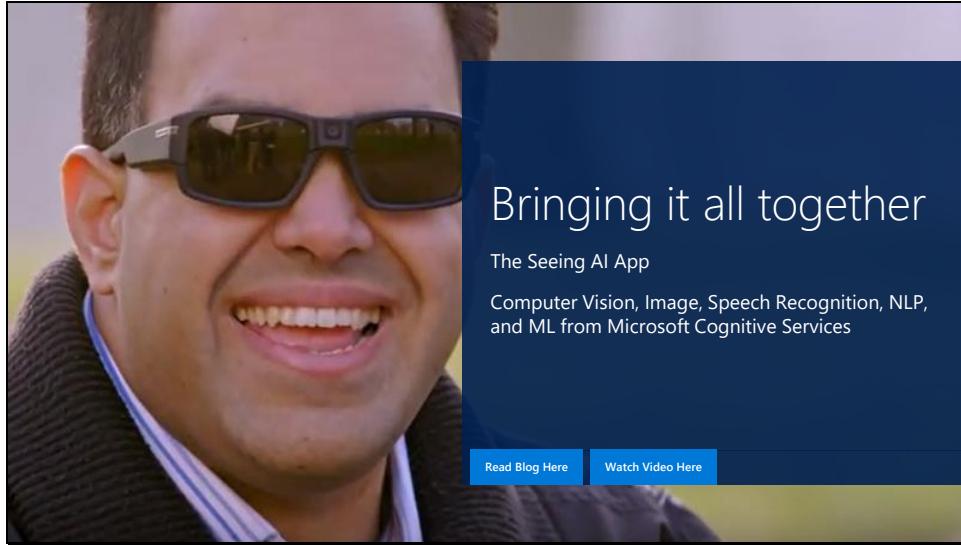
```
{"language": "en", "orientation": "up", "photosWithText": [{"text": "LIFE IS LIKE RIDING A BICYCLE TO KEEP YOUR BALANCE YOU MUST KEEP MOVING"}], "fineGrainedLocation": [{"line": "LIFE", "bbox": "41, 102, 150, 150"}, {"line": "IS", "bbox": "160, 102, 210, 150"}, {"line": "LIKE", "bbox": "220, 102, 270, 150"}, {"line": "RIDING", "bbox": "280, 102, 330, 150"}, {"line": "A", "bbox": "340, 102, 390, 150"}, {"line": "BICYCLE", "bbox": "400, 102, 450, 150"}, {"line": "TO", "bbox": "460, 102, 510, 150"}, {"line": "KEEP", "bbox": "520, 102, 570, 150"}, {"line": "YOUR", "bbox": "580, 102, 630, 150"}, {"line": "BALANCE", "bbox": "640, 102, 690, 150"}, {"line": "YOU", "bbox": "700, 102, 750, 150"}, {"line": "MUST", "bbox": "760, 102, 810, 150"}, {"line": "KEEP", "bbox": "820, 102, 870, 150"}, {"line": "MOVING", "bbox": "880, 102, 930, 150"}]}
```

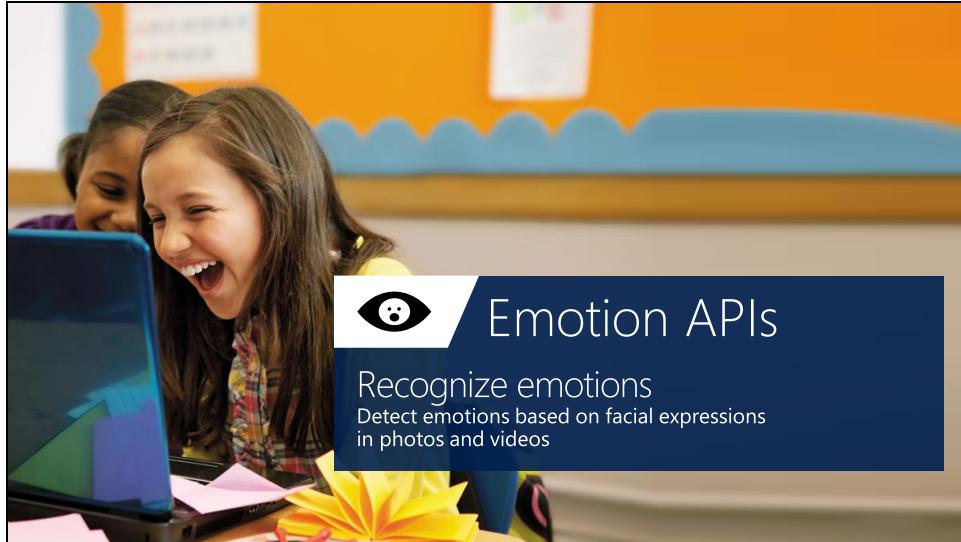




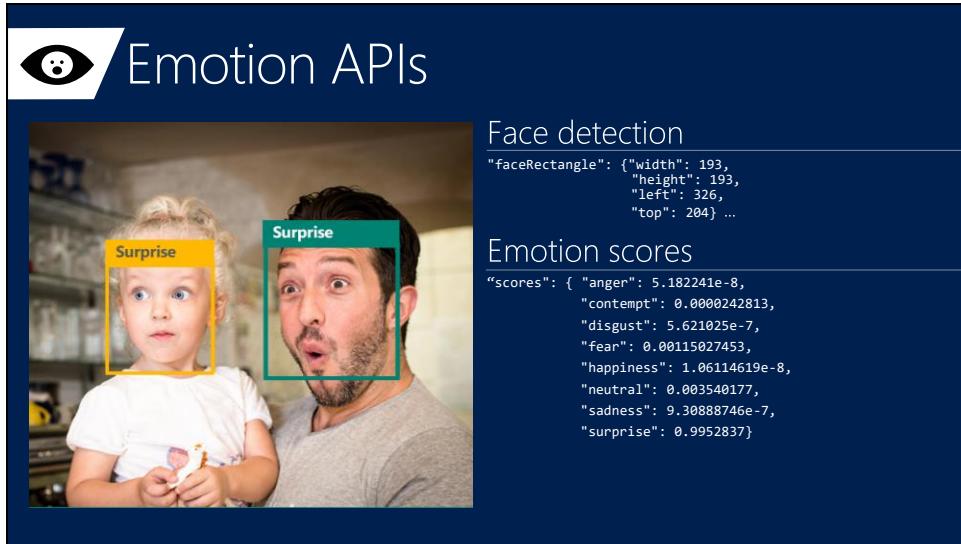
Vigiglobe

Vigiglobe saw an opportunity to analyze not only what was being said on social media, but also the context in which it was being discussed. The team created proprietary algorithms to accurately interpret and contextualize social media messages in real time. Using the Computer Vision API of Microsoft





How are you feeling? Can your app tell? With the Emotion API, you can build an app that recognizes emotions according to facial expressions—giving you the capability to provide an amazing, personalized experience. Using facial expressions, this cloud-based API can detect happiness, neutrality, sadness, contempt, anger, disgust, fear, and surprise. The AI understands these emotions based on universal facial expressions, and it functions cross-culturally, so your app will work around the world.



The slide features a logo of an eye with a speech bubble and the text 'Emotion APIs'.

