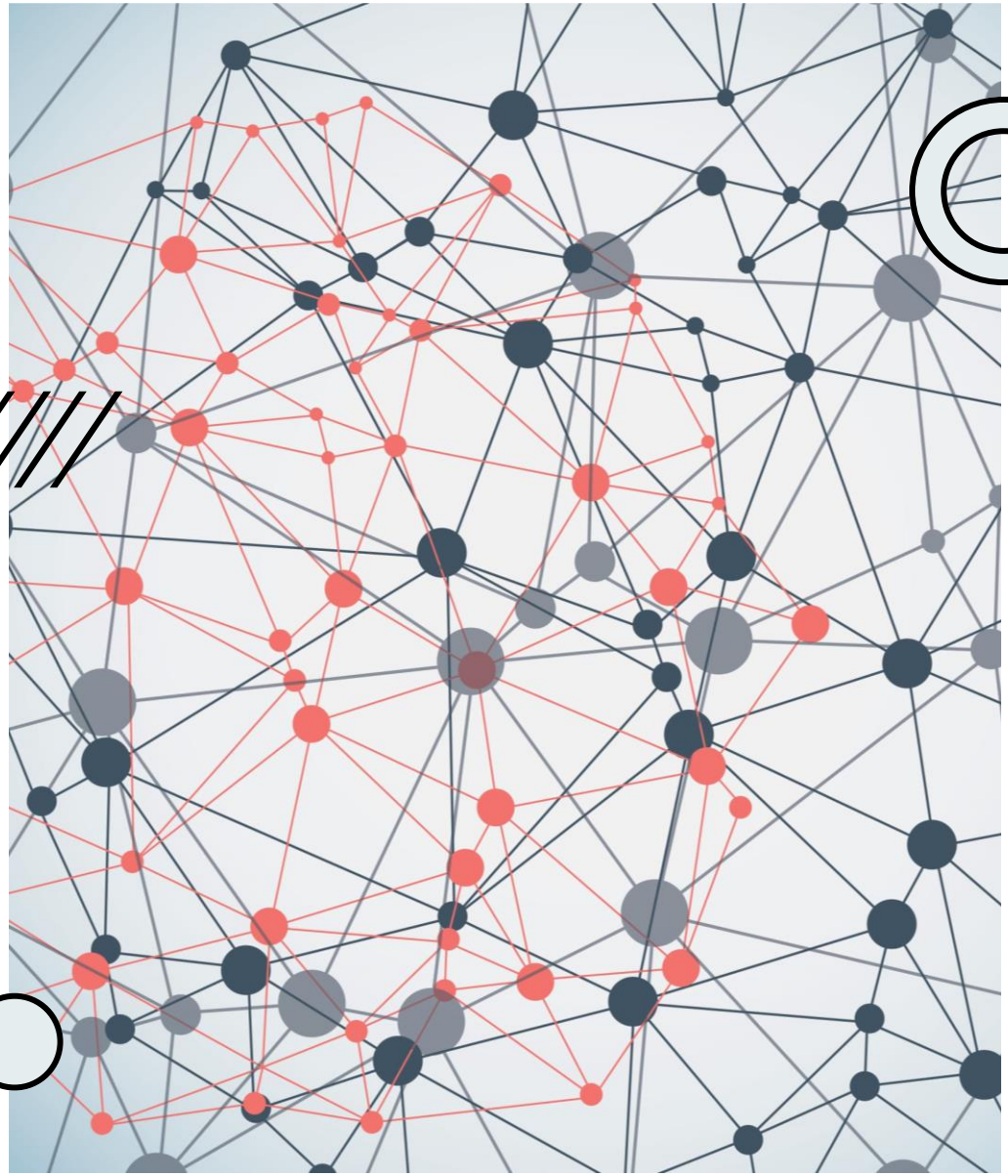
COGNITIVE SERVICES

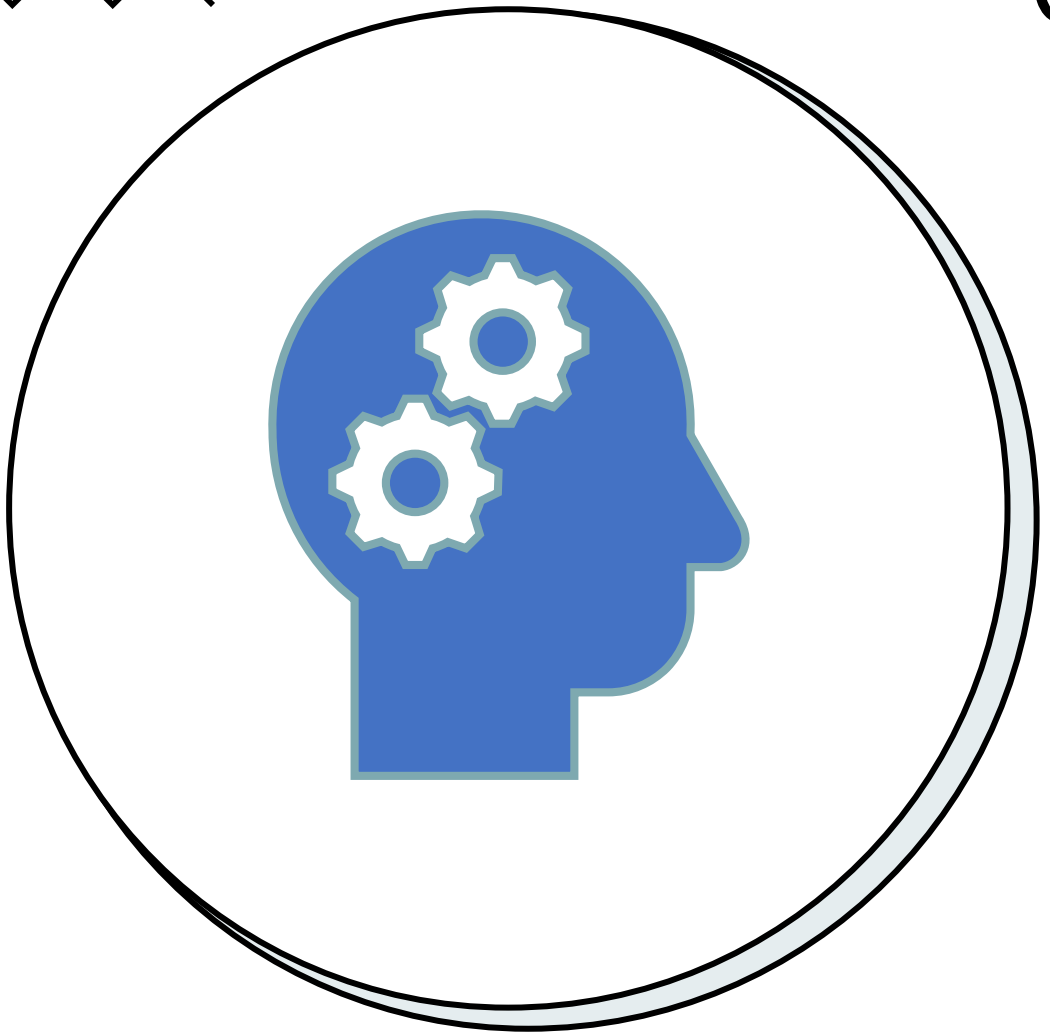
COMPUTER VISION

CONF.DR.CRISTIAN KEVORCHIAN

UNIVERSITATEA BUCUREȘTI



Cognitive Computing



Cognitive Computing (CC) is a new paradigm of intelligent computing platforms and approaches for creating cognitive and autonomous systems that mirror the brain's functional principles.

We must learn to see computers if we want them to think.

Fei Fei Li, Professor at Stanford and Chief Scientist AI/ML at Google Cloud

Problem

- In most circumstances, we have a family of

objects that need to be categorized, tagged, and classed into different subgroups of grou ps in the computer vision sector.

Human operators have developed very good skills in these tasks, but the machines are still in an early stage.

How do we give a computer a sample of things and ask it to categorize them into two groups? It may be a picture or a drawing, a corgi or a stray dog - the possibilities are endless..

Use Cases

Detecting emotions in retail

In order to locate missing children, facial recognition is used.

Sentiment analysis applied in sales processes

Facial detection in calculating the ratio of women/men in a nightclub

Language processing to allow voice assistants (bot) to understand natural language

Object recognition to help those with severe visual impairments read more easily

Vision

Computer Vision API

Processing of information in images

Face API

Detection, identification, analysis, organization and tagging of faces in photos

Emotion API

Personalizing experience with emotion recognition

Video API

Analysis, editing, and video processing within applications

Azure Cognitive Services Custom Vision

The Custom Vision service uses a machine learning algorithm for analyzing images.

Whether or not to show the explicit attributes is loaded into groups of photos.

When the photographs are uploaded to the project, they are labeled. The algorithm then uses this data to train and calculate the evaluation's correctness by testing on the same photos.

When the photographs are uploaded to the project, they are labeled. The algorithm then uses this data to train and calculate the evaluation's correctness by testing on the same photos.

Classification and detection of objects

- **Custom Vision** functionality can be divided into two features:
  - **Image classification** applies one or more labels to an image.
  - **Object detection** is similar, but unlike what classification also returns the coordinates in the image where the applied labels can be found.

