Cognitive Services

Jani Nevalainen
Technical Evangelist @ Microsoft
www.j4ni.com
Twitter: @jn1974

Microsoft

Cognitive Services

Give your solutions a human side

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

Cognitive Services

Give your solutions a human side

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 | Knowledge Exploration | Recommendations

Search

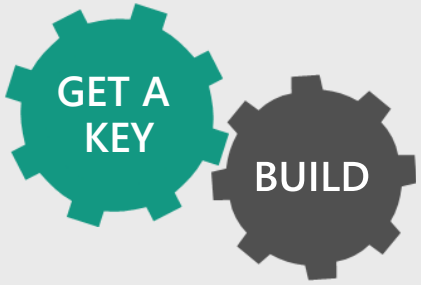
Bing Auto Suggest | Bing Image Search | Bing News Search | Bing Video Search | Bing Web Search

Why Microsoft Cognitive Services?

Easy

Roll your own with REST APIs

Simple to add: just a few lines of code required



Flexible

Integrate into the language and platform of your choice

Breadth of offerings helps you find the right API for your app

Tested

Built by experts in their field from Microsoft Research, Bing, and Azure Machine Learning

Quality documentation, sample code, and community support

Vision

Computer Vision API

Distill actionable information from images

Face API

Detect, identify, analyze, organize, and tag faces in photos

Emotion API

Personalize experiences with emotion recognition

Video API

Analyze, edit, and process videos within your app

Content Moderator

Automated image, text, and video moderation

Speech

Bing Speech API

Convert speech to text and back again, and understand its intent

Speaker Recognition API

Give your app the ability to know who's talking

Custom Recognition Intelligent Service

Fine-tune speech recognition for anyone, anywhere

Language

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

Knowledge

Academic Knowledge API

Explore relationships among academic papers, journals, and authors

Knowledge Exploration Service

Add interactive search over structured data to your project

Entity Linking Service

Contextually extend knowledge of people, locations, and events

Recommendations API

Provide personalized product recommendations for your customers

Search

Bing Web Search API

Connect powerful search to your apps

Bing Autosuggest API

Give your app intelligent autosuggest options for searches

Bing Image Search API

Bring advanced image and metadata search to your app

Bing Video Search API

Trending videos, detailed metadata, and rich results

Bing News Search API

Link your users to robust and timely news searches

Vision

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

Computer Vision | Emotion | Face | Video | Content Moderator

Computer Vision APIs

Analyze an image
Understand content within an image

OCR
Detect and recognize words within an image

Generate thumbnail
Scale and crop images, while retaining key content

Recognize Celebrities
Thanks to domain specific models, ability to recognize 200K celebrities from business, politics, sports and entertainment around the World

Analyze image

Type of image

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

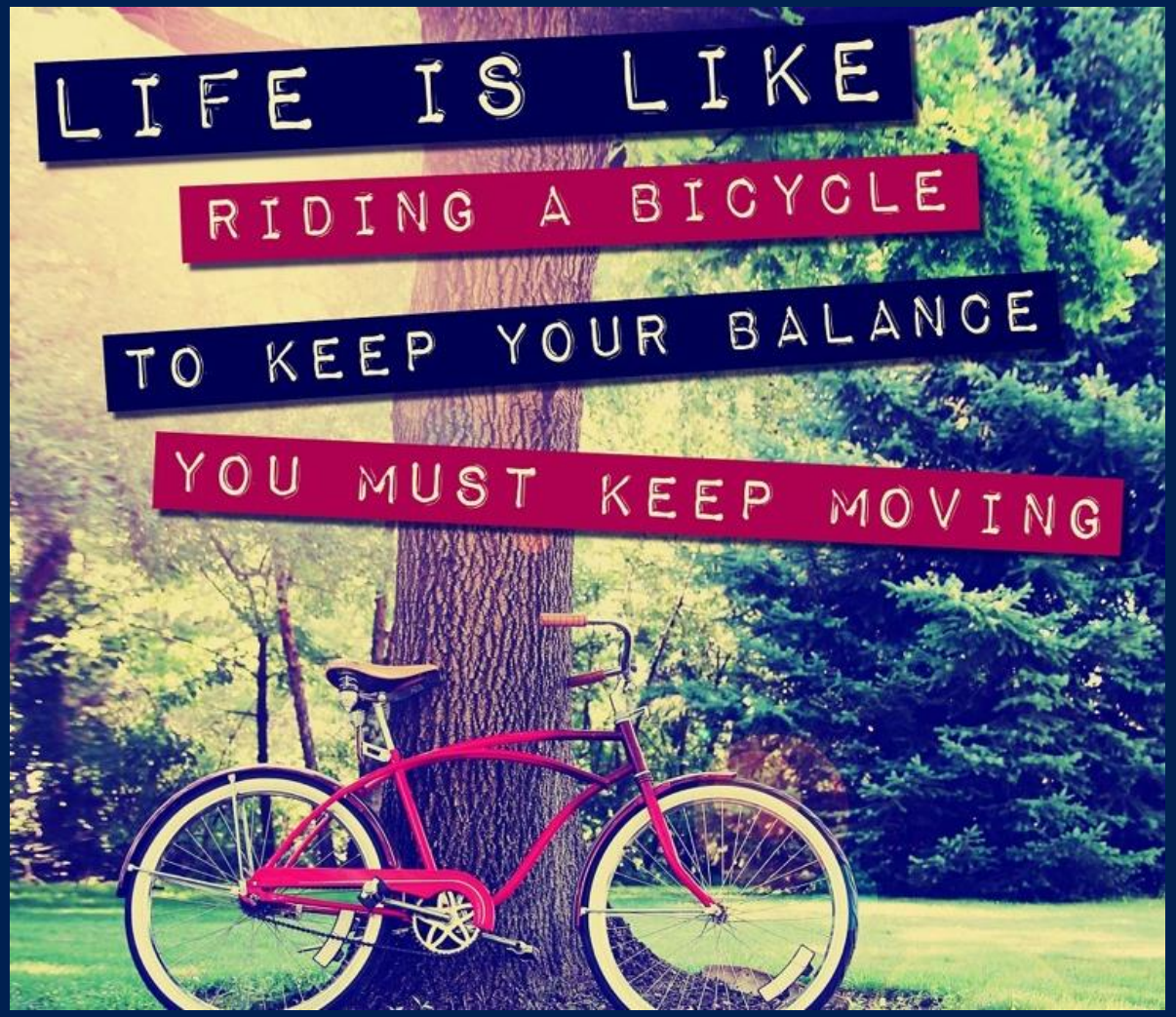
Content of image

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

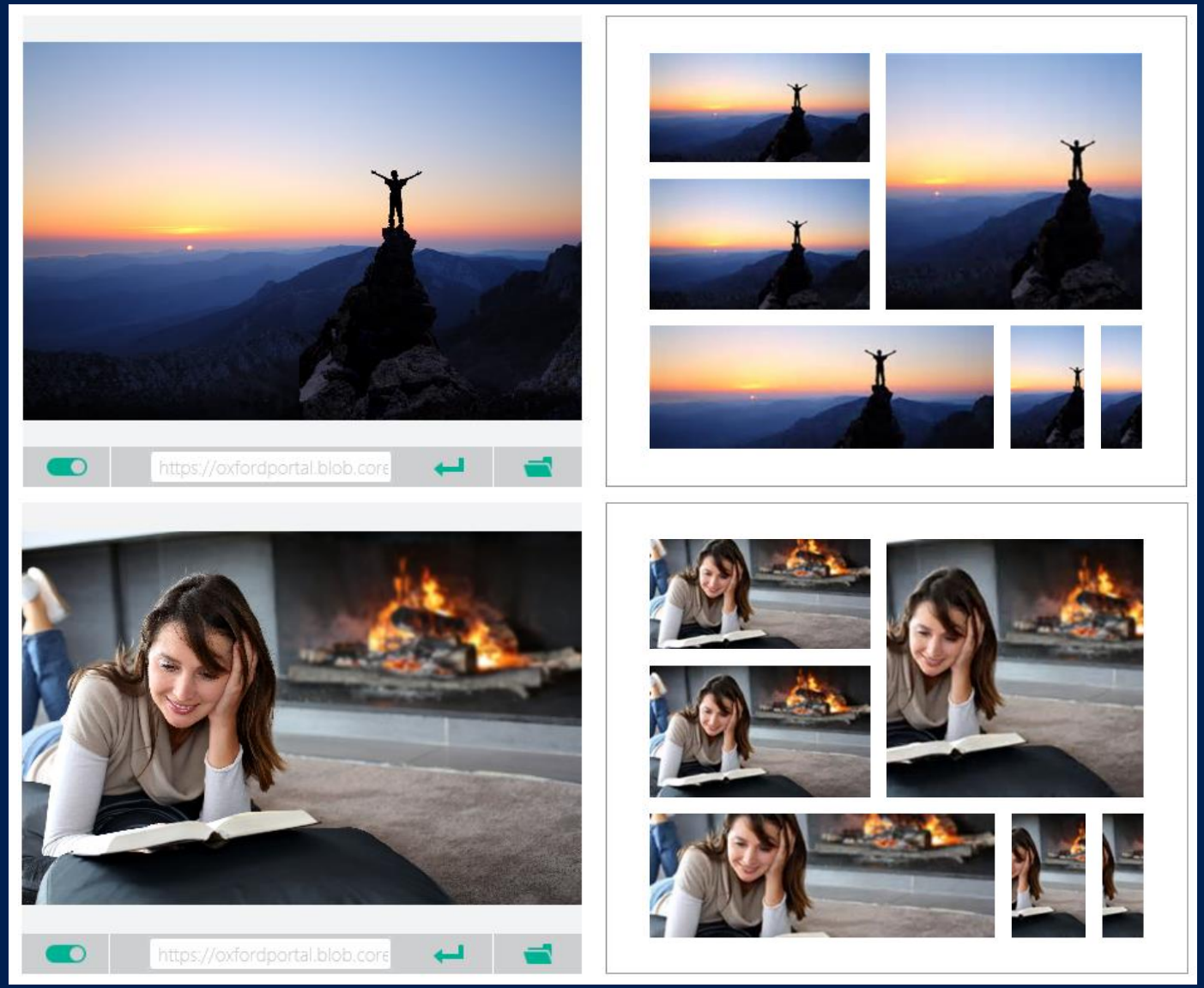
- Scanned documents
- Photos with text
- Fine grained location information