Face detection

```
"faceRectangle": { "width": 193, "height": 193, "left": 326, "top": 204} ...
```

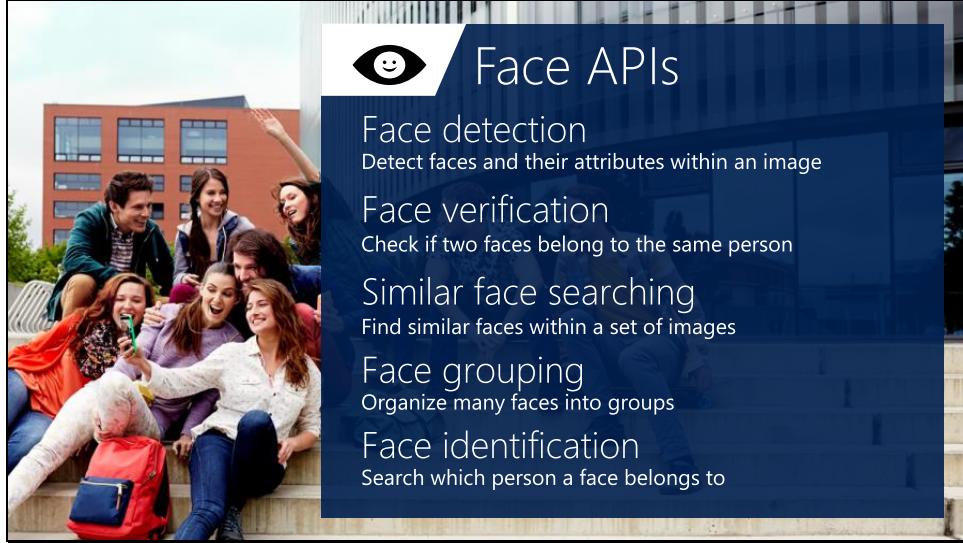
Emotion scores

```
"scores": { "anger": 5.182241e-8, "contempt": 0.0000242813, "disgust": 5.621025e-7, "fear": 0.00115027453, "happiness": 1.06114619e-8, "neutral": 0.003540177, "sadness": 9.30888746e-7, "surprise": 0.9952837}
```

Lab

Consuming Emotion API

Go through Emotion python notebook at <https://notebooks.azure.com/library/python-cognitive>



Face APIs

Face detection
Detect faces and their attributes within an image

Face verification
Check if two faces belong to the same person

Similar face searching
Find similar faces within a set of images

Face grouping
Organize many faces into groups

Face identification
Search which person a face belongs to

Detect human faces and compare similar ones, organize people into groups according to visual similarity, and identify previously tagged people in images.

Face APIs



Detection

```
"faceRectangle": {"width": 193, "height": 193, "left": 326, "top": 204}  
...  
"faceId": "12345678901234567890123456789012",  
"faceWidth": 193, "faceHeight": 193, "faceLeft": 326, "faceTop": 204}
```

Feature attributes

```
"attributes": { "age": 42, "gender": "male",  
"headPose": { "roll": "8.2", "yaw": "-37.8", "pitch": "0.0" }}
```

Grouping

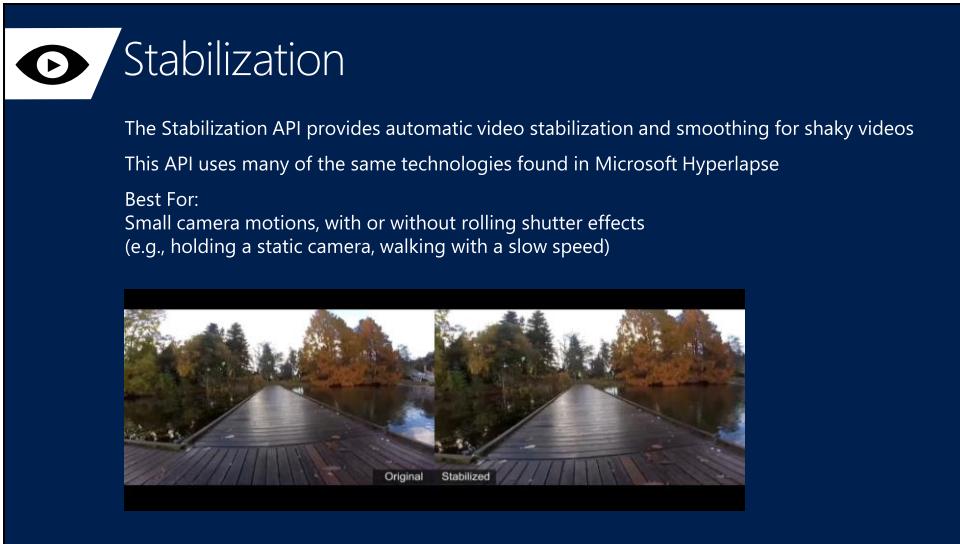


Identification

Jasper Williams



Intelligent video processing produces stable video output, detects motion, creates intelligent thumbnails, and detects and tracks faces.



The Stabilization API provides automatic video stabilization and smoothing for shaky videos. This API uses many of the same technologies found in Microsoft Hyperlapse. Best For: Small camera motions, with or without rolling shutter effects (e.g., holding a static camera, walking with a slow speed)

Original Stabilized

A side-by-side comparison of two images of a wooden pier over a lake. The left image, labeled 'Original', shows noticeable vertical shake and slight rolling shutter effects. The right image, labeled 'Stabilized', appears much smoother and more stable.

Face detection and tracking

High precision face location detection and tracking
Can detect up to 64 human faces in a video (no smaller than 24x24 pixels)
Detected and tracked faces are returned with coordinates and a Face ID
to track throughout the video

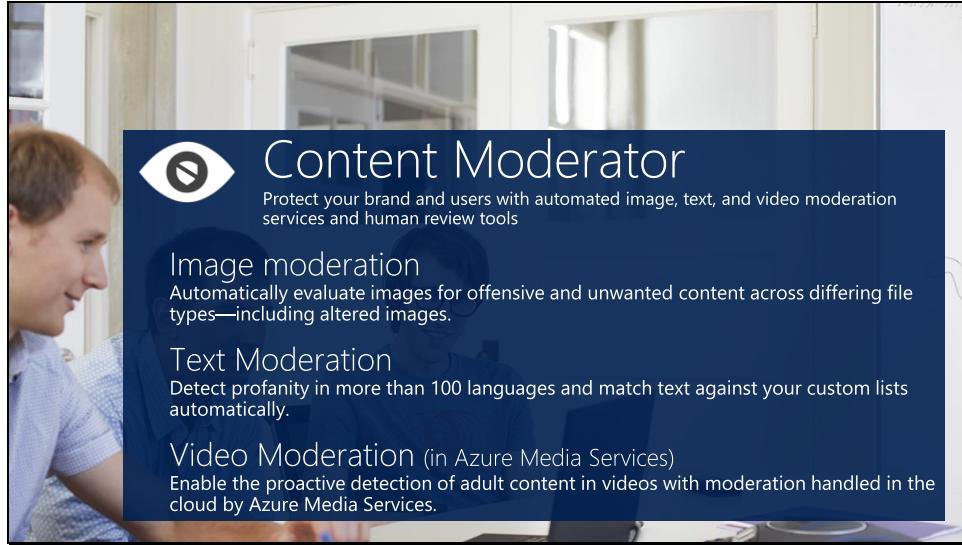
Time (sec)	Face ID	x, y	Width, height
0	0	0.59, 0.23	0.09, 0.16
0	1	0.38, 0.15	0.07, 0.12
1	0	0.54, 0.25	0.09, 0.15
1	1	0.23, 0.18	0.07, 0.12

Motion detection

Indicates when motion occurs against a fixed background (e.g., surveillance video)
Trained to reduce false alarms, such as lighting and shadow changes



Start time	End time	In region
1.9	3.6	0
5.2	15.1	0



Powered by intelligent machine learning, Microsoft Content Moderator automatically filters out offensive content in images, text, and video across platforms and includes human review tools for more nuanced cases.

