



从零入坑—Azure 给了我 AI 的眼睛

认知服务：视觉

胡浩





胡浩



Award Categories

Microsoft Azure

First year awarded:

2004

Number of MVP Awards:

14



多年多方向微软最有价值专家，多年从事云及数据中心基础架构、全栈虚拟化、企业移动管理等架构和咨询工作，热爱有趣的新技术并乐于分享。

2004, Microsoft MVP for Windows Server - Networking

2006, Microsoft MVP for Windows - Shell/User

2007, Microsoft MVP for Windows - Mobile Device

2008, Microsoft MVP for Windows - Mobile Device

2009, Microsoft MVP for Management Infrastructure

2010, Microsoft MVP for System Center Configuration Manager

2011, Microsoft MVP for Management Infrastructure

2012, Microsoft MVP for Management Infrastructure

2013, Microsoft MVP for System Center Cloud and Datacenter Management

2014, Microsoft MVP for System Center Cloud and Datacenter Management

2015, Microsoft MVP for System Center Cloud and Datacenter Management

2016, Microsoft MVP for Cloud and Datacenter Management

2017, Microsoft MVP for Cloud and Datacenter Management

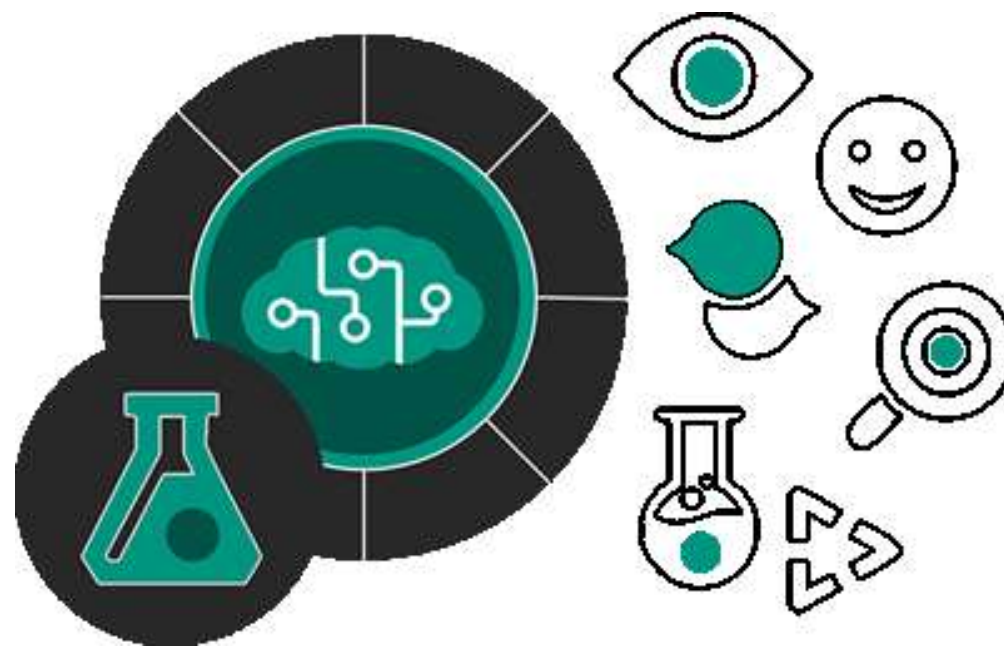
2018, Microsoft MVP for Microsoft Azure



微软的认知服务体系



- 影像 Vision
- 知识 Knowledge
- 语言 Language
- 语音 Speech
- 搜索 Search



影像服务 类别



计算机影像

[免费试用](#) | [了解更多](#)

图像分类

识别图像中的场景和活动 | [演示](#)

识别图像中的名人和地标性建筑 | [演示](#)

图像光学字符识别 (OCR) | [演示](#)

识别手写体 | [演示](#)



视频索引器 预览版

[免费试用](#) | [了解更多](#)

检测视频中的人脸

检测视频中的对象、场景和活动

提取并分析元数据、音频和关键帧



自定义视觉 预览版

[免费试用](#) | [了解更多](#)

识别可自定义图像 | [演示](#)



人脸

[免费试用](#) | [了解更多](#)

检测图像中的人脸 | [演示](#)

识别图像中人物

识别图像中的表情 | [演示](#)

识别图像中的相似人脸并分组 | [演示](#)



内容审查器

[免费试用](#) | [了解更多](#)