JSON:


```
{
  "language": "en",
  "orientation": "Up",
  "regions": [
    {
      "boundingBox": "41,77,918,440",
      "lines": [
        {
          "boundingBox": "41,77,723,89",
          "words": [
            {
              "boundingBox": "41,102,225,64",
              "text": "LIFE"
            },
            {
              "boundingBox": "356,89,94,62",
              "text": "IS"
            },
            {
              "boundingBox": "539,77,225,64",
              "text": "LIKE"
            }
. . .
```

Smart thumbnail

Smart cropping **off**

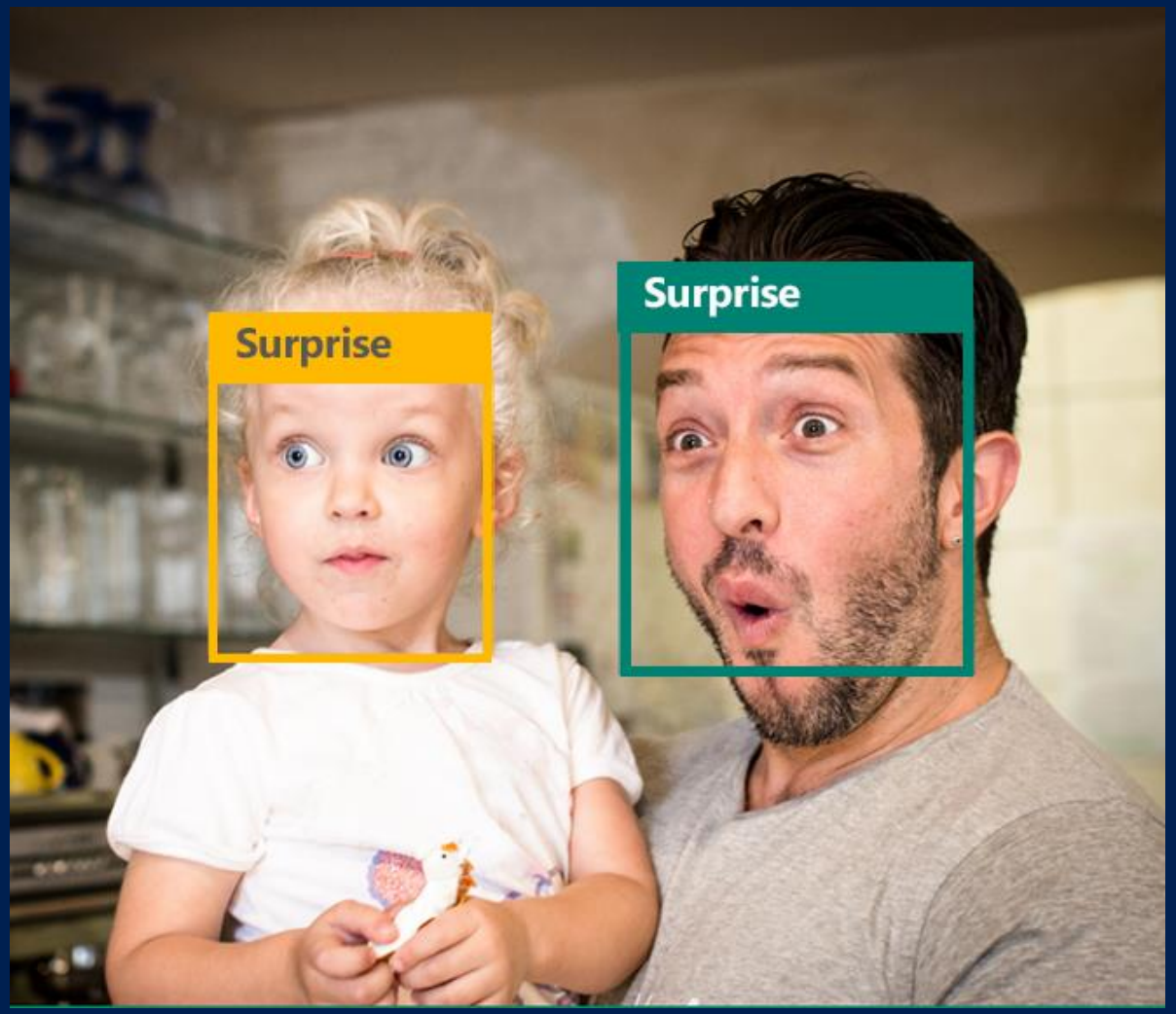


Emotion APIs

Recognize emotions

Detect emotions based on facial expressions in photos and videos

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}
```

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

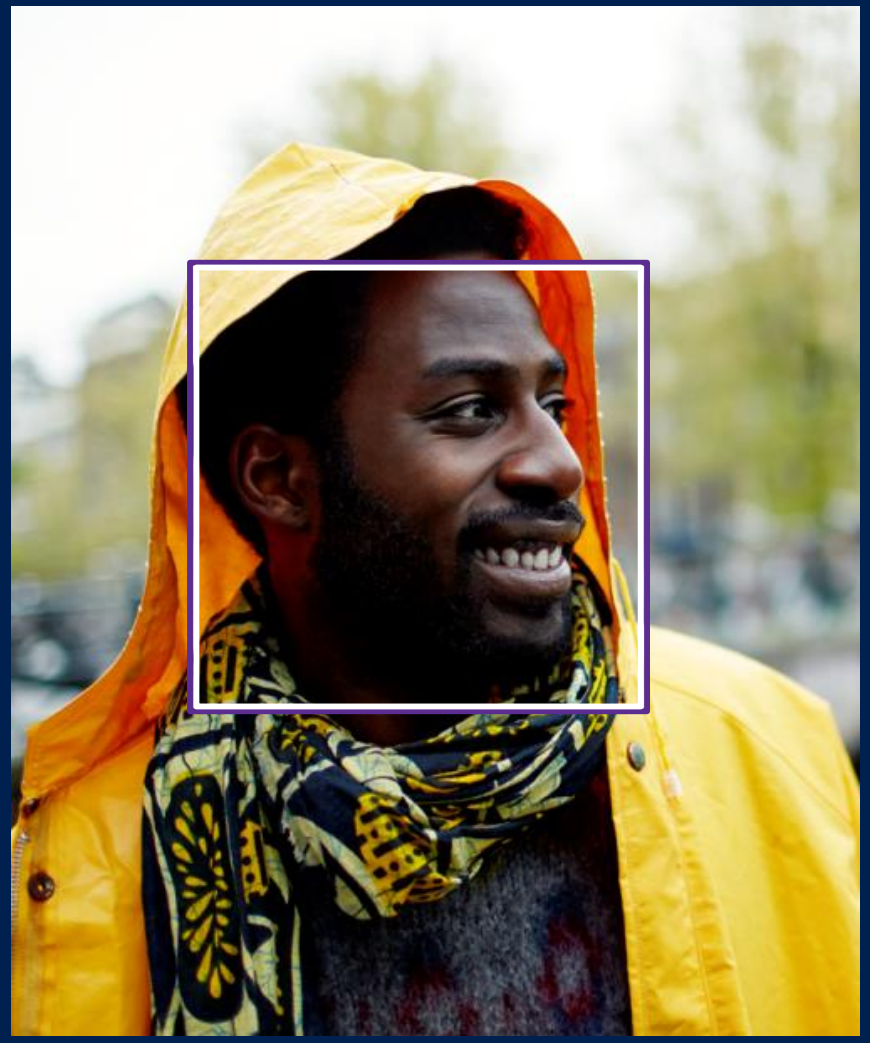
Face grouping

Organize many faces into groups

Face identification

Search which person a face belongs to

Face APIs



Detection

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

Feature attributes

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

Grouping

Identification

Jasper Williams

Video APIs

Stabilization

Smooth and stabilize shaky video

Face detection and tracking

Detect and track faces in videos

Motion detection

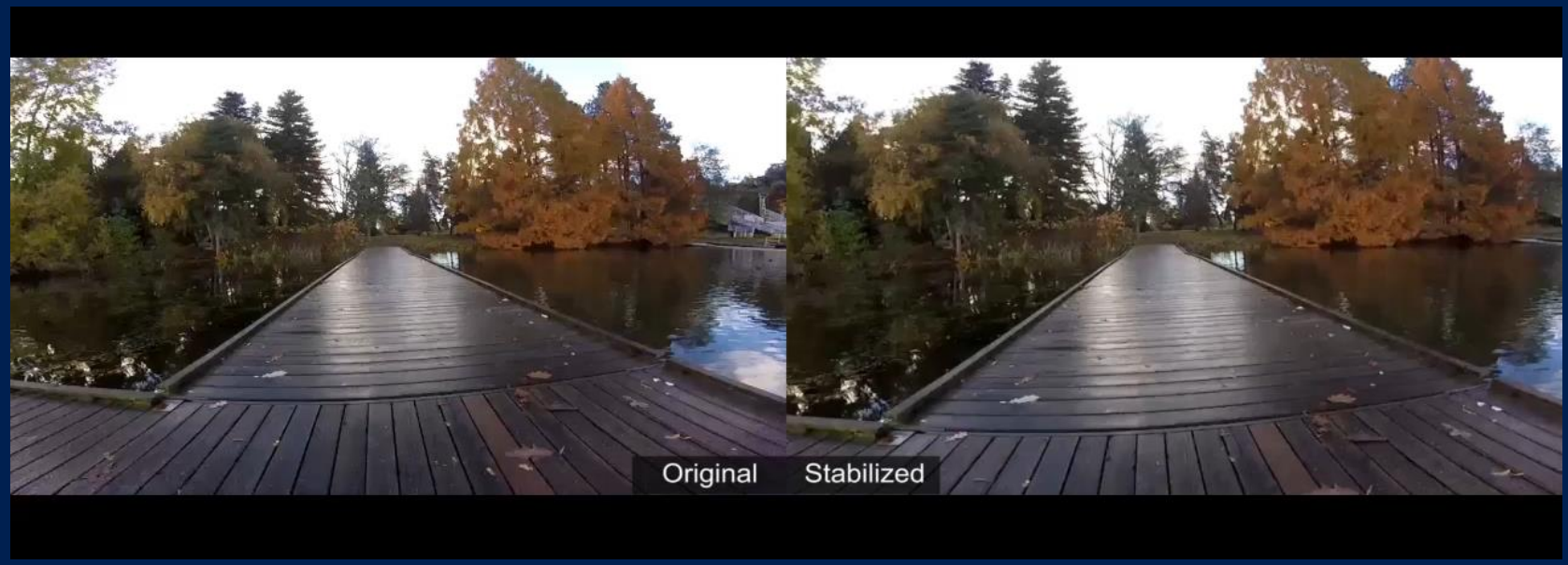
Detect when motion occurs

Stabilization

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)



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

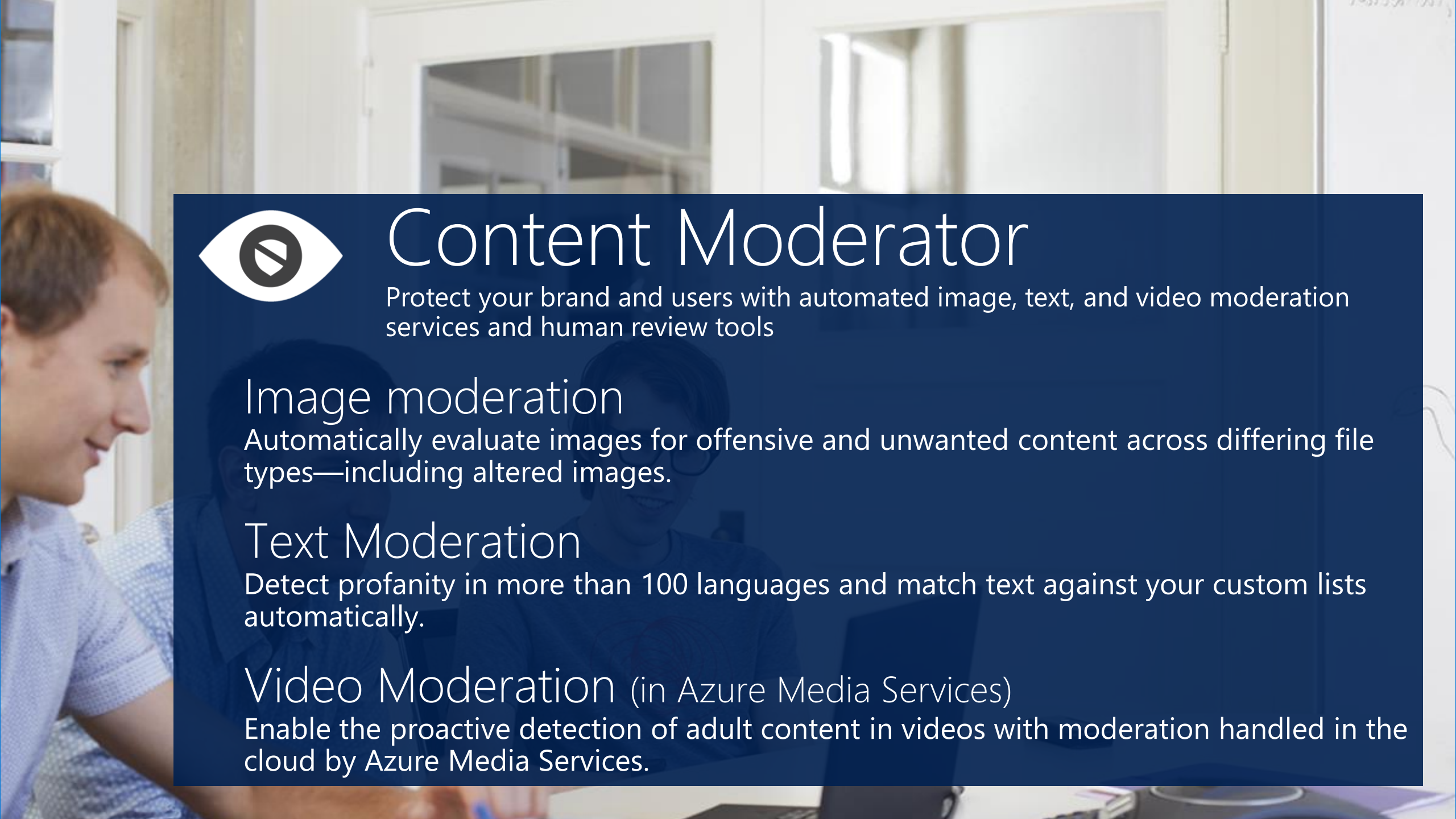
Time (sec)	Face ID	x, y	Width, height
0	0	0.59, 0.23	0.09, 0.16
0	1	0.38, 0.15	0.07, 0.12
1	0	0.54, 0.25	0.09, 0.15
1	1	0.23, 0.18	0.07, 0.12

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



Content Moderator