Image moderation API

Automatically evaluate images for offensive and unwanted content across differing file types—including altered images. Content Moderator includes optical character recognition (OCR) and face detection to improve the chances of detecting unwanted or exploitative content and implements fuzzy matching against custom blacklists and whitelists.

Text moderation API

Detect profanity in more than 100 languages and match text against your custom lists automatically. Content Moderator also checks for malware and phishing URLs plus personally identifiable information (PII).

Video moderation API

Enable the proactive detection of adult content in videos with moderation handled in the cloud by Azure Media Services.

Human review tool beta

Enjoy the freedom and control of human oversight while still benefiting from automated moderation. Review tools let you review automated results with your team and approve or change tags to override the automated results. Thanks to machine learning and custom lists, your moderation process gets smarter the more you use it.

[Try the review tool beta](#)

Lab

Consuming Computer Vision API

Go through Computer Vision python notebook at <https://notebooks.azure.com/library/python-cognitive>

Speech



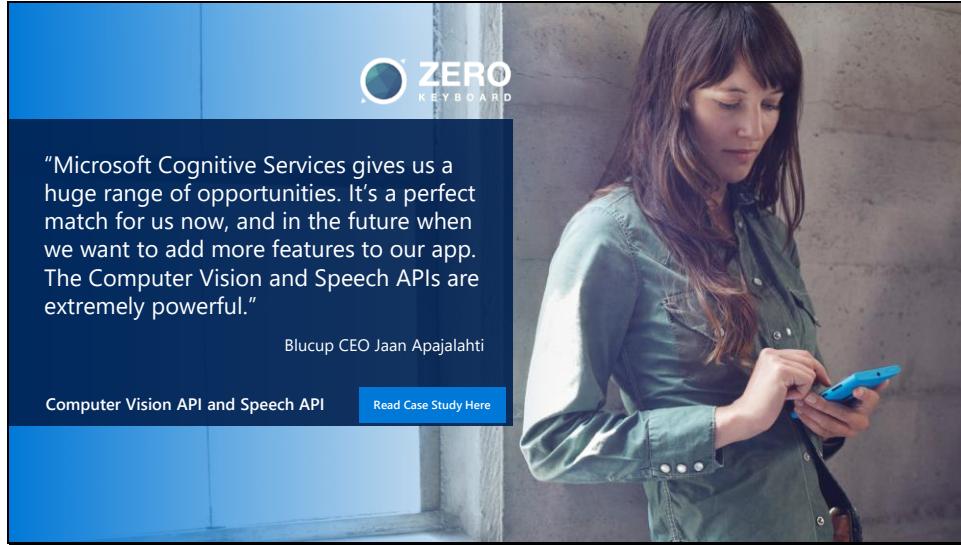
Hear and speak to your users by filtering noise,
identifying speakers, and understanding intent

Custom Recognition | Speaker Recognition | Speech



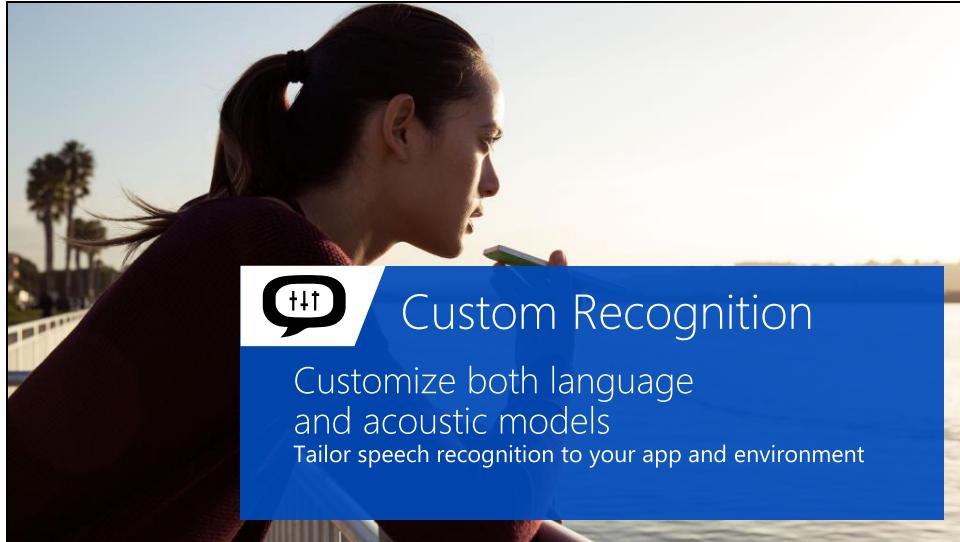
Convert audio to text, understand intent, and convert text back to speech for natural responsiveness.

Speech Intent Recognition can, in addition to returning recognized text from audio inputs, also return structured information about the speech to apps that parse the intent of the speaker and drive further actions by the app. Models trained by [LUIS](#) service are used to generate the intent.



Blucup

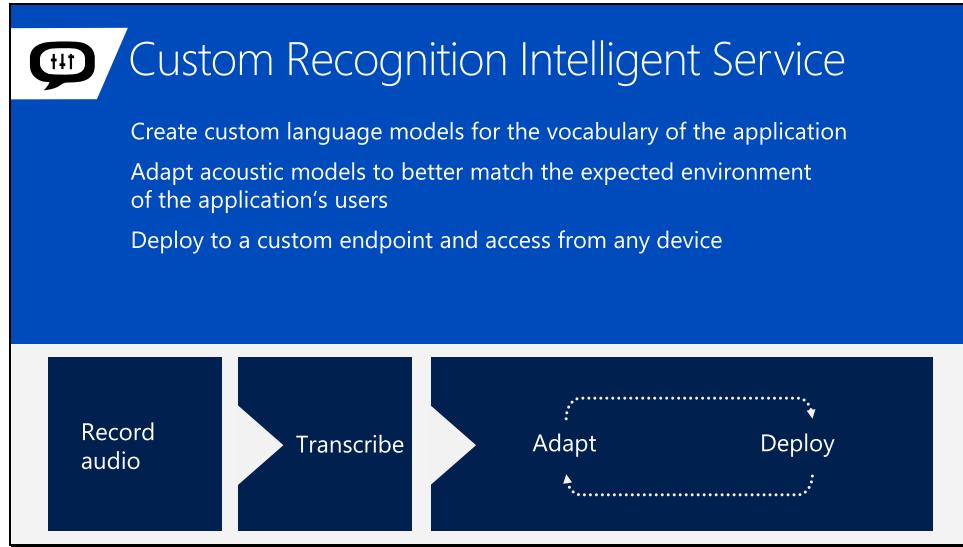
Finland-based Blucup was on a mission to solve a common problem: how can salespeople capture data while on the go? The company developed the Zero Keyboard app, which sales reps could use to record customer information quickly and add it automatically to their customer relationship management (CRM) systems using touch gestures, voice, and pictures. Using the Speech and Computer Vision APIs from Microsoft Cognitive Services, Blucup provides customers with accurate results and rich features—all while speeding development internally.