审查图像和视频的色情和冒犯性内容

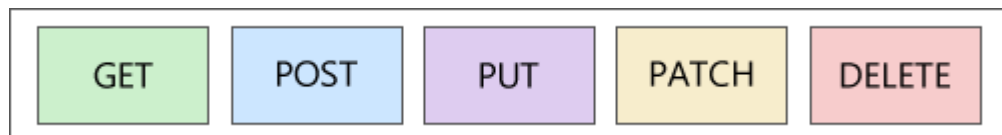
自定义图像和文本列表，可阻止或允许匹配内容

用于包含审阅人反馈的工具

认知服务，如此简单



REST服务



- 所有Microsoft认知服务都是通过REST (Representational State Transfer) 端点公开的
- 可以在任何平台上访问服务的特性和功能，以及任何能够进行简单HTTPS Web调用的设备。

☒ Python

☒ C#

☒ Visual Basic

☒ JavaScript

☒ Node.js

☒ Ruby

☒ PHP

☒ cURL

影像服务 API 功能



Computer Vision API提供的方法常分为以下几类:

- 图像和视频分析
 - 分析, 描述和标记图像和视频, 包括内容审核
 - OCR
 - 检测图像中的文本并提取识别的字符
- 缩略图
- 使用基于感兴趣区域的“智能”裁剪生成缩略图

The screenshot shows the 'APIs Documentation > API Reference' page. On the left, a list of API endpoints is visible, including 'POST Analyze Image', 'POST Describe Image', 'GET Get Handwritten Text Operation Result', 'POST Get Thumbnail', 'GET List Domain Specific Models', 'POST OCR', 'POST Recognize Domain Specific Content', 'POST Recognize Handwritten Text', and 'POST Tag Image'. A blue arrow points from the 'POST Analyze Image' endpoint to a sample image analysis result. The result is a JSON object with the following structure:

FEATURE NAME	VALUE
Description	{ "Tags": ["water", "swimming", "sport", "pool", "person", "man", "frisbee", "ocean", "blue", "bird", "riding", "top", "standing", "wave", "young", "body", "large", "game", "glass", "pond", "playing", "board", "catch", "clear", "boat", "white"], "Captions": [{ "Text": "a man swimming in a pool of water", "Confidence": 0.8909298 }] }
Tags	[{ "Name": "water", "Confidence": 0.9997857 }, { "Name": "swimming", "Confidence": 0.955619633 }, { "Name": "sport", "Confidence": 0.953807831 }, { "Name": "pool", "Confidence": 0.9515978 }, { "Name": "person", "Confidence": 0.889862537 }, { "Name": "water sport", "Confidence": 0.664259 }] }
Image format	"Jpeg"



影像服务 API 场景



```
conn = httplib.HTTPSConnection('westcentralus.api.cognitive.microsoft.com')
conn.request("POST", "/vision/v1.0/analyze?%s" % params, body)
response = conn.getresponse()
data = response.read()
```

计算机视觉API是利用人工智能的力量的完美方式:

- 根据内容标记图像
- 分类图像
- 确定图像的类型和质量
- 检测人脸并返回其坐标
- 识别特定于域的内容
- 生成内容的描述
- 使用光学字符识别识别图像中的打印文本
- 识别手写文本
- 区分配色方案
- 标记成人内容
- 剪裁要用作缩略图的照片



要使用，先授权



- Ocp-Apim-Subscription-Key
- 常见于Microsoft Cognitive Services中
- Subscription-Key更具描述性，特定于服务的版本

<https://azure.microsoft.com/try/cognitive-services/>

开放云原则（OCP） Azure API管理（APIM） 订阅密钥



还没有账号？ 免费的！

<https://signup.microsoft.com>



欢迎，我们开始吧



创建 Microsoft 帐户

如果使用 Outlook.com 或 Xbox Live，或者已经拥有 Microsoft 帐户，请输入该帐户以[登录](#)



创建工作或学校帐户

如果你的组织使用 Office 365、Dynamics 365 或 Power BI，请使用工作或学校帐户[登录](#)

分类功能



Object											
Food				Trans				Drink			
Bread		Pisa		Bus		TrainStation		Drink		Can	
Grilled		Fastfoo		Market		Car		Bicycle		Object	
								Screen		Sculpture	
Animal				Text				Plant			
Cat		Panda		Mag		Text		Tree		Branches	
Bird		Dog		Horse		Menu		Sign		Flower	
						Map				Leave	
People						Building					
Crowd		Tattoo		Swimming		DoorWindow		Stair			
Young		People		Group		Show		Corner		Pillar	
Baby		Many		Hand		Portrait		Arch		Brickwall	
										Church	

Abstract					Others
Net	Nonphol	Rect	Shape	Texture	
Scene					
Indoor			Dark		
Venue	MarketStore		Fire	Firworks	LightIND:
DoorWindow	Room		Sky		
			Rainbow	Object	
ChurchWind	Court		Cloud	Sun	
Outdoor					
Field	House	StoneRock	SwimmingPool		
			Watersid	Street	
OceanBeacl	Grass	Road			
City	Water	Mountain	Palygroun	Railway	