Protect your brand and users with automated image, text, and video moderation services and human review tools

Image moderation

Automatically evaluate images for offensive and unwanted content across differing file types—including altered images.

Text Moderation

Detect profanity in more than 100 languages and match text against your custom lists automatically.

Video Moderation (in Azure Media Services)

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

Speech

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

Custom Recognition | Speaker Recognition | Speech

Bing Speech API

Voice recognition (speech to text)

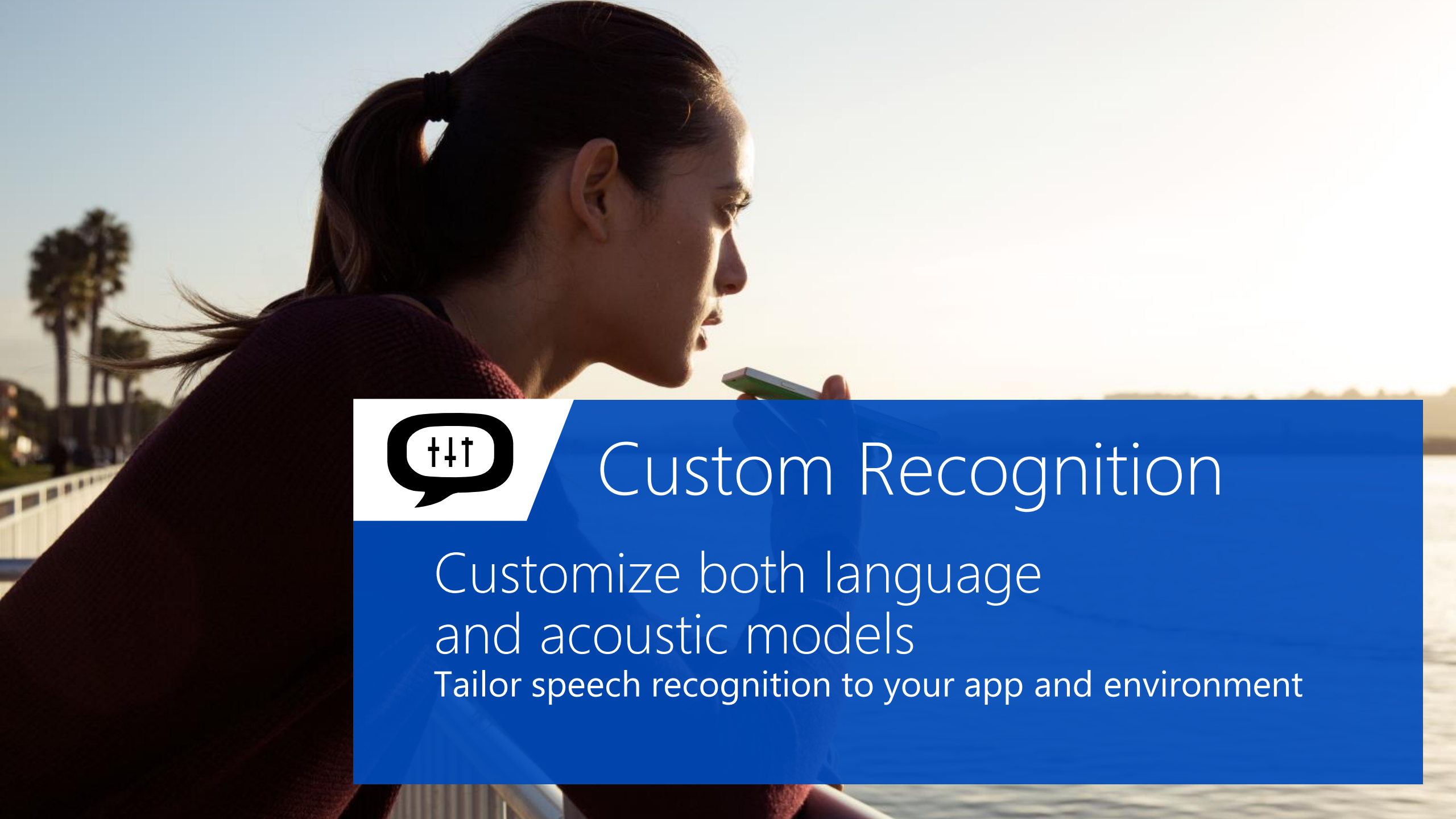
Converts spoken audio to text

Voice output (text to speech)

Synthesize audio from text

Speech intent recognition

Convert spoken audio to intent



Custom Recognition

Customize both language and acoustic models

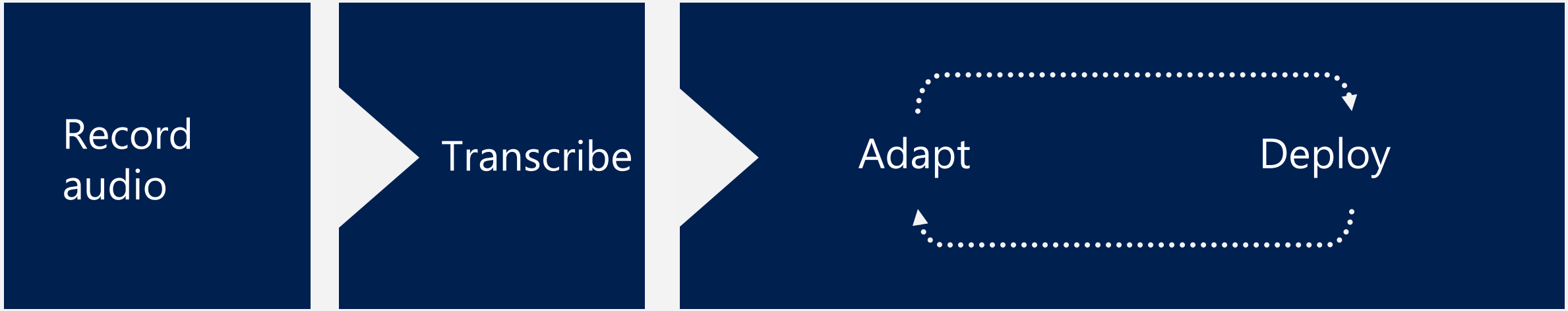
Tailor speech recognition to your app and environment

Custom Recognition Intelligent Service

Create custom language models for the vocabulary of the application

Adapt acoustic models to better match the expected environment of the application's users

Deploy to a custom endpoint and access from any device



Speaker recognition API

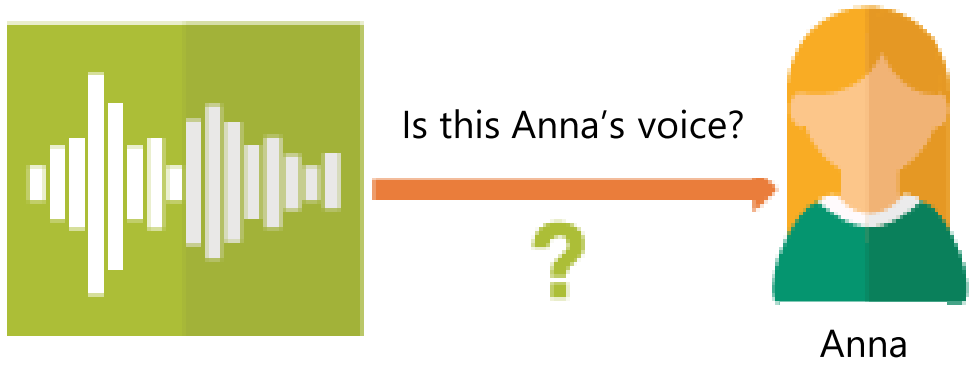
Speaker verification

Check if two voices are the same

Speaker identification

Identify who is speaking

Speaker recognition APIs



Enrollment

Create a unique voiceprint for a profile

Recognition

After enrolling one or more voices, identify who is speaking from an audio clip

Verification

Confirm if a voice belongs to a previously enrolled profile

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

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

Bing spell check APIs

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" }
    ] }
```