Eliminate speech recognition barriers like speaking style, background noise, and vocabulary.

Does your speech recognition work with varied user populations, vocabularies, or with background noise? The Custom Recognition Intelligent Service (CRIS) helps you create custom speech recognition endpoints—so accents and environments are features, not challenges.

Customize your speech recognition by vocabulary and speaking style, create custom acoustic models to match the expected environment of your users, and tap into the API's powerful intelligence to create speech recognition endpoints customized to your app's needs. Turn language barriers into features with CRIS.





Your users' voices are their passports with the Speaker Recognition API. Your app can authenticate identities by using someone's voice, giving your users the capability to interact securely through speech.

Speaker recognition APIs

The diagram illustrates the three main components of speaker recognition:

- Enrollment:** Create a unique voiceprint for a profile. It shows a green waveform icon with a question mark and an arrow pointing to a person labeled "Anna".
- Recognition:** After enrolling one or more voices, identify who is speaking from an audio clip. It shows a green waveform icon with a question mark and a checkmark, with arrows pointing to "Mike" and "Anna".
- Verification:** Confirm if a voice belongs to a previously enrolled profile. It shows a green waveform icon with a question mark and a checkmark, with arrows pointing to "Mike" and "Anna".

Language



Process text and learn how to recognize what
users want

Bing Spell Check | Language Understanding |
Linguistic Analysis | Text Analytics | Web Language Model |
Translator Text and Speech



9 # Bing spell check APIs

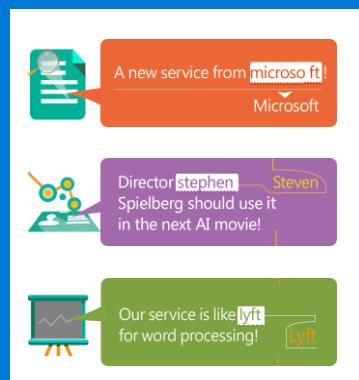
State-of-the-art cloud based spelling algorithms
Recognizes a wide variety of spelling errors

Recognize name errors and homonyms in context
Difficult to spot errors that use the context of the words around them

Updates over time
Support for new brands and coined expressions as they emerge

The Bing Spell Check API corrects spelling errors, contextually recognizes names and slang, understands homonyms, and supports brand names.

#9 / Bing spell check APIs



A new service from Microsoft

Director **stephen** Steven Spielberg should use it in the next AI movie!

Our service is like **lyft** for word processing!

Check a single word or a whole sentence
"Our engineers developed this **four** you!"
Corrected Text: "four" → "for"

Identify errors and get suggestions

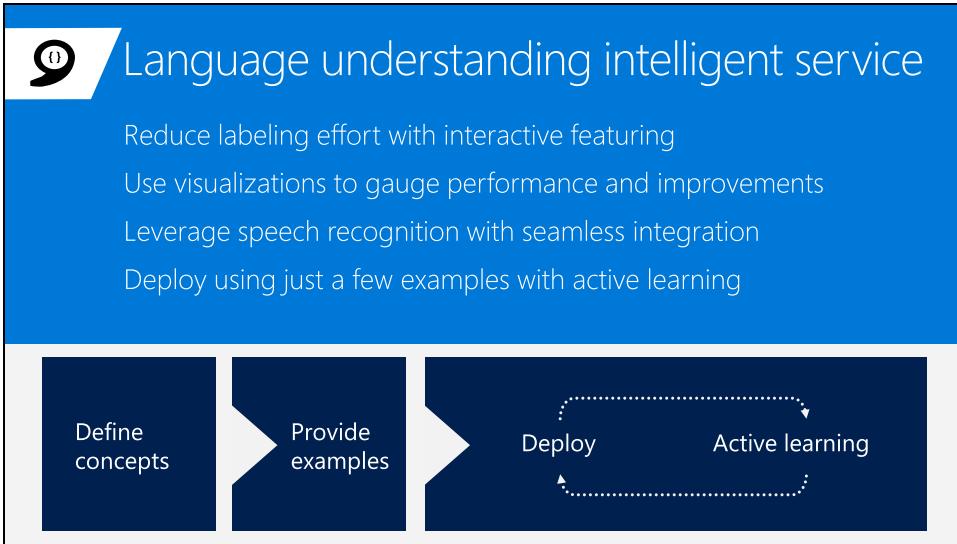
```
spellingErrors": [ { "offset": 5, "token": "gona", "type": "UnknownToken", "suggestions": [ { "token": "gonna" } ] }
```



Understand language contextually, so your app communicates with people in the way they speak.

Do your apps understand language in the way people speak it—contextually? With the Language Understanding Intelligent Service (LUIS) API, you can integrate language models that understand your users quickly and easily. And if one of our preexisting models won't work, it will guide you through building your own.

Prebuilt models will recognize places, times, numbers, and temperatures, and handle common requests like "set an alarm for 8 AM." LUIS supports dialogue and action fulfillment, so your users can carry on a conversation with your app. For example, the input "schedule a meeting with Allison" results in the question "when?," allowing the user to respond "3 PM," and the meeting gets scheduled.



The diagram illustrates a process flow for a Language understanding intelligent service. It begins with a blue header section containing a black speech mark icon and the text 'Language understanding intelligent service'. Below this, four bullet points describe features: 'Reduce labeling effort with interactive featuring', 'Use visualizations to gauge performance and improvements', 'Leverage speech recognition with seamless integration', and 'Deploy using just a few examples with active learning'. The main process flow consists of four dark blue rectangular boxes connected by white arrows. The first box on the left is labeled 'Define concepts'. The second box is labeled 'Provide examples'. The third box is labeled 'Deploy'. The fourth box is labeled 'Active learning'. The arrows indicate a sequential flow from 'Define concepts' to 'Provide examples', then to 'Deploy', and finally to 'Active learning'. There are also small dotted arrows pointing from 'Deploy' back to 'Define concepts' and from 'Active learning' back to 'Deploy', suggesting a feedback loop.