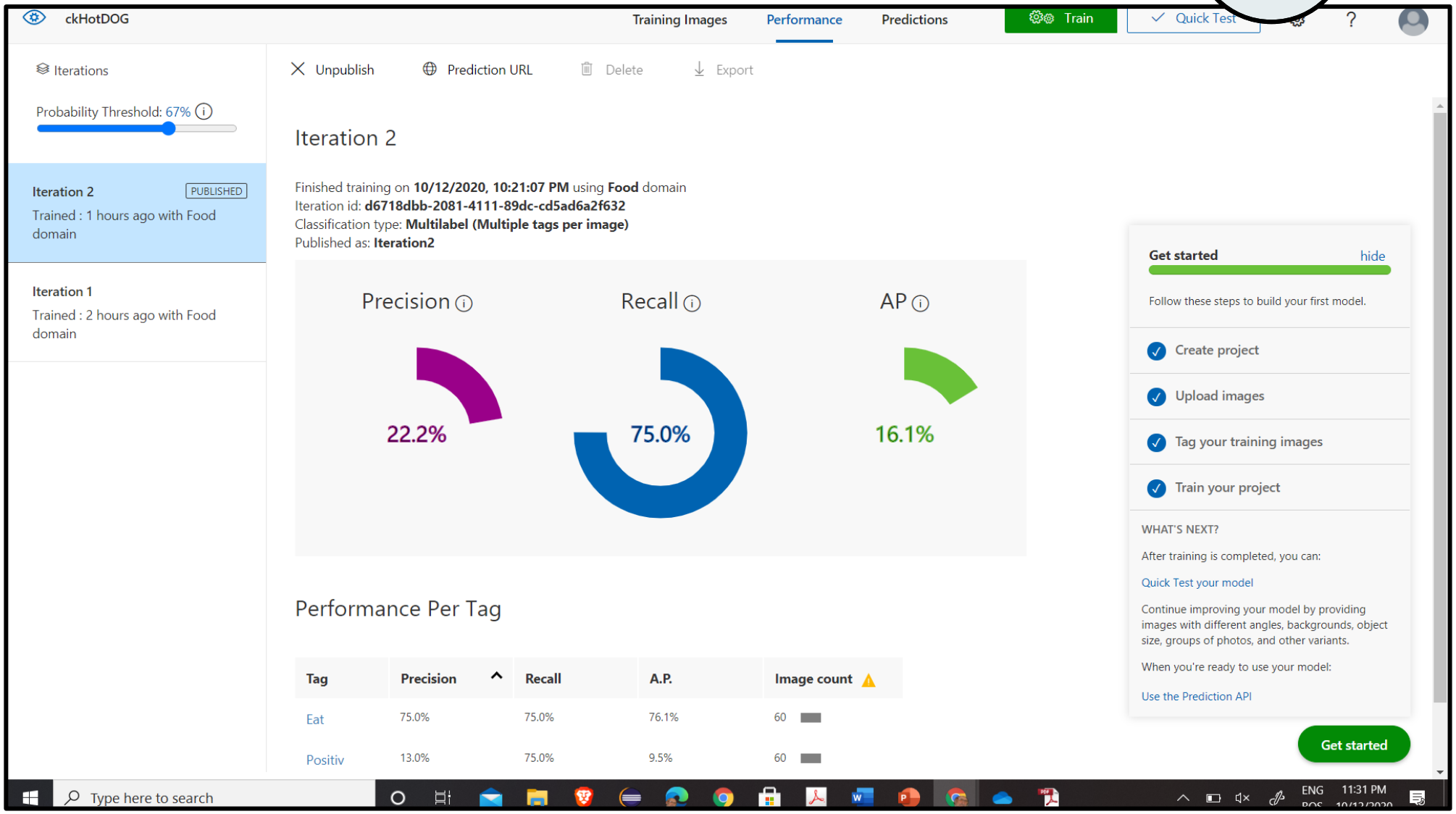
Precision and Recall

- Precision is the degree of exactness of the model in identifying only relevant objects. It is the ration of TPs over all detections made by the model.

$$P = \frac{TP}{TP + TN}$$

- Recall measures the ability of the model to detect all ground truths— proposition of TPs among all ground truths.

$$R = \frac{TP}{TP + FP}$$



https://www.customvision.ai/projects/e5834bd6-47df-4d39-b37f-530d7b2c9544#/performance

VISION

sursa: Microsoft.com

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

Computer Vision | Content Moderator | Emotion | Face | Video Indexer | Custom Vision Service

Computer Vision API

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}}]