Language Understanding

Understand what your users are saying

Use pre-built Bing and Cortana models or create your own

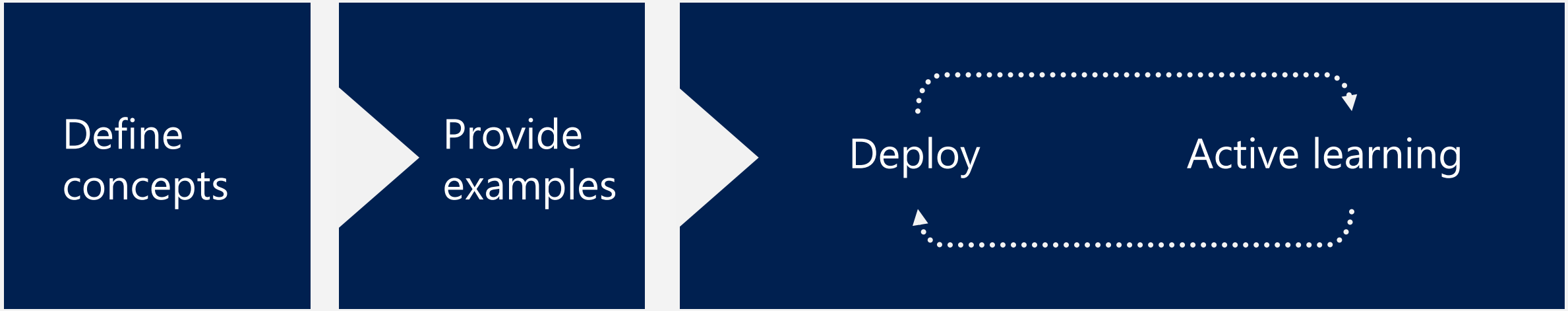
Language understanding intelligent service

Reduce labeling effort with interactive featuring

Use visualizations to gauge performance and improvements

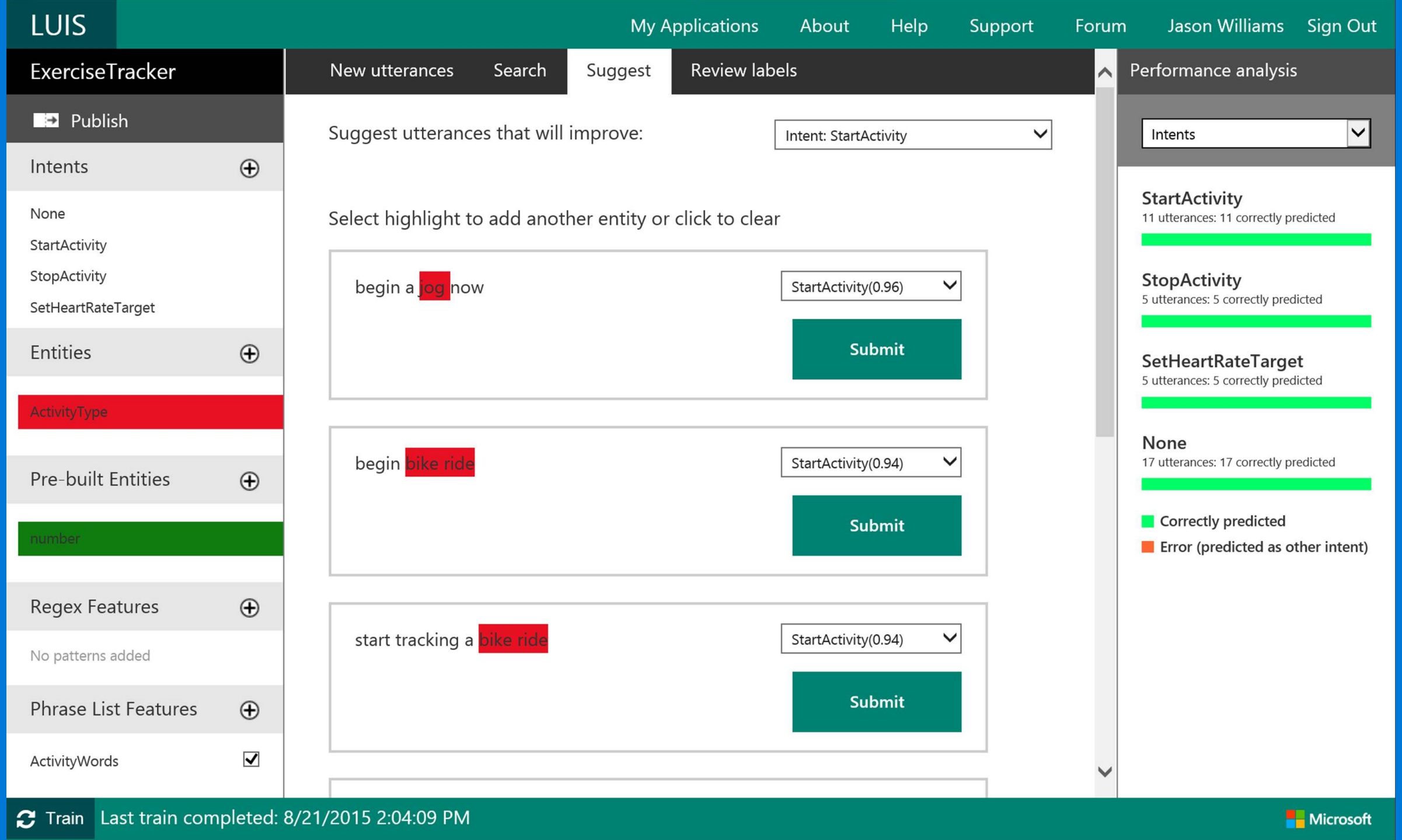
Leverage speech recognition with seamless integration