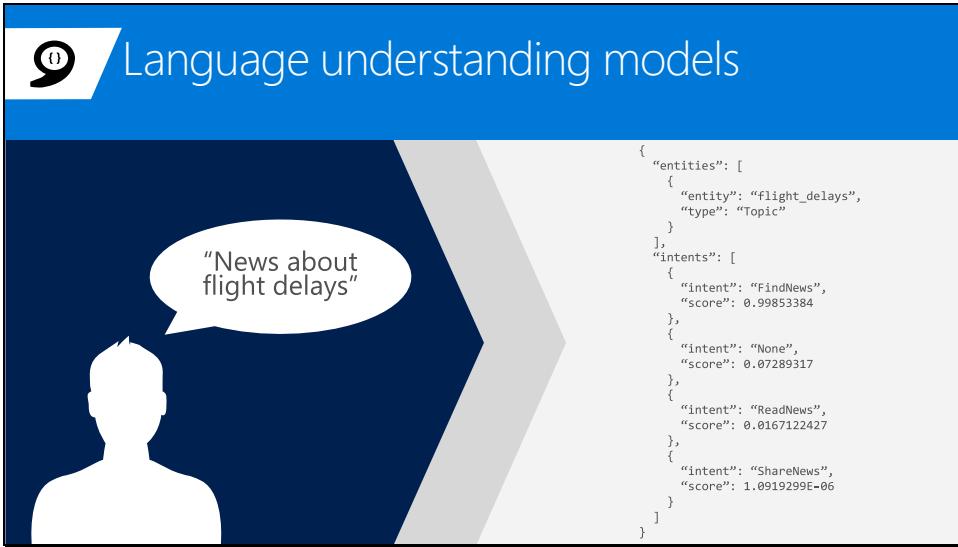
- Reduce labeling effort with interactive featuring
- Use visualizations to gauge performance and improvements
- Leverage speech recognition with seamless integration
- Deploy using just a few examples with active learning

Define concepts

Provide examples

Deploy

Active learning



Language understanding models

```
{ "entities": [ { "entity": "flight_delays", "type": "Topic" } ], "intents": [ { "intent": "FindNews", "score": 0.99853384 }, { "intent": "None", "score": 0.07289317 }, { "intent": "ReadNews", "score": 0.0167122427 }, { "intent": "ShareNews", "score": 1.0919299E-06 } ] }
```

Slide 51

LUIS

ExerciseTracker

Intents

- None
- StartActivity
- StopActivity
- SetHeartRateTarget

Entities

- ActivityType

Pre-built Entities

- None

Regex Features

No patterns added

Phrase List Features

ActivityWords

Suggest

New utterances

Search

Review labels

Select highlight to add another entity or click to clear

begin a now

Intent: StartActivity(0.96)

Submit

begin now

Intent: StartActivity(0.94)

Submit

start tracking a now

Intent: StartActivity(0.94)

Submit

Performance analysis

Intents

StartActivity

11 utterances: 11 correctly predicted

StopActivity

5 utterances: 5 correctly predicted

SetHeartRateTarget

5 utterances: 5 correctly predicted

None

17 utterances: 17 correctly predicted

Correctly predicted

Error (predicted as other intent)

Train Last train completed: 8/21/2015 2:04:09 PM

Microsoft

Lab

Create a LUIS app

<https://luis.ai> and follow the short “Luis Application” section here <https://github.com/Microsoft/BotBuilder-Samples/tree/master/Node/intelligence-LUIS#luis-application>. Test the model with queries and enter in some new utterances.



You know what your users are saying, but do you know what it means? The Linguistic API uses advanced linguistic analysis tools for natural language processing, giving you access to part-of-speech tagging and parsing. These tools allow you to hone in on important concepts and actions. The API can tap into traditional linguistic analysis tools that allow you to identify the concepts and actions in your text with part-of-speech tagging, and find phrases and concepts using natural language parsers. Whether you're mining customer feedback, interpreting user commands, or consuming web text, understanding the structure of the text is a critical first step. Try it out below!

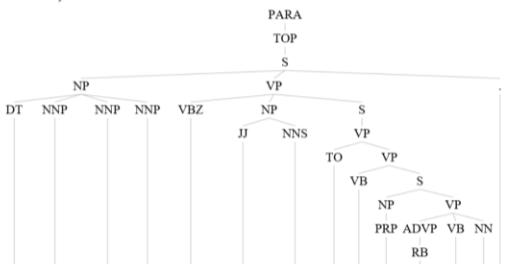
Slide 54

 Linguistic Analysis

Enter a sentence
The Linguistic Analysis API simplifies complex languages to help you easily parse text.

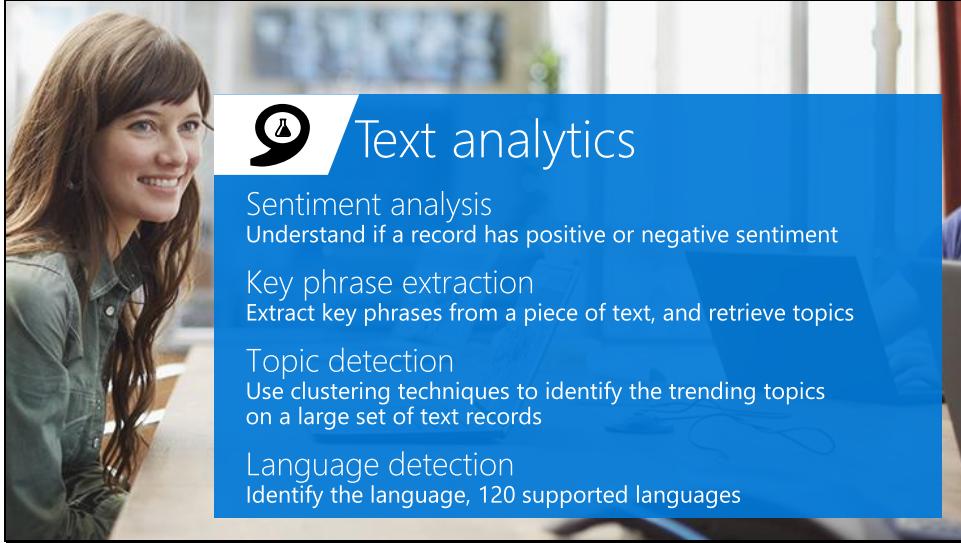
POS tags
[["DT", "NNP", "NNP", "NNP", "VBZ", "JJ", "NNS", "TO", "VB", "PRP", "RB", "VB", "NN", "."]]

Constituency tree



```
graph TD; PARA --- TOP; TOP --- S1[S]; S1 --- VP1[VP]; VP1 --- NP1[NP]; NP1 --- DT1[DT]; NP1 --- NNP1[NNP]; NP1 --- NNP2[NNP]; NP1 --- NNP3[NNP]; VP1 --- VBZ1[VBZ]; VP1 --- JJ1[JJ]; VP1 --- NNS1[NNS]; VP1 --- TO1[TO]; VP1 --- VB2[VB]; VP1 --- NP2[NP]; VP1 --- VP2[VP]; VP2 --- PRP1[PRP]; VP2 --- ADVP1[ADVP]; VP2 --- VB3[VB]; VP2 --- RB1[RB]; VP2 --- NN1[NN]; S1 --- S2[S]; S2 --- VP3[VP]; VP3 --- VB4[VB]; VP3 --- NP3[NP]; VP3 --- VP4[VP];
```

The Linguistic Analysis API simplifies complex languages to help you easily parse text .



Text analytics

Sentiment analysis
Understand if a record has positive or negative sentiment

Key phrase extraction
Extract key phrases from a piece of text, and retrieve topics

Topic detection
Use clustering techniques to identify the trending topics on a large set of text records

Language detection
Identify the language, 120 supported languages

Understanding and analyzing unstructured text is an increasingly popular field and includes a wide spectrum of problems such as sentiment analysis, key phrase extraction, topic modeling/extraction, aspect extraction and more.