描述和标记功能



```
"tags": [  
  {  
    "confidence": 0.9926539659500122,  
    "name": "\u5efa\u7b51"  
  },  
  {  
    "confidence": 0.9916537404060364,  
    "name": "\u6237\u5916"  
  },  
  {  
    "confidence": 0.9839764833450317,  
    "name": "\u5929\u7a7a"  
  },  
  {  
    "confidence": 0.8205077052116394,  
    "name": "\u7ea2"  
  }  
]
```



A large red building with tiananmen square in the background

```
"description": {  
  "captions": [  
    {  
      "confidence": 0.8968637267424504,  
      "text": "a large red building with Tiananmen Square in the background"  
    }  
  ],  
}
```

描述: a large red building with Tiananmen Square in the background

标记: 建筑, 户外, 天空, 红



图片类型功能

Is Line Drawing? **TRUE**



Is Line Drawing? **FALSE**



Clip Art Type: **3 GOOD QUALITY**

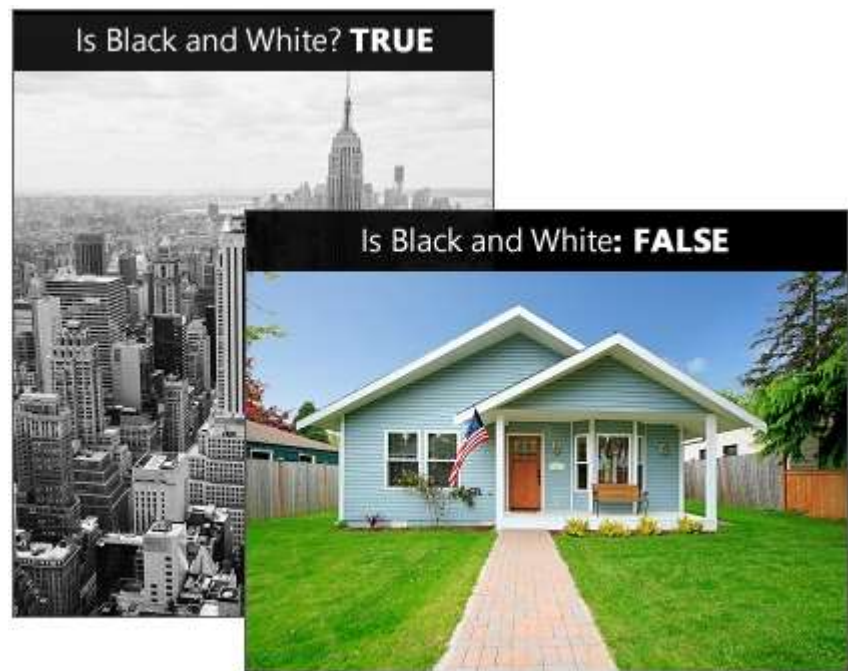


Clip Art Type: **0 NON CLIP ART**



颜色功能

BLACK	BLUE	BROWN	GRAY	GREEN	ORANGE
PINK	PURPLE	RED	TEAL	WHITE	YELLOW



人脸功能



几位女性?

几个人戴了眼镜?