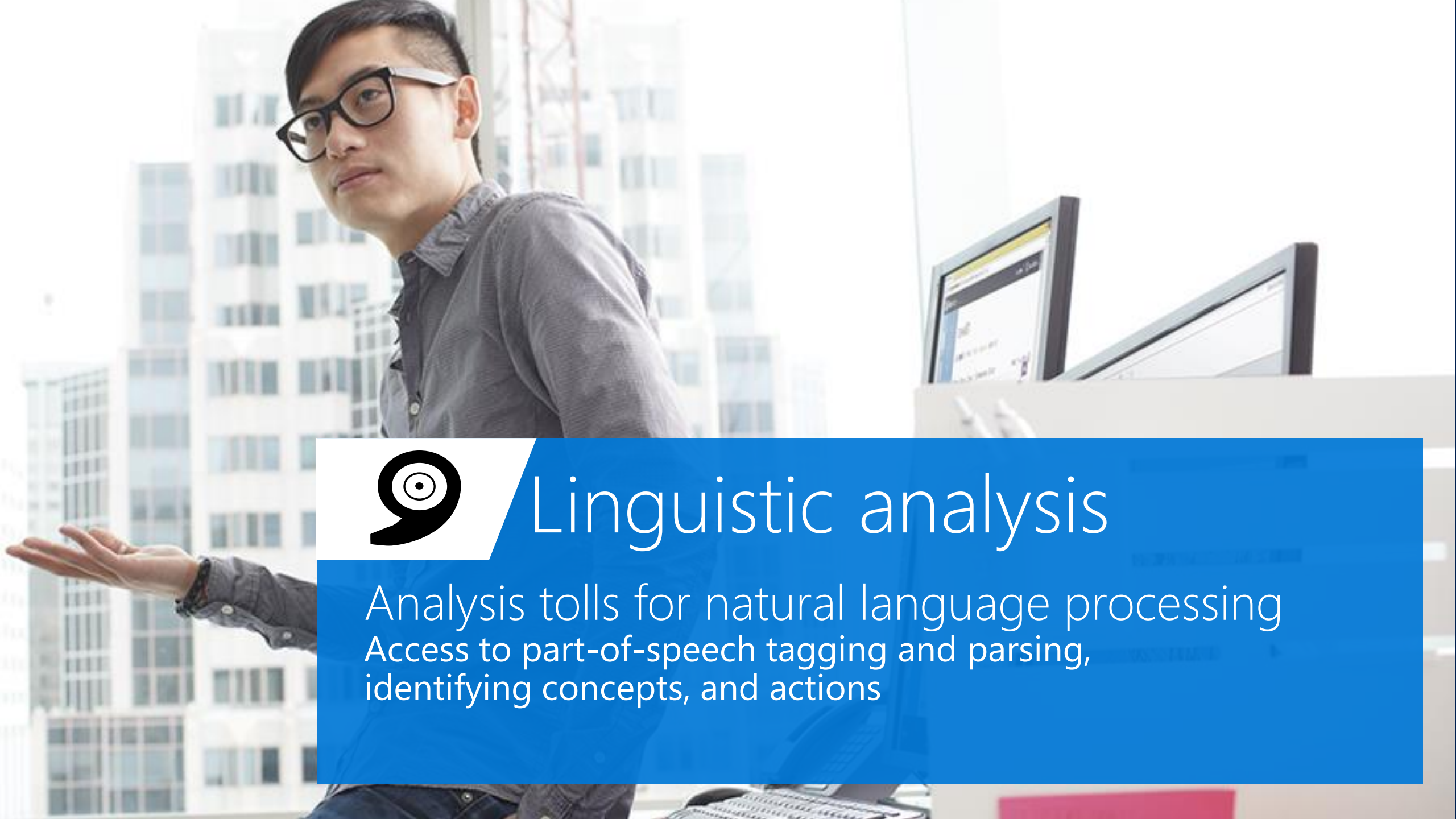
Deploy using just a few examples with active learning