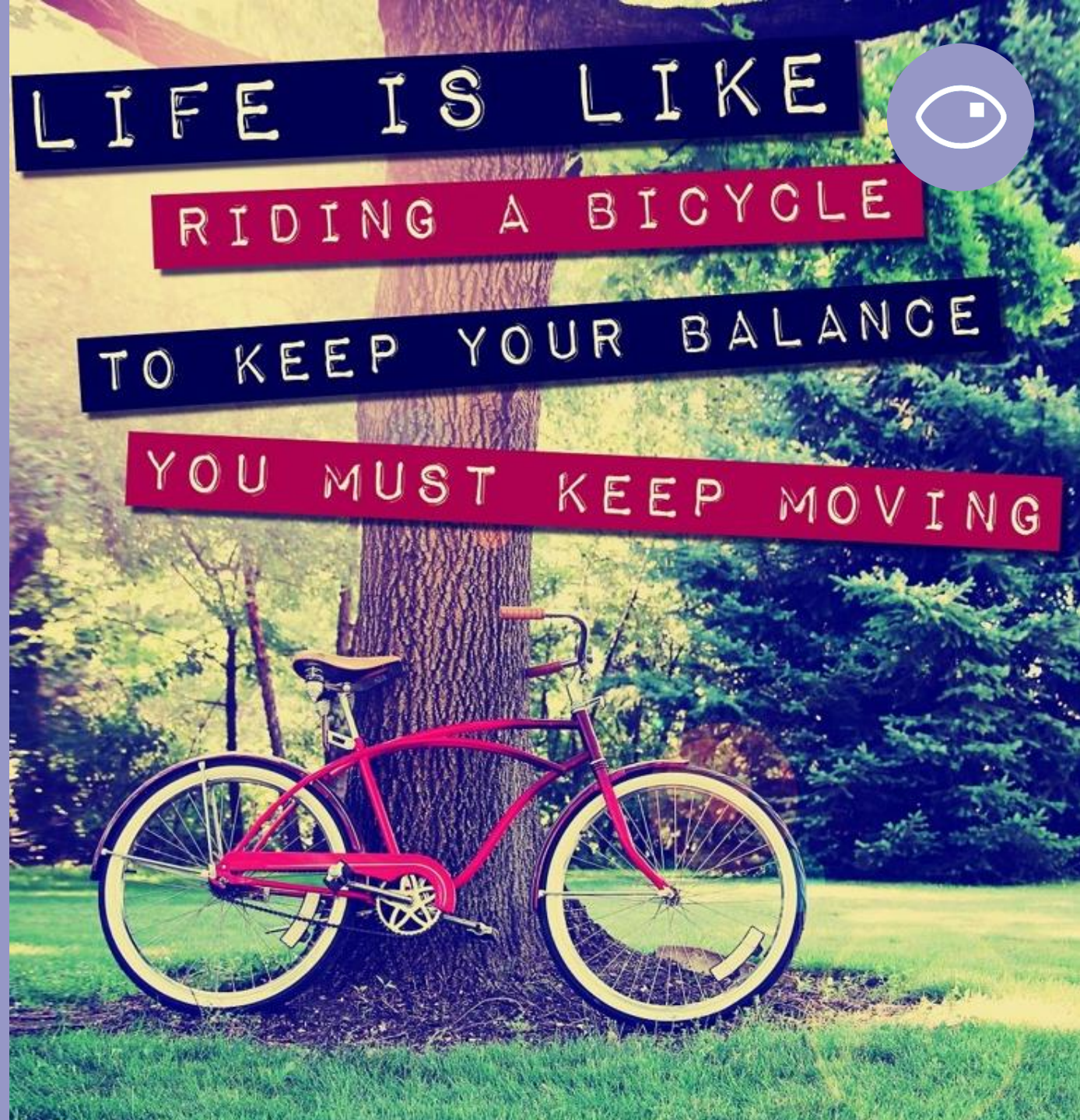
Image colors

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

OCR

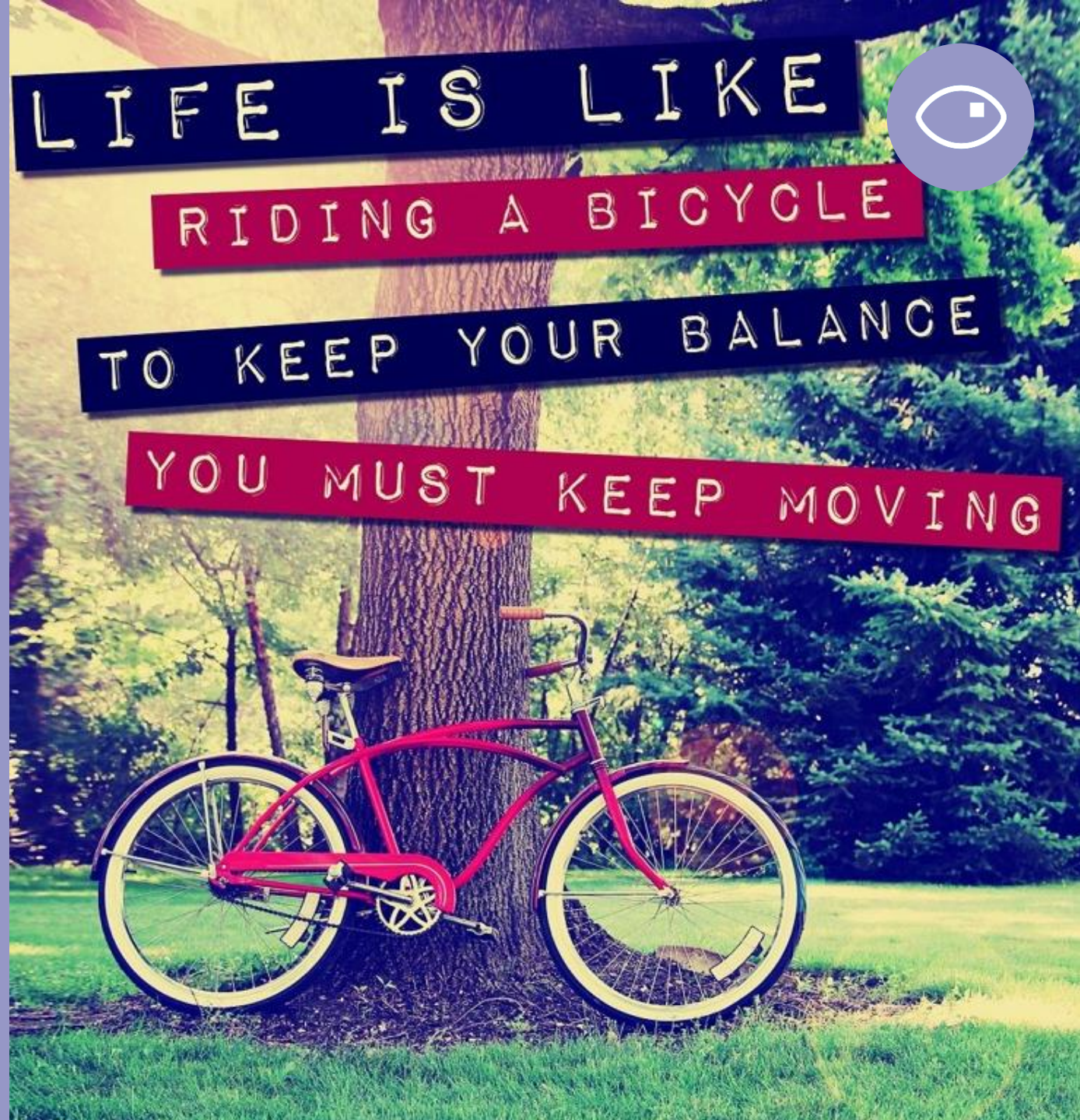
Life is like riding a bicycle

To keep your balance you must keep moving



OCR


```
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"
      }
. . .
```