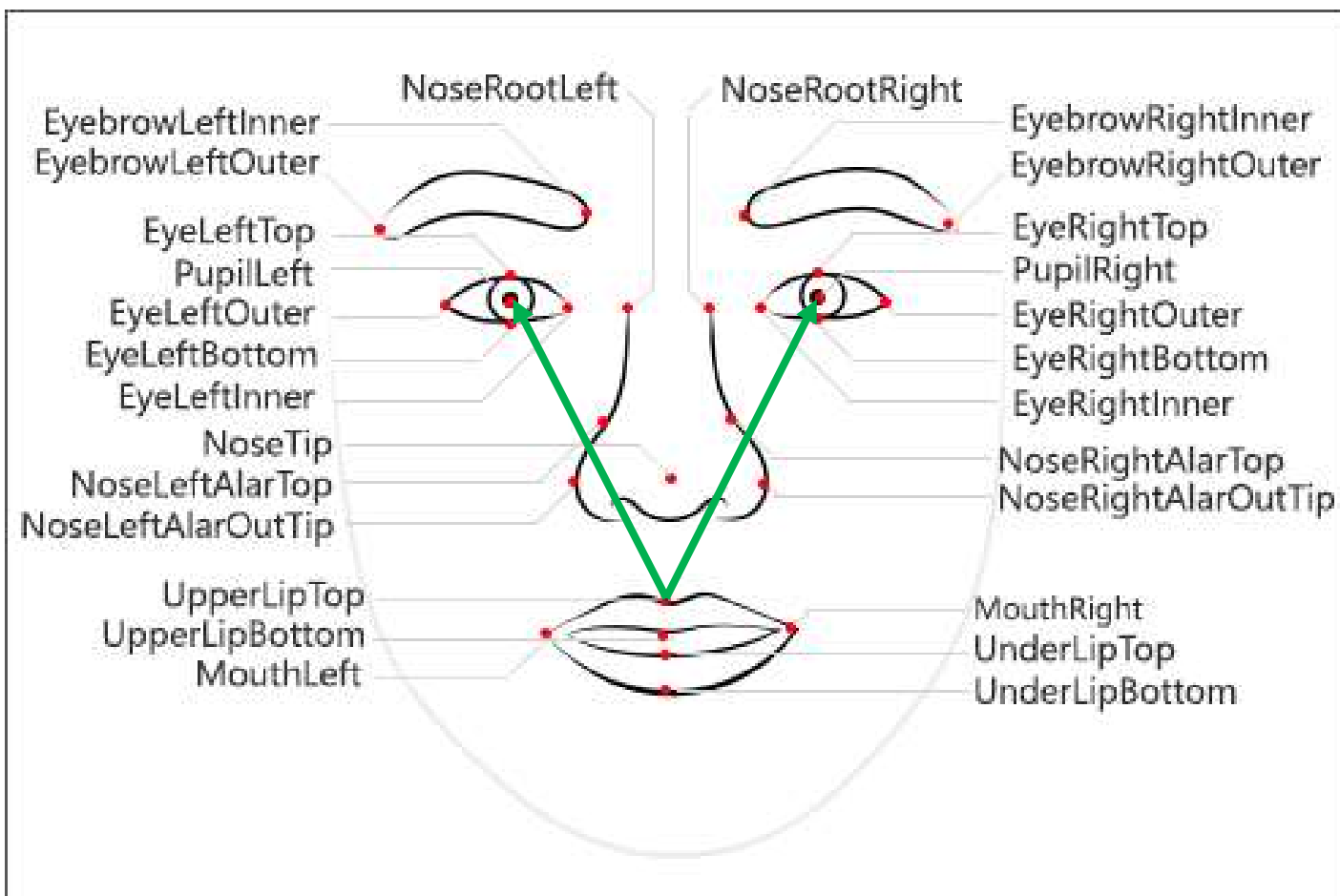
为什么没能检测?

<https://azure.microsoft.com/zh-cn/services/cognitive-services/face/#detection>

- 矩形区域像素过低就无法识别面部特征
- $\leq 4\text{MB}$
- 如果需要识别人脸，一次不要超过10个



这可不是看相



27个特征点

- 人脸特征点是人脸上的系列具体特征点
- 通常是瞳孔、眼角或鼻子等人脸组成部分特征点
- 人脸特征点是可在人脸检测期间分析的选属性
- 人脸方向检测



人脸识别



- 使用多个faceID训练一个personID
- 使用faceID识别personID

```
[Running] python "c:\Personal\Hao.Hu\My Codes\Python\Cognitive\FaceAPI\faces.py"
Satya Nadella
personID:
c3c6de37-7eec-4dde-895b-1904f8869293
faceID:
5ad8cc20-ff01-4b64-9d4e-7b7be6769d77
642a1ef5-cef7-459c-8a07-8d4e8bcb6d03
a1abdd3a-150a-40d7-9ff4-8afd8a29ba35

[Running] python "c:\Personal\Hao.Hu\My Codes\Python\Cognitive\FaceAPI\identify.py"
Based on confidebce 0.76732, I found Hao Hu in the photo.
personID:
06576e9b-dea7-44e5-8073-9e48158491de
Based on confidebce 0.80839, I found XiangYang Shen in the photo.
personID:
c153727b-919b-49bd-afbe-f0b3b5fb015a
Based on confidebce 0.88311, I found Satya Nadella in the photo.
personID:
c3c6de37-7eec-4dde-895b-1904f8869293
Based on confidebce 0.65814, I found ZhengLei Pan in the photo.
personID:
e0655727-9766-465c-b58f-5ed8067ae961

[Done] exited with code=0 in 10.436 seconds
```