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
    }
  ]
}
```


LUIS
My Applications
About
Help
Support
Forum
Jason Williams
Sign Out
ExerciseTracker
New utterances
Search
Suggest
Review labels
Performance analysis
Publish
Intents
None
StartActivity
StopActivity
SetHeartRateTarget
Entities
ActivityType
Pre-built Entities
number
Regex Features
No patterns added
Phrase List Features
ActivityWords
Suggest utterances that will improve:
Intent: StartActivity
Select highlight to add another entity or click to clear
begin a jog now
StartActivity(0.96)
Submit
begin bike ride
StartActivity(0.94)
Submit
start tracking a bike ride
StartActivity(0.94)
Submit
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



Linguistic analysis

Analysis tolls for natural language processing

Access to part-of-speech tagging and parsing,
identifying concepts, and actions

Linguistic Analysis

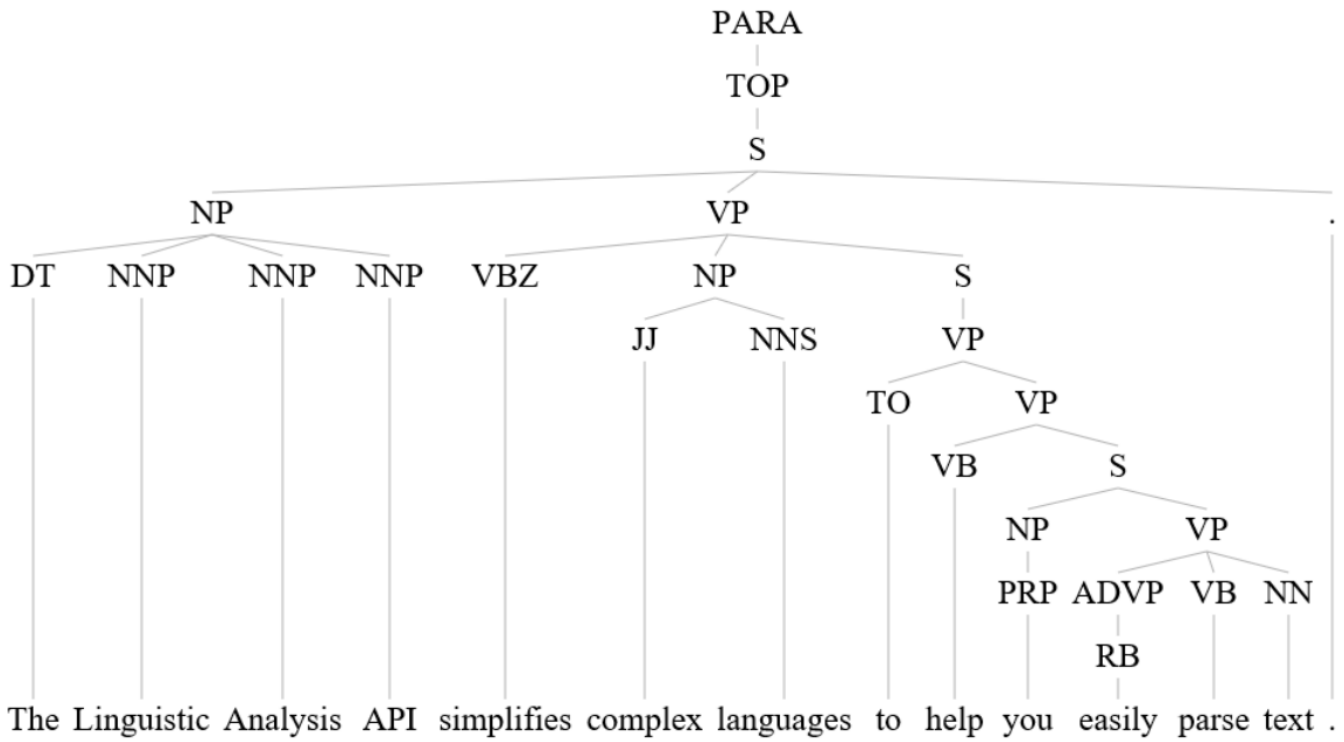
Enter a sentence

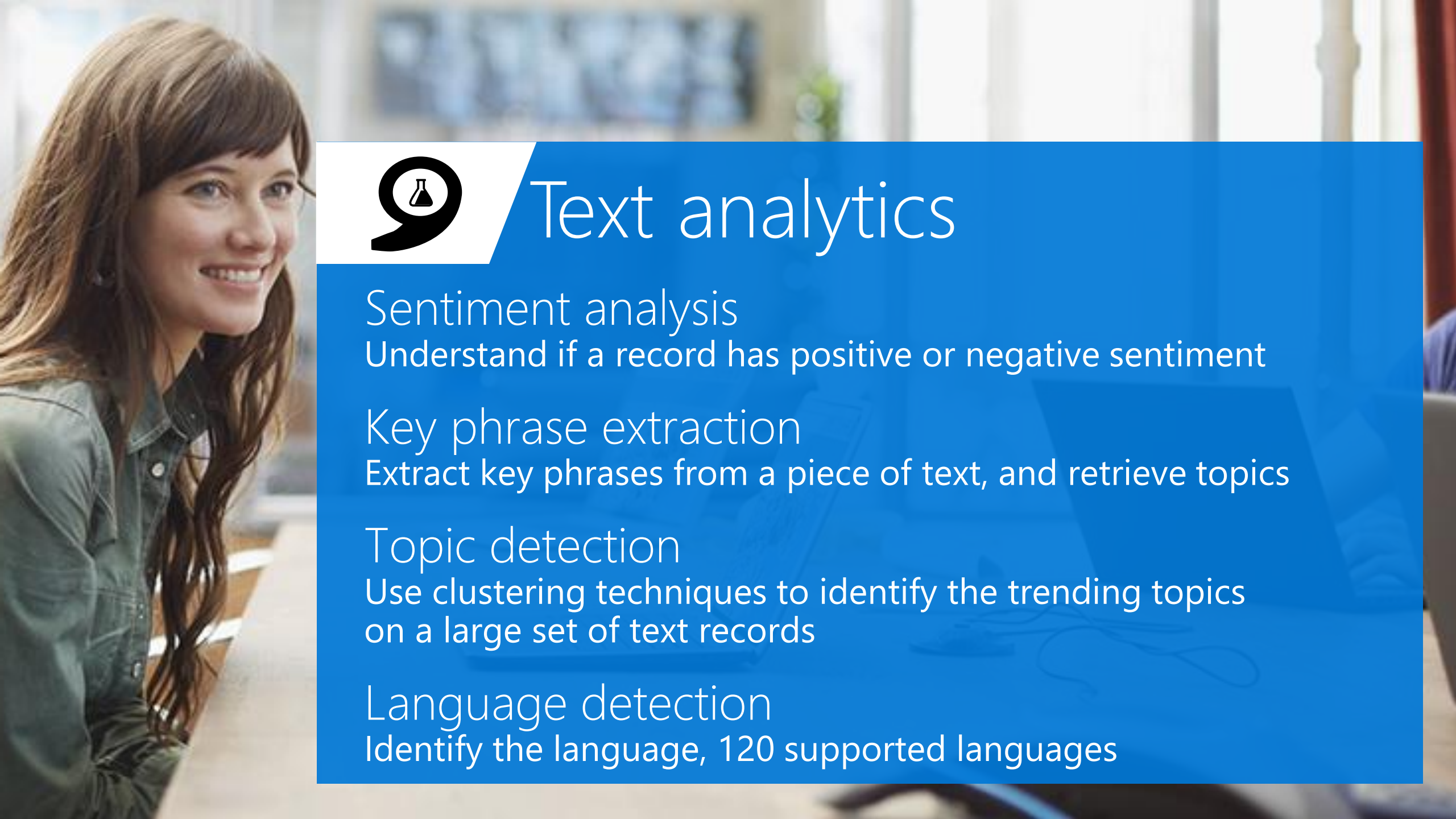
The Linguistic Analysis API simplifies complex languages to help you easily parse text.

POS tags

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

Constituency tree





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

Microsoft Translator

Translator Text API

Automatically detect language & easily power translation to and from 60 supported text languages

Translator Speech API

Easily translate real-time speech conversations in 9 support languages

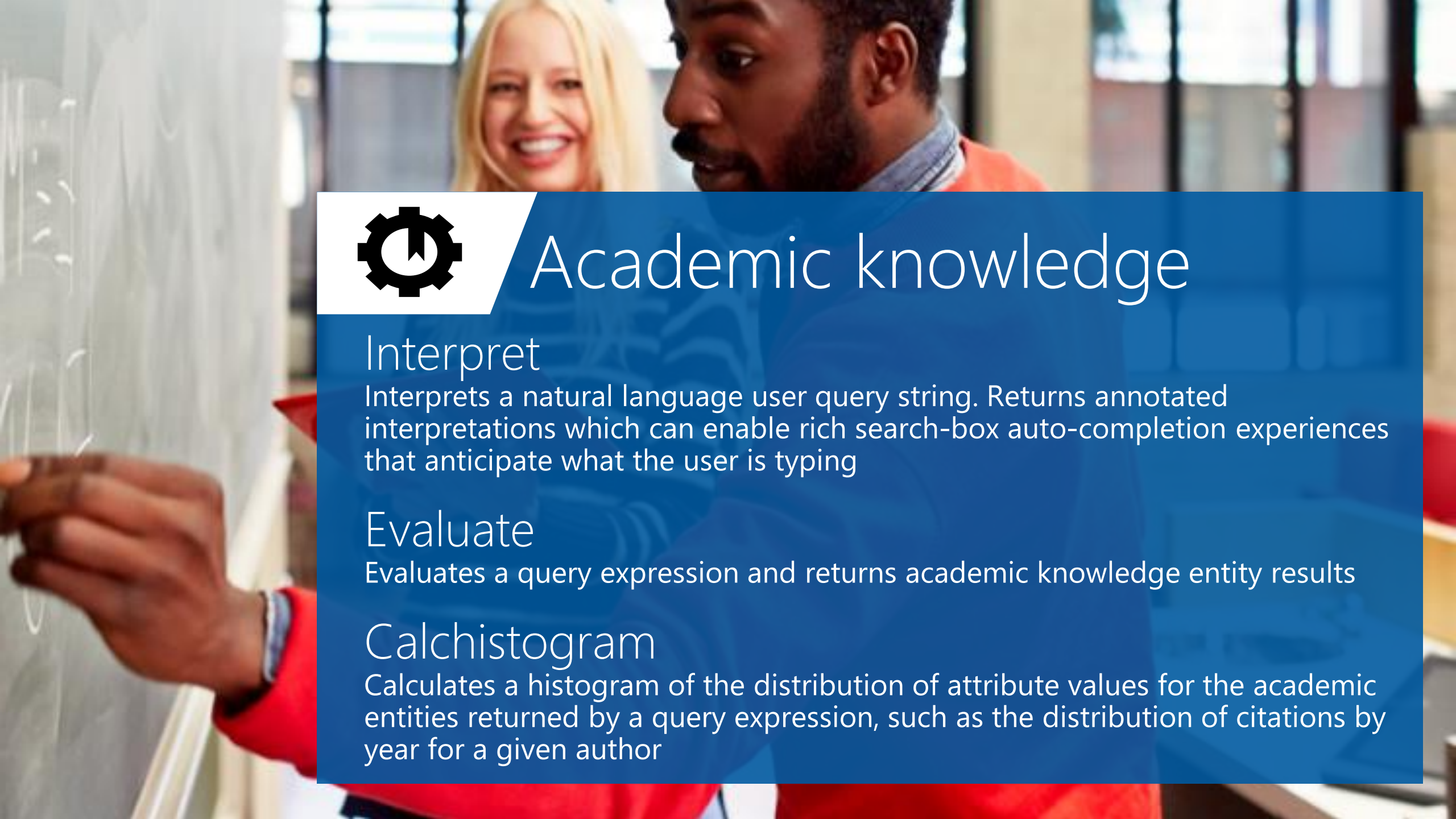
Knowledge

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