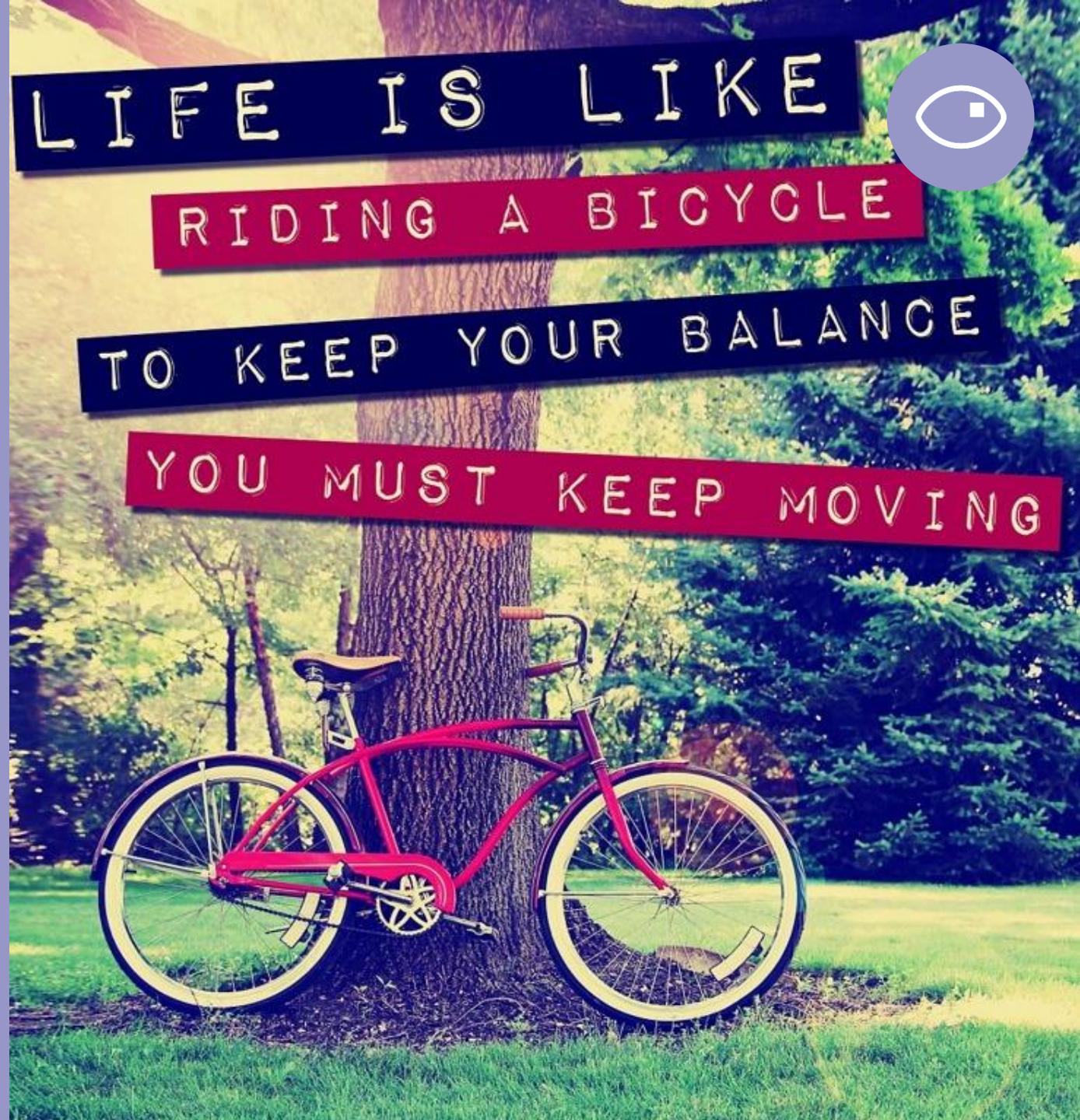
OCR

Good at

Scanned documents

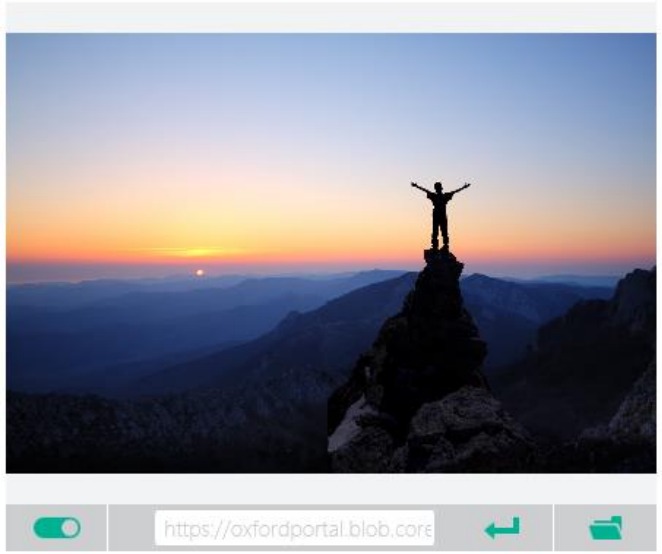
Photos with text

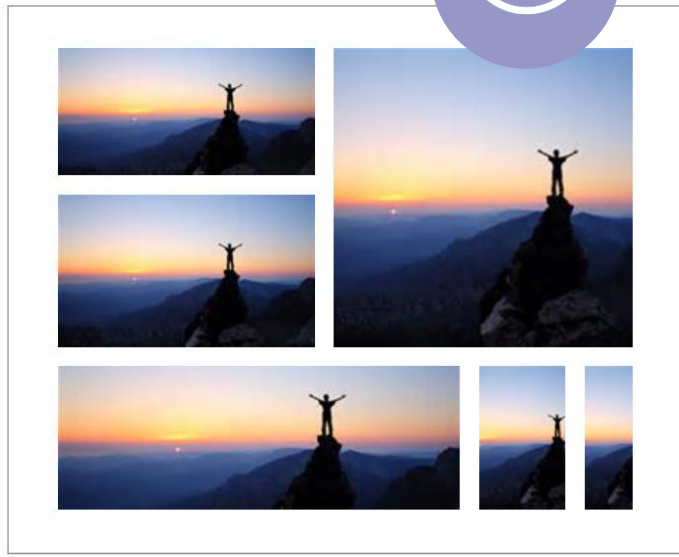
Fine grained
location information

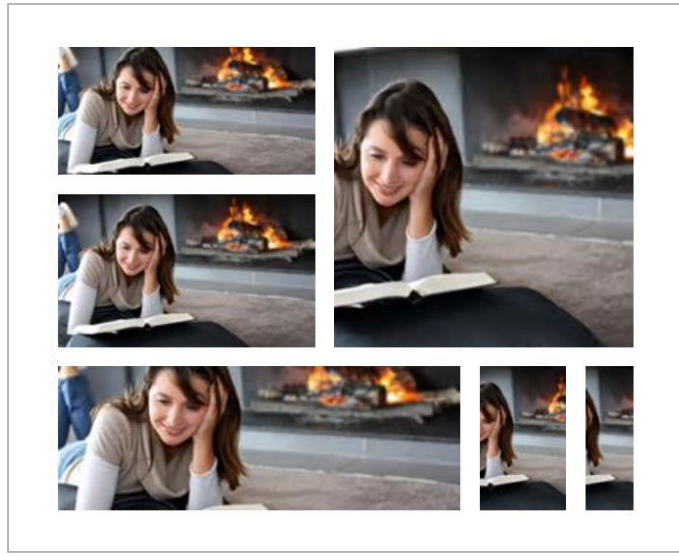


Smart thumbnail

Smart cropping **off**







Emotion API

Recognize emotions

Understand content within an image