ceo01.jpg



ceo02.jpeg



ceo03.png



pz01.jpg



pz02.jpg



pz03.jpg



sxy01.jpg



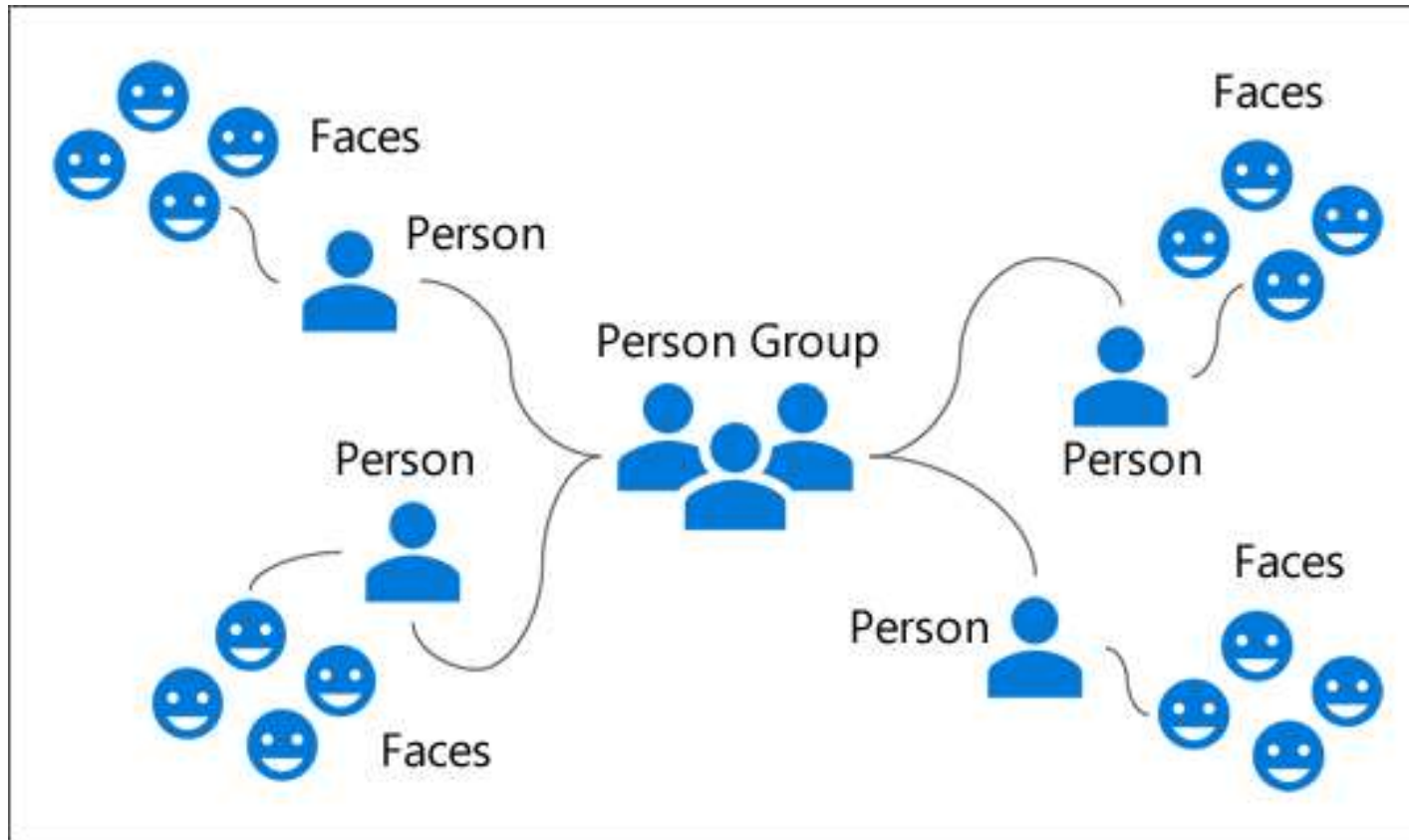
sxy02.jpg



sxy03.jpg



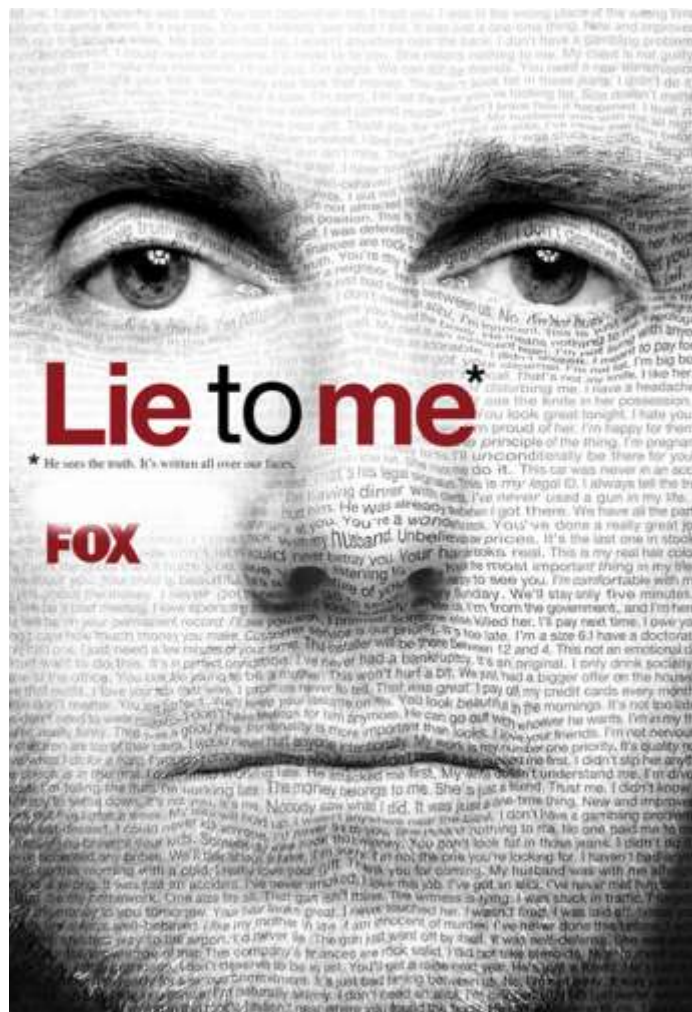
千人千面



- **Face List**
 - Up to 1,000 distinct faces