Academic Knowledge

Entity Linking | Knowledge Exploration

Recommendations



Academic knowledge

Interpret

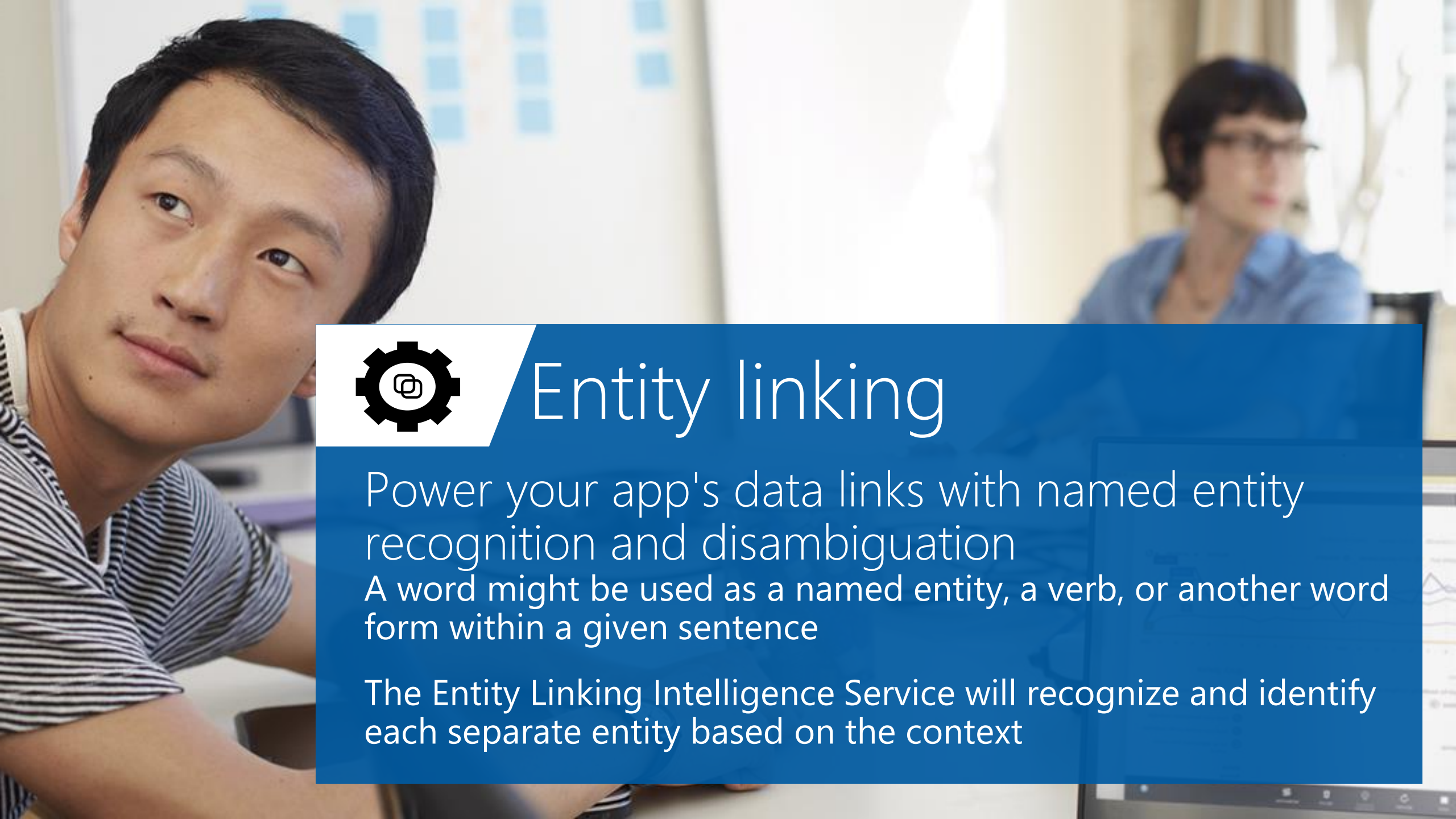
Interprets a natural language user query string. Returns annotated interpretations which can enable rich search-box auto-completion experiences that anticipate what the user is typing

Evaluate

Evaluates a query expression and returns academic knowledge entity results

Calchistogram

Calculates a histogram of the distribution of attribute values for the academic entities returned by a query expression, such as the distribution of citations by year for a given author



Entity linking

Power your app's data links with named entity recognition and disambiguation

A word might be used as a named entity, a verb, or another word form within a given sentence

The Entity Linking Intelligence Service will recognize and identify each separate entity based on the context

Knowledge exploration

Enable interactive search experiences over structured data via natural language inputs

Attribute histograms

To enable rich visualization and interactive faceted experience

Structured query evaluation

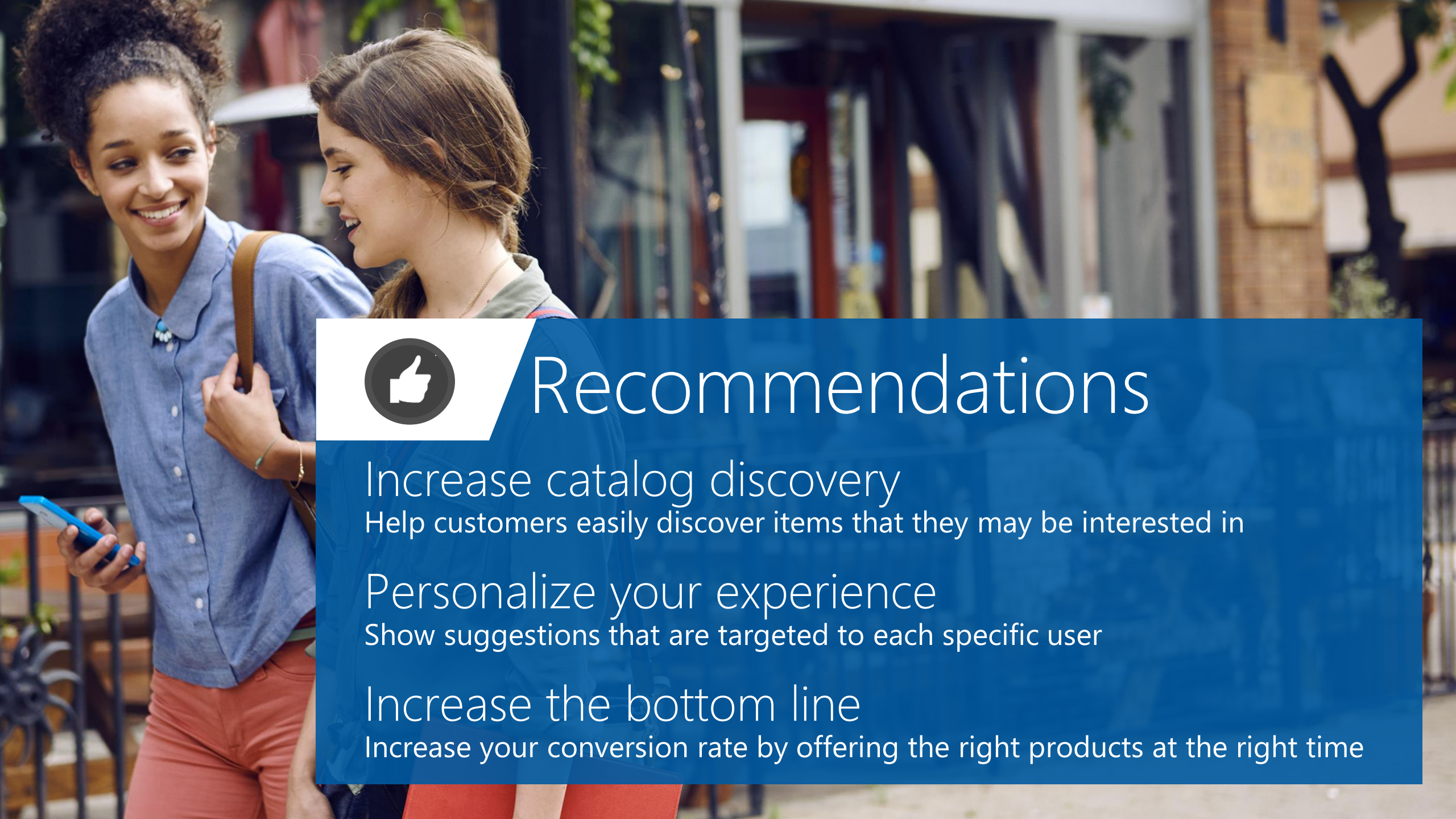
To efficiently retrieve detailed information about matching objects

Query auto-completion

To reduce user effort and help with discovery of rich capabilities

Natural language understanding

To interpret natural language queries into structured query expressions



Recommendations

Increase catalog discovery

Help customers easily discover items that they may be interested in