Lab

Text Analytics Key and API Reference

Follow https://github.com/Azure/bot-education/blob/master/Student-Resources/Labs/CSharp/CognitiveServicesLab_API_ref.md

Lab

Consuming Text Analytics API

Go through Text Analytics python notebook at <https://notebooks.azure.com/library/python-cognitive>



Ziosk

*"Thanks to **Text Analytics** by Azure Machine Learning, we're able to incorporate guest sentiment into our actionable guest feedback platform to deliver a comprehensive view of guest satisfaction and server performance."*

Al Pappa
Head of Business Intelligence
Ziosk
Text Analytics API

Get your appetizer order to the kitchen ASAP

Order another round as soon as your glass is empty

Order dessert when the craving strikes

Case study coming soon



Automate a variety of standard natural language processing tasks using state-of-the-art language modeling APIs. Do you need to know how frequently certain words appear together? Or figure out which words a user might type next? Or how to break a hashtag into individual words? The Web Language Model API lets your app do all of this quickly and accurately.



- . Try now the Speech Translator demo App on : <https://github.com/MicrosoftTranslator/SpeechTranslator>
- . Try now the Document Translator demo, translating (batches of) Word or pdf documents while preserving the formatting : <https://www.microsoft.com/en-us/translator/doctranslator.aspx>

Language detection

The Translator Text API automatically detects the language of the text that's sent before translating it. If your application simply needs to know what language the text is in, you can also call the API to detect the language of any text string.

Translation

Add speech translation, for any of the 9 supported languages, and text translation, for any of the 60 supported languages, to your app. Grow your potential user base by localizing your app and its content with clear translations.

Custom translation system

Build a custom translation system, using as little as 1,000 parallel sentences or start out simply by providing a dictionary of company specific words.

Collaborative Translation Framework (CTF)

Improve translations by creating a specific user group that provides suggestions to improve the translations. Users suggest translations and designated approvers either approve or deny changes. These updated translations can then be used for the company's specific Hub to further improve its custom system.

Knowledge

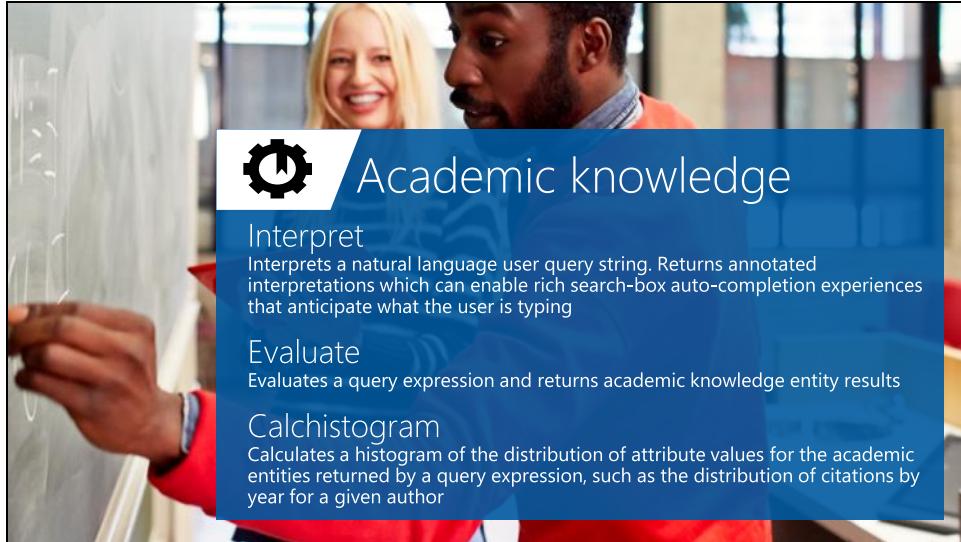


Tap into rich knowledge amassed from
the web, academia, or your own data

Academic Knowledge

Entity Linking | Knowledge Exploration

Recommendations | QnA Maker



Tap into the wealth of academic content in the Microsoft Academic Graph using the Academic Knowledge API. Do your users know who the top scholars have been in machine learning over the last three years? What about every paper authored by an expert like Li Deng from Microsoft? The Academic Knowledge API answers these and other questions by applying the Knowledge Exploration Service to the Microsoft Academic Graph. Users can start from natural language queries, or you can ping the graph directly through structured query expressions. Additionally, the Academic Knowledge API can auto-complete natural language queries and return entity results, helping users narrow research results faster. It can also create a histogram of attribute values for academic entries returned by a query—for example, the distribution of papers by year for an author.



Provide better user experiences by adding the Entity Linking Service to your app so that it can provide additional knowledge and facts from the web to supplement the text in context. The Entity Linking Service uses a prebuilt knowledge base to build links, and the option to acquire knowledge from your own data.

For example, your app may need to understand that "London, the capital" is the city of London in the United Kingdom and not London, Ontario, or Jack London, the author.

The Entity Linking Service provides this information quickly and within context, offering a faster, more intelligent user experience.

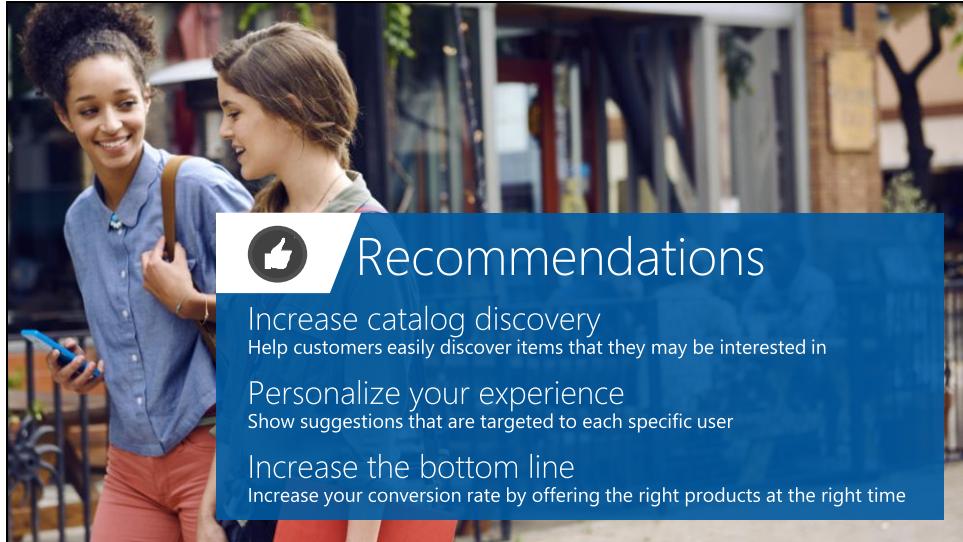
Lab

Consuming Text Analytics API

Go through Entity Linking python notebook at <https://notebooks.azure.com/library/python-cognitive>



Do you have structured data for users to explore via natural language? The Knowledge Exploration Service takes structured data and linguistic resources you provide and creates a service that enables interactive search. For example, as your users enter queries in a search box, the Knowledge Exploration Service offers auto-complete suggestions and semantic annotations. You can retrieve the top matching objects from the data, and you can create histograms of attribute values among the matches.