- **Person Group**
 - Up to 10,000 persons
- **Person**
 - Up to 248 faces



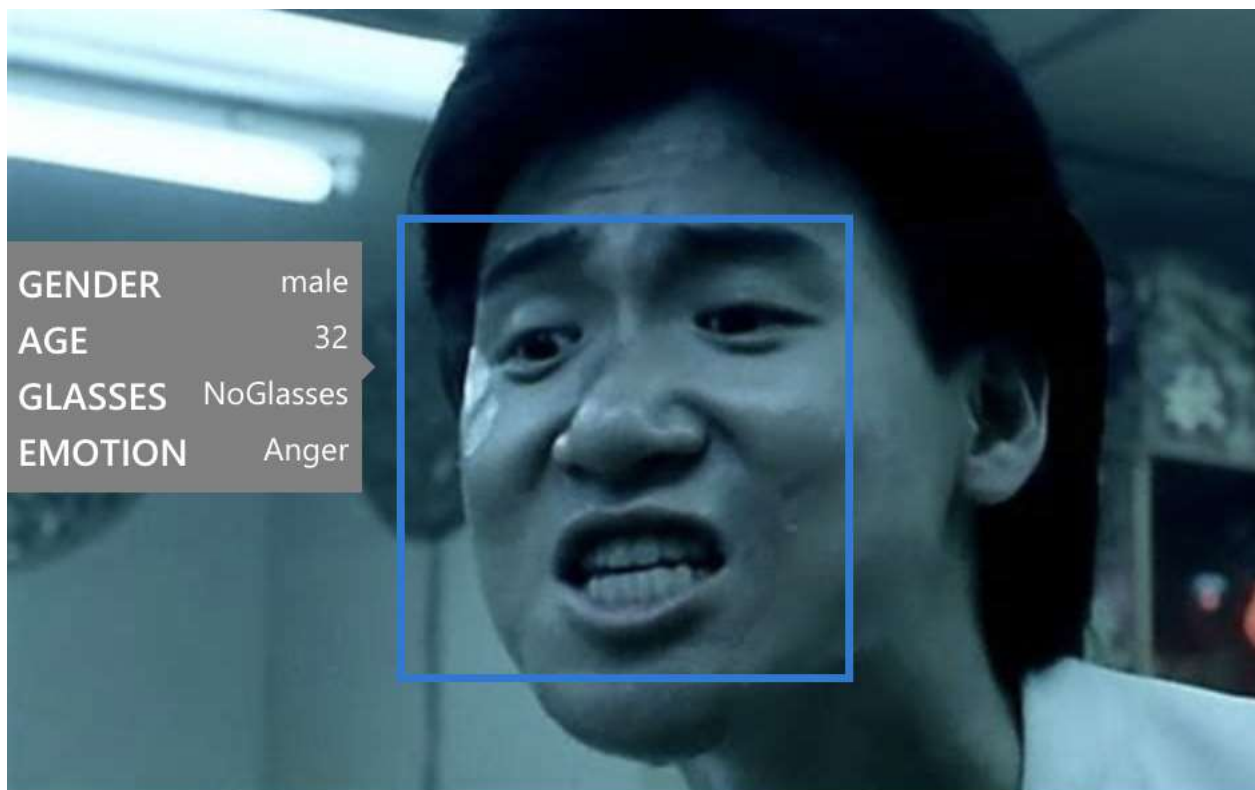
表情识别



- 七情六欲
 - 喜、怒、哀、恐、惊、恶、轻+没表情~
 - happiness, anger, sadness, fear, surprise, disgust, contempt and neutral
- 以评分标记可能的表情