Personalize your experience

Show suggestions that are targeted to each specific user

Increase the bottom line

Increase your conversion rate by offering the right products at the right time

Recommendations API

*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.

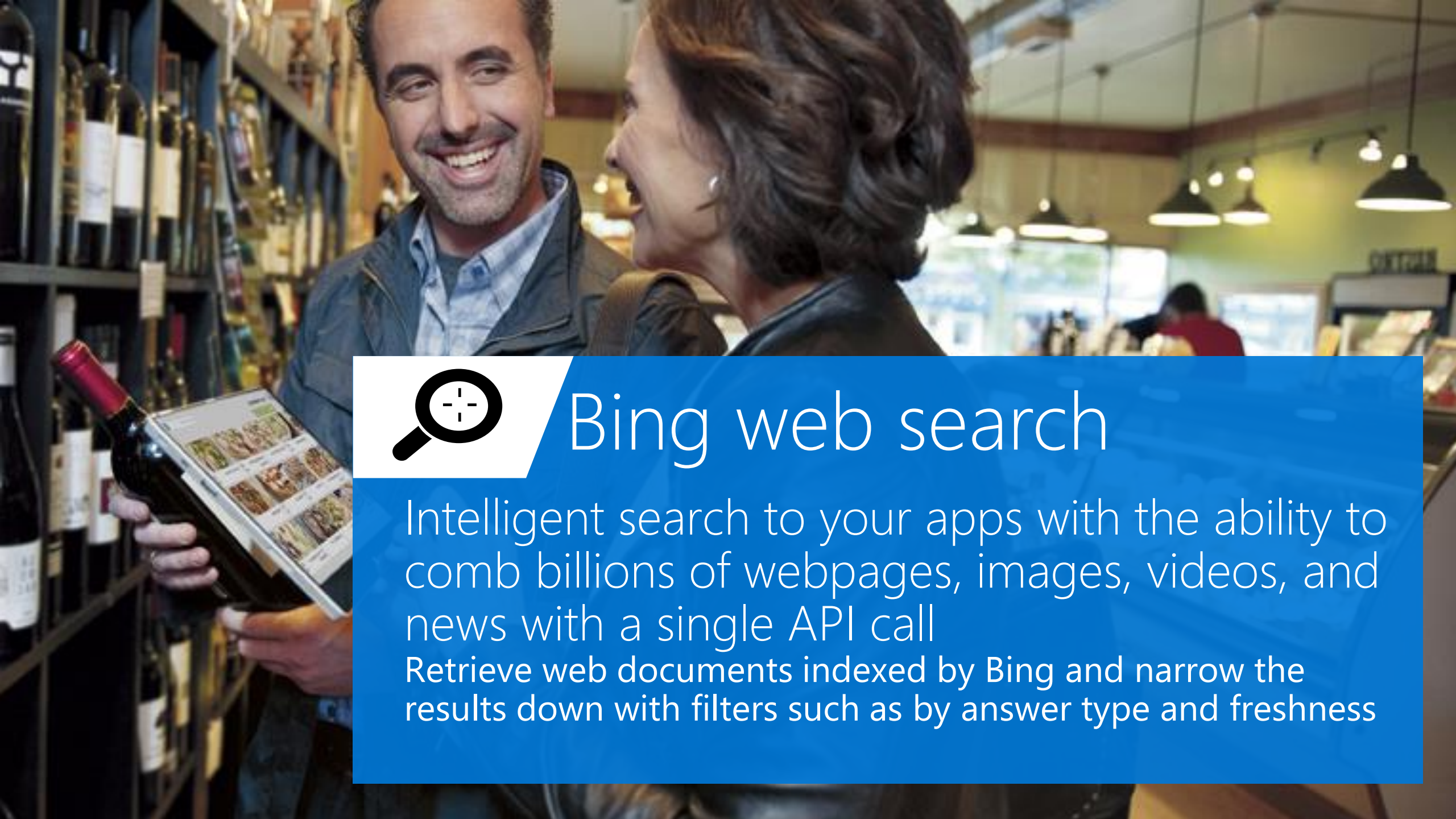
John Keane
CTO, Allrecipes.com

allrecipes.com®

Search

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

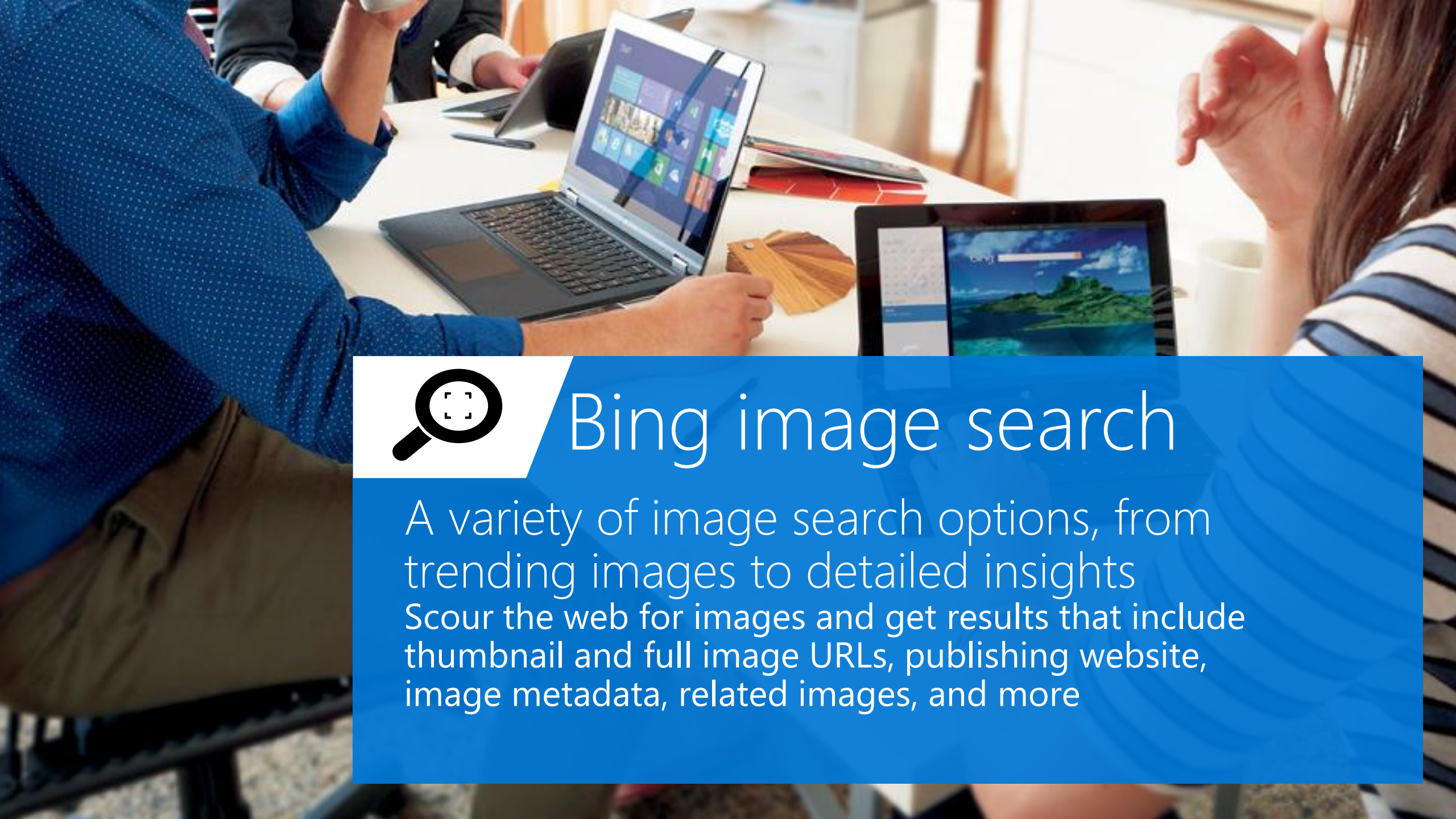
Bing Web Search | Bing Image Search
Bing News Search | Bing Video Search
Bing Auto Suggest



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



Bing image search

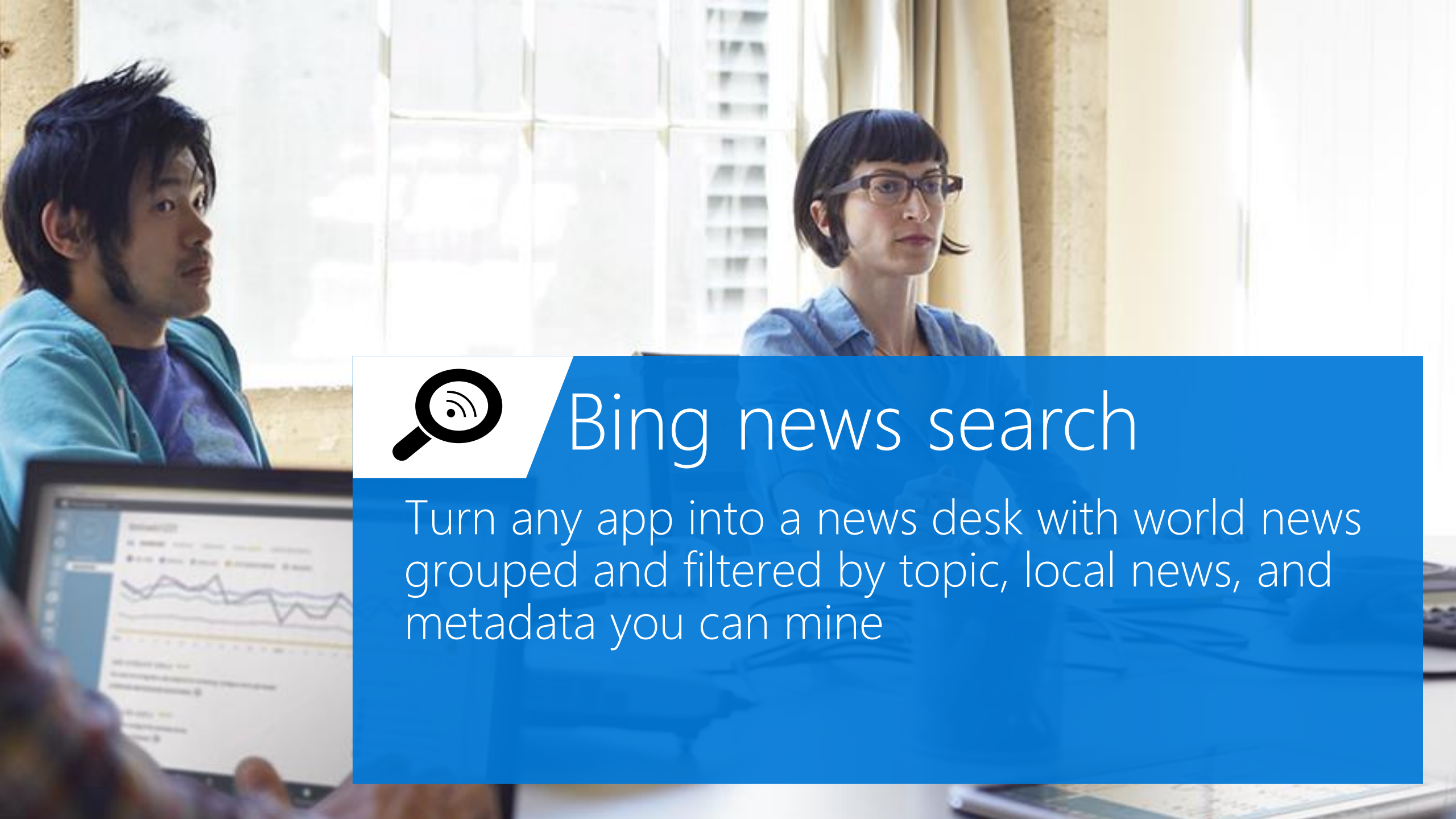
A variety of image search options, from trending images to detailed insights

Scour the web for images and get results that include thumbnail and full image URLs, publishing website, image metadata, related images, and more

Bing video search

A variety of advanced video search features, including trending videos, price, and other useful metadata

Find videos from across the web and get responses that provide useful metadata including creator, encoding format, video size and quality, and source view count



Bing news search

Turn any app into a news desk with world news grouped and filtered by topic, local news, and metadata you can mine

Bing auto suggest

Query completion suggestions capabilities, so users can type less and get to what they want faster

Microsoft

Sample on Cognitive Services and Xamarin can be found at:

www.j4ni.com

© 2016 Microsoft Corporation. All rights reserved.