Emotion API

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 API

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 API

Detection

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

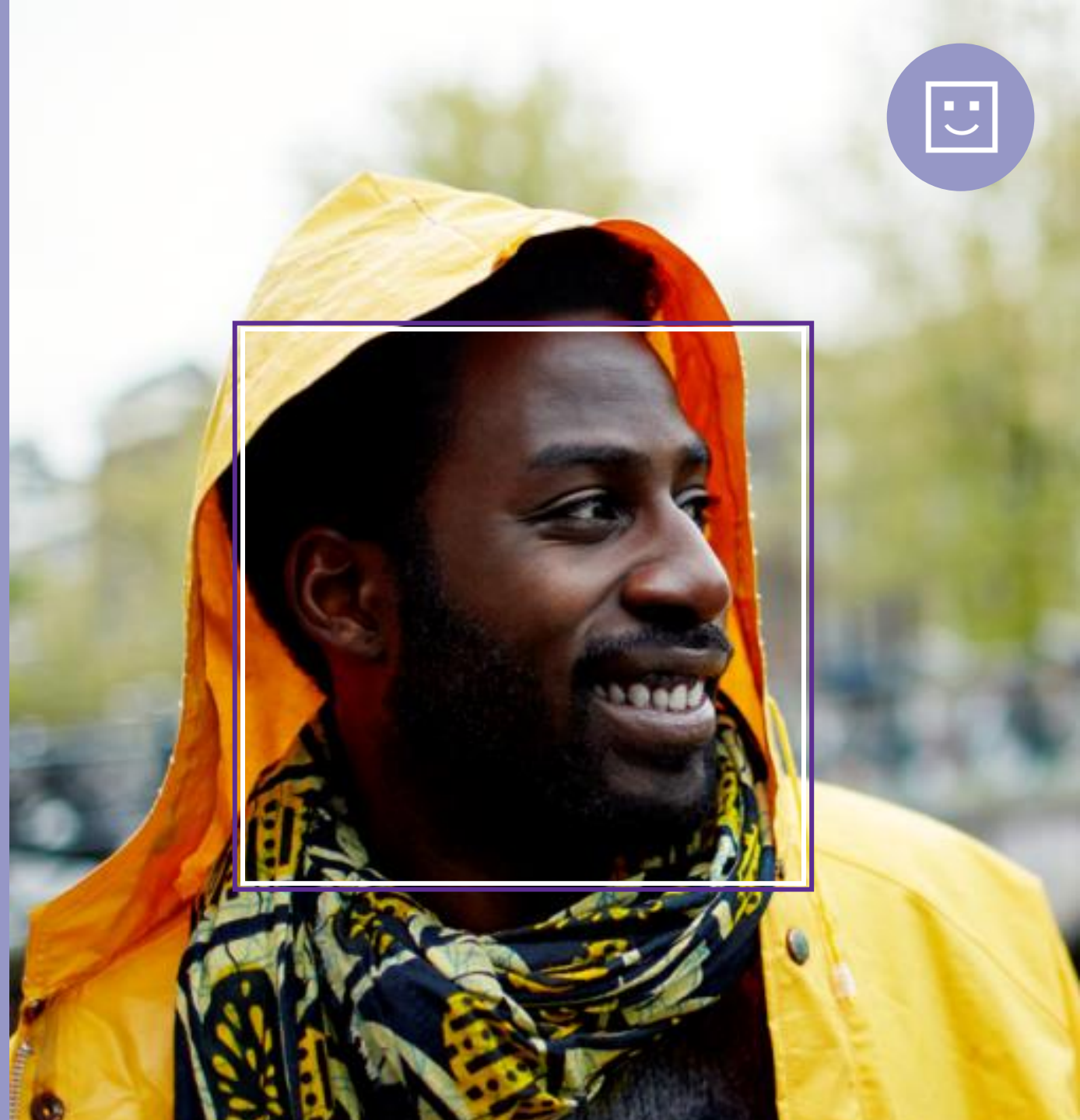
Feature attributes

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

Grouping

Identification