表情包是啥表情



```
"glasses": "NoGlasses",  
"makeup": {  
  "eyeMakeup": false,  
  "lipMakeup": false  
},  
"emotion": {  
  "anger": 0.374,  
  "contempt": 0.002,  
  "disgust": 0.086,  
  "fear": 0.031,  
  "happiness": 0.062,  
  "neutral": 0.12,  
  "sadness": 0.309,  
  "surprise": 0.015  
},  
"occlusion": {  
  "foreheadOccluded": false,  
  "eyeOccluded": false,  
  "mouthOccluded": false  
},
```



呵呵一下



成人功能



Sola Aoi - Wikipedia

安全 | https://en.wikipedia.org/wiki/Sola_Aoi

Not logged in | Talk | Contributions | Create account | Log in

Article | Talk | Read | Edit | View history | Search Wikipedia

Sola Aoi


From Wikipedia, the free encyclopedia

For the manga character, see Sora Aoi (Aki Sora)

Sola Aoi (Japanese: 蒼井 そら Hepburn: *Aoi Sora*, literally **Blue Sky**)^[2] is the stage name of a Japanese AV idol, nude model, and film and television actress. Her popularity as an adult video actress has led to celebrity status as a media personality in Japan and abroad.

Contents [hide]

- 1 Life and career
 - 1.1 Early life and AV debut
 - 1.2 Into mainstream entertainment
 - 1.3 International celebrity
- 2 Filmography
 - 2.1 TV drama
 - 2.2 V-Cinema
 - 2.3 Theatrical releases
- 3 References
- 4 Bibliography
- 5 External links

Sola Aoi


Native name 蒼井 そら
Born November 11, 1983

Life and career

 [edit]