The Recommendations API helps your customer discover items in your catalog.

Customer activity in your digital store is used to recommend items and to improve conversion in your digital store. The recommendation engine may be trained by uploading data about past customer activity or by collecting data directly from your digital store. When the customer returns to your store you will be able to feature recommended items from your catalog that may increase your conversion rate.



Recommendations API

“By leveraging Cortana Intelligence **Recommendations capabilities** combined with Azure Machine Learning processing power, we have enabled retailers with a Personalized Commerce Experience, allowing them to grow shopper engagement and conversions across all channels.”

Frank Kouretas,
Chief Product Officer at Orkestra
Orkestra.com



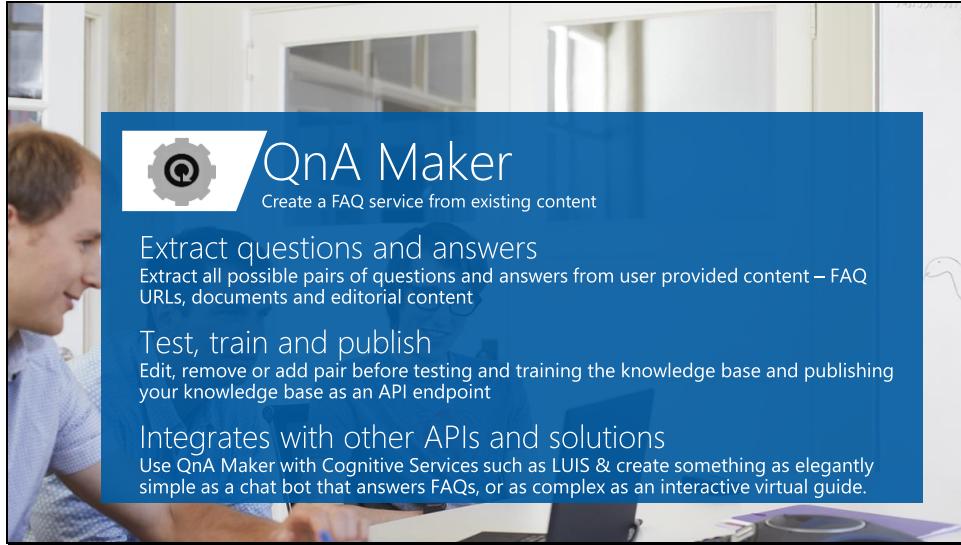


*The **Recommendations API** allows Allrecipes.com to harness billions of user-shared experiences to deliver highly personalized recipe solutions that answer busy families most pressing question, 'What's for dinner tonight?'*

We are able to do this with minimal investment due to the proven capabilities of the Cortana Analytics platform. This helps us further realize our vision of providing highly personalized cooking inspiration for home cooks.

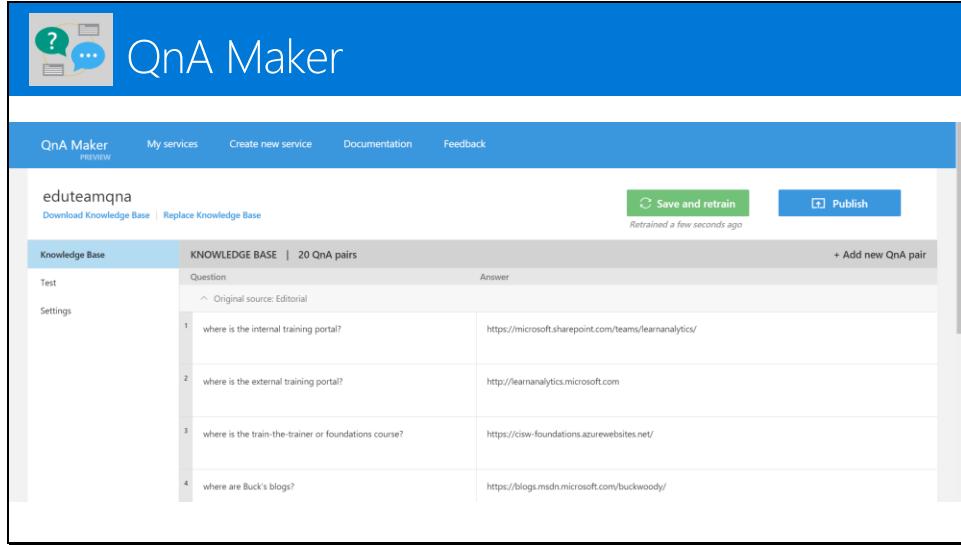
John Keane
CTO, Allrecipes.com





QnA Maker is primarily meant to provide a FAQ data source which you can query from your Bot/Application. Although developers will find this useful, content owners will especially benefit from this tool. QnA Maker is a completely no-code way of managing the content that powers your Bot/Application.

Slide 70



The screenshot shows the QnA Maker Knowledge Base interface. At the top, there is a blue header with the QnA Maker logo and the word "PREVIEW". Below the header, the page title is "QnA Maker". The main content area is titled "eduteamqna" and shows a "KNOWLEDGE BASE" with "20 QnA pairs". On the left, there is a sidebar with "Test" and "Settings" buttons. The main table lists four QnA pairs:

Question	Answer
1 where is the internal training portal?	https://microsoft.sharepoint.com/teams/learnanalytics/
2 where is the external training portal?	http://learnanalytics.microsoft.com
3 where is the train-the-trainer or foundations course?	https://ciscw-foundations.azurewebsites.net/
4 where are Buck's blogs?	https://blogs.msdn.microsoft.com/buckwood/

At the top right, there are "Save and retrain" and "Publish" buttons. The "Save and retrain" button is green and says "Retrained a few seconds ago". The "Publish" button is blue. Below the table, there is a "+ Add new QnA pair" button.

Editing the QnA Maker Knowledge Base

Showing, one of "My services" at <https://qnamaker.ai> or you can "Create a new service" here as well.

Lab

Create a knowledge base in the QnA Maker

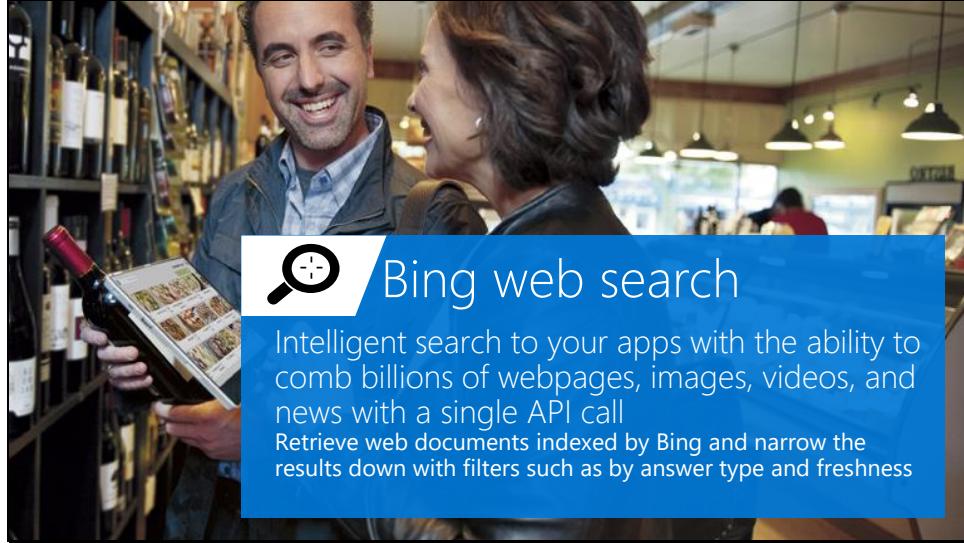
<https://qnamaker.ai>

Search



Access billions of web pages, images, videos,
and news with the power of Bing APIs