Jasper Williams



Content Moderator

Machine-assisted moderation of text and images, augmented with human review tools

Image moderation

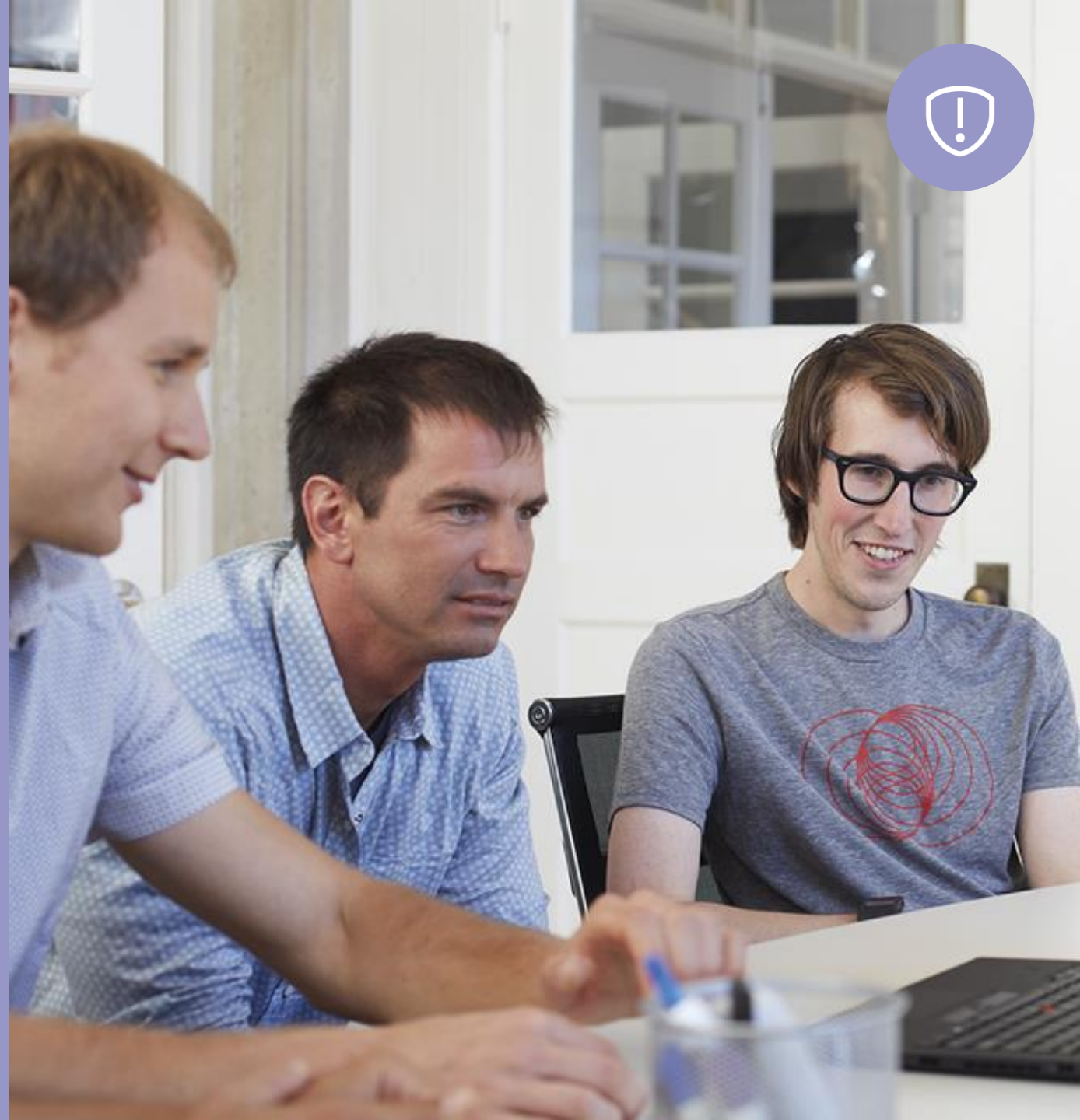
Enhance your ability to detect potentially offensive or unwanted images through machine-learning based classifiers, custom blacklists, and Optical Character Recognition (OCR)

Text moderation

Helps you detect potential profanity in more than 100 languages and match text against your custom lists automatically. Content Moderator also checks for possible Personally Identifiable Information (PII)

Video moderation (in Azure Media Services)

Enable the scoring of possible adult content in videos. Video moderation is currently deployed in preview on Azure Media Services



Custom Vision Service

A customizable web service that learns to recognize specific content in imagery

Upload images

Upload your own labeled images, or use Custom Vision Service to quickly tag any unlabeled images

Train

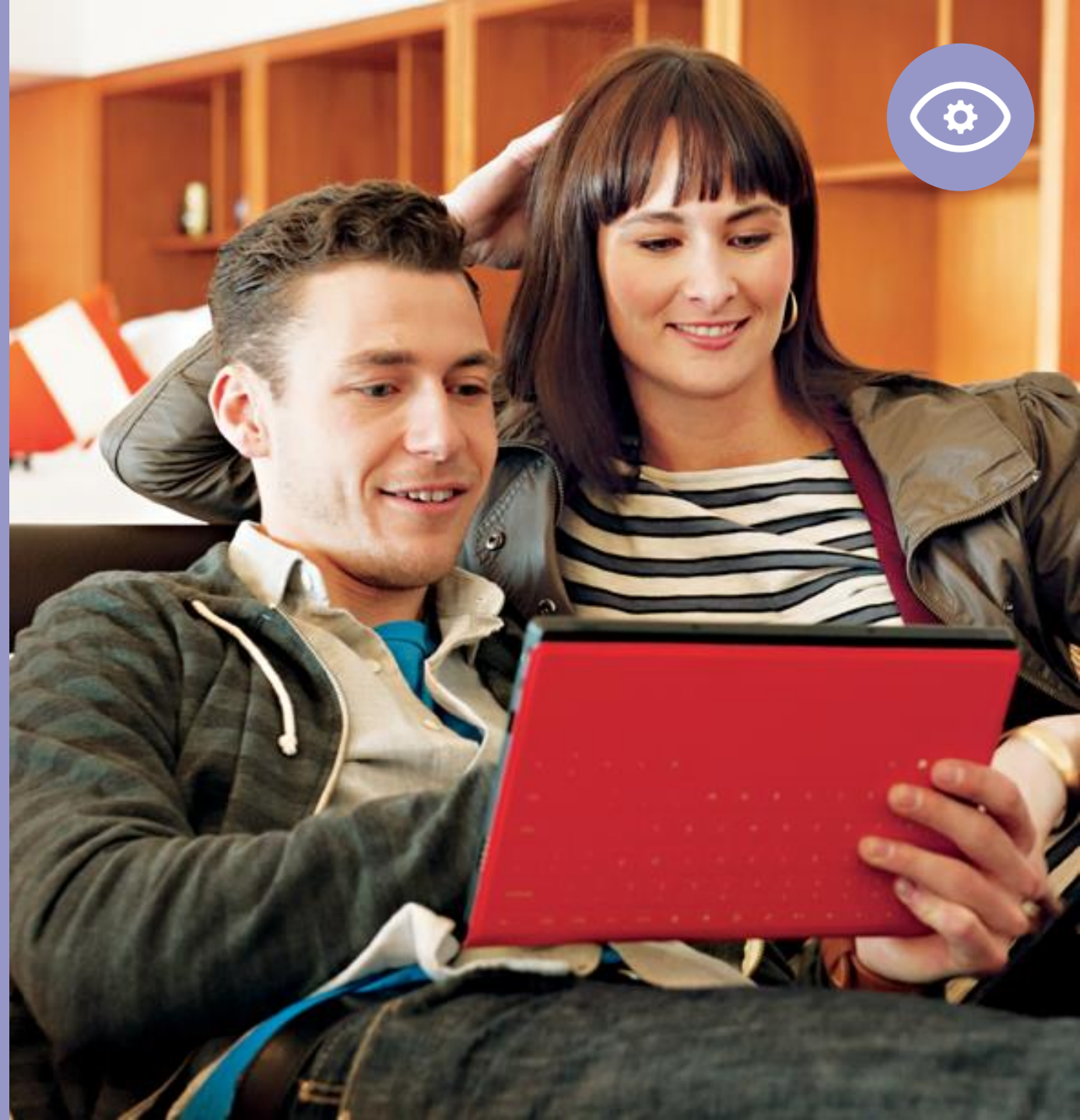
Use your labeled images to teach Custom Vision Service the concepts you want it to learn

Evaluate

Use simple REST API calls to quickly tag images with your new custom computer vision model

Active learning

Images evaluated through your custom vision model become part of a feedback loop you can use to keep improving your classifier



Video Indexer

Unlock video insights

Upload your video and go

Start turning your video into insights right away. No more tedious and error-prone manual indexing. And no need for specialized expertise. With Video Indexer, just upload your video, and start finding insights right away, without writing a single line of code

Make your content more discoverable

Quickly and easily extract insights from videos using artificial intelligence. Enhance content discovery experiences such as search results by detecting spoken words, faces, characters, and emotions

Improve engagement with your video

Metadata extracted by Video Indexer can be used to build powerful engagement experiences with recommendations, highlight clips, and interactive videos

Thak you,
for your attention!

Cristian Kevorchian
cristian.kevorchian@unibuc.ro