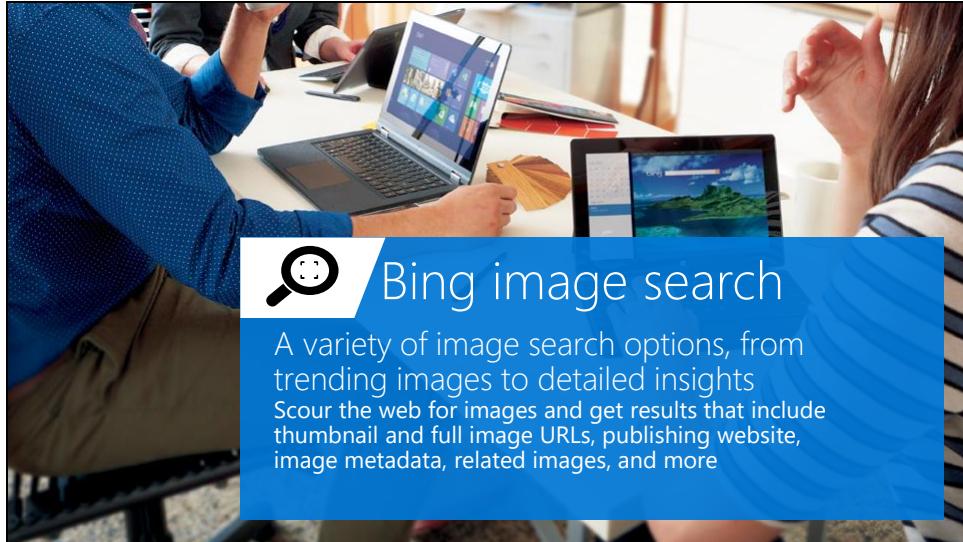
Bing Web Search | Bing Image Search
Bing News Search | Bing Video Search
Bing Auto Suggest



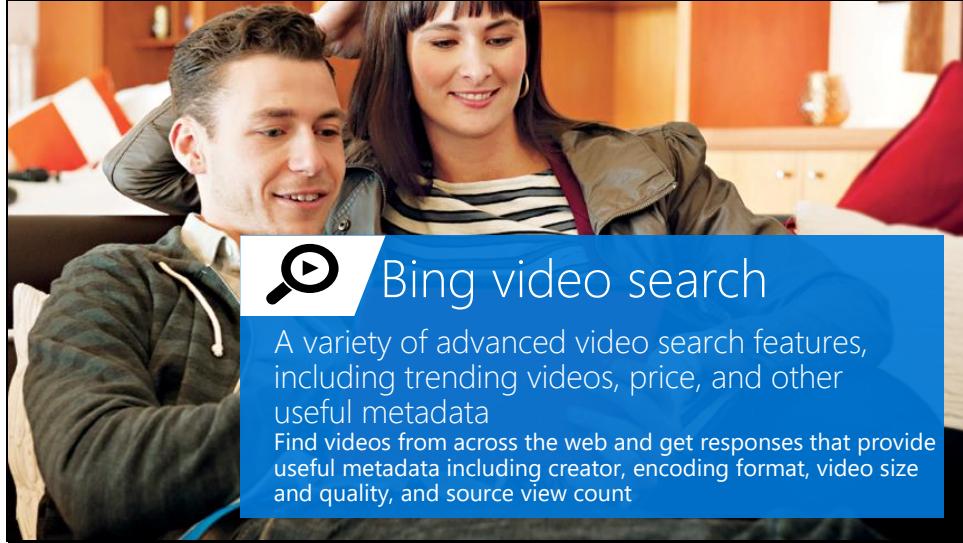
Bing web search

Intelligent search to your apps with the ability to comb billions of webpages, images, videos, and news with a single API call
Retrieve web documents indexed by Bing and narrow the results down with filters such as by answer type and freshness

The Bing Search API adds intelligent search to your app, combing hundreds of billions of webpages, images, videos, and news to provide relevant results with no ad requirements. The results can be automatically customized to your users' locations or markets, increasing relevancy by staying local.



The Bing Image Search API gives you powerful image searching tools with a single call. You can tap into trending images of people, places, and things from around the world, and filter results by image style, size, layout, date added, and license type.



The Bing Video Search API offers robust video searching features with a single API call. You can receive information from around the world about trending videos, updated on a daily basis. Search results can be returned by either a static image or a motion thumbnail, allowing you to customize how your users see what they're looking for.



The Bing News Search API can help turn your app into an up-to-date news center. Results from a single call bring trending news from around the world, which is updated in near-real time, so users can be kept up to date on whatever's happening in their neighborhood—or across the globe.



Whether you're searching the web, a local set of data, or just asking users to enter an input into your app, the Bing Autosuggest API helps narrow the search quickly by allowing your users to see suggestions for popular search terms. It can correct perceived mistakes, and returns detailed contextual suggestions according to other searches people have found useful.



The image is a collage consisting of two main parts. On the left is a vertical blue sidebar with four items, each featuring a white icon and text. On the right is a photograph of two people, a man and a woman, smiling and looking at a screen together. The blue sidebar contains the following text:

-  Get started for free at <http://microsoft.com/Cognitive>
-  Learn more on the Cortana Intelligence Suite [website](#) and Cognitive Services [website](#)
-  Schedule a workshop to identify areas in your business where analytics and intelligence can drive transformation
-  Talk with your Microsoft contact about licensing options and partners

Developer Resources

Pricing

<https://azure.microsoft.com/en-us/pricing/details/cognitive-services/>

Documentation

<https://www.microsoft.com/cognitive-services/en-us/computer-vision-api/documentation>

Client SDKs

<https://github.com/Microsoft/ProjectOxford-ClientSDK>

<https://github.com/felixriesenberg/project-oxford> (nodejs)

<https://github.com/southwood/project-oxford-python>

Example Code

<https://github.com/1sturtevant/happy-image-tester-django>

<https://github.com/1sturtevant/happy-image-tester-nodejs>

Community

<https://stackoverflow.com/questions/tagged/microsoft-cognitive>

<https://social.msdn.microsoft.com/forums/azure/en-US/home?forum=mlapi>

<https://cognitive.uservoice.com/>

Access to strong documentation, sample code and community resources is critical for developers to be able to understand and become users of Cognitive Services. Customize these links based on your own resources or use the ones listed here.

Q&A

Lab

Create a sample Knowledge Exploration Service

<https://www.microsoft.com/cognitive-services/en-us/kes/documentation/GettingStarted>