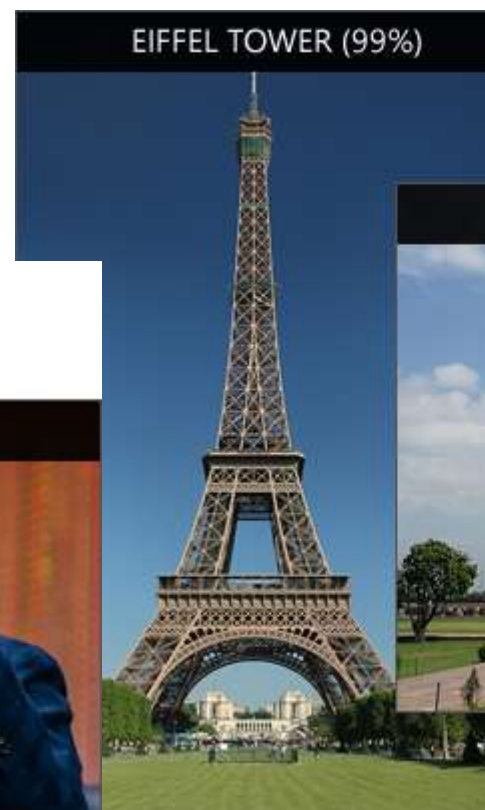
Sola Aoi lying on a bed



特定领域



名人

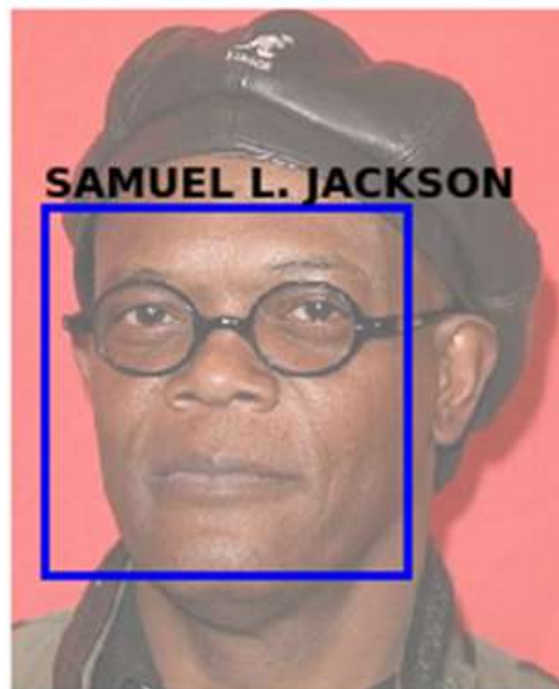


名胜



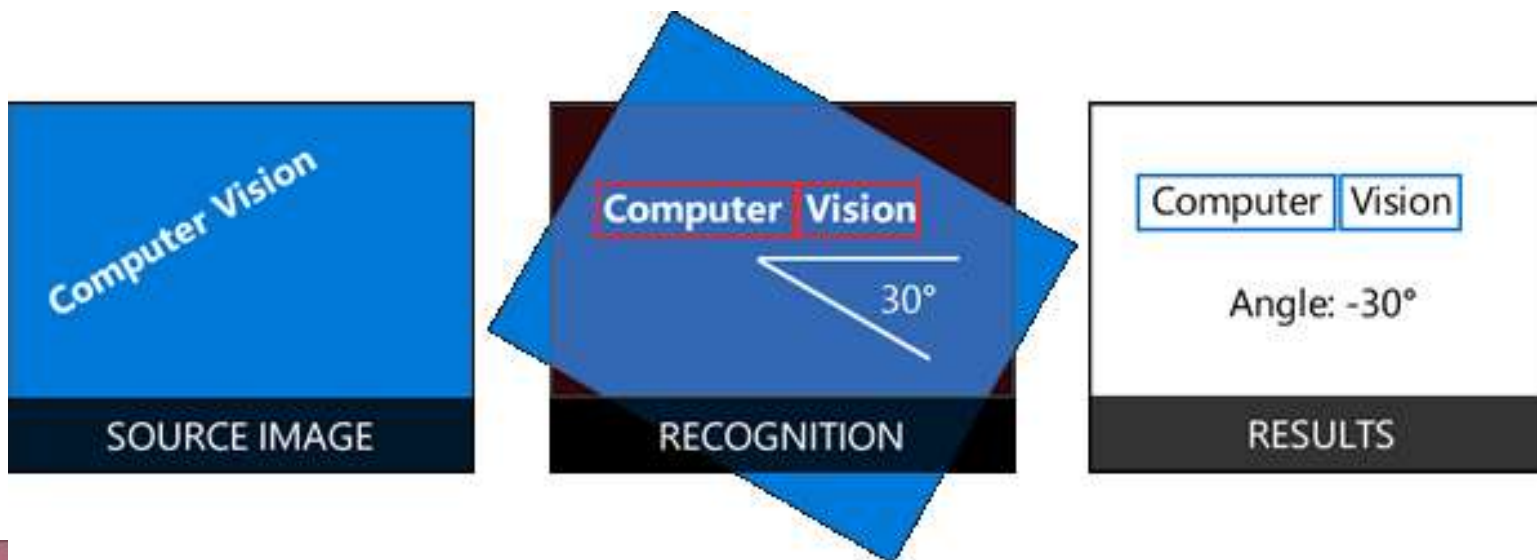
特定领域示例

Space needle



OCR 功能

- 文字识别
- 嵌入文字识别
- 手写文字识别



那么，举个🍌吧



```
14 headers = {
15     'Content-Type': 'application/octet-stream',
16     'Ocp-Apim-Subscription-Key': API_Key
17 }
18 params = {
19     'language': 'zh-Hans',
20     'detectOrientation': 'true'
21 }
22 imageFile = 'C:\\Personal\\Hao.Hu\\My
Codes\\Python\\Cognitive\\ComputerVisionAPI\\recidemo1.png'
23 imageData = open(imageFile, 'rb')
24
25 response = requests.post(API_URI, headers=headers, params=params,
data=imageData)
26 response.raise_for_status()
27 analysis = response.json()
28 print(analysis)
29 #parsed = json.loads(response.text)
30 #print(parsed)
31
32 line_infos = [region["lines"] for
33 word_infos = []
34 for line in line_infos:
35     for word_metadata in line:
36         for word_info in word_met
37             word_infos.append(word
38 word_infos
```



预览	JSON
餐饮服务	
60、38	
1	
60、38	
3、62	
计	
¥3、62	
¥60-38	
陆拾肆圆整	
价税合计（大写）	
（小写）¥64、00	
名	
称：北京增信优有限公司	

```
无标题 - 记事本
文件(F) 编辑(E) 格式(O) 查看(V) 帮助(H)
'238,280,21,19'], ['text': '：', 'boundingBox': '263,288,3,11'], ['boundingBox': '123,280,143,19'], ['words': [{'text':
'方', 'boundingBox': '75,300,19,19'}], ['boundingBox': '75,300,19,19'], ['words': [{'text': '开', 'boundingBox':
'124,320,19,19'}], ['text': '户', 'boundingBox': '148,318,16,21'], ['text': '行', 'boundingBox': '169,319,21,20'], ['text': '
及', 'boundingBox': '193,320,19,17'], ['text': '账', 'boundingBox': '215,319,21,20'], ['text': '号', 'boundingBox':
'238,318,20,22'], ['text': '：', 'boundingBox': '263,327,3,10'], ['boundingBox': '124,318,142,22'], ['words': [{'text':
'价', 'boundingBox': '158,687,21,21'}, ['text': '税', 'boundingBox': '180,688,21,19'}, ['text': '合', 'boundingBox':
'204,687,20,21'], ['text': '计', 'boundingBox': '227,687,19,21'], ['text': '（', 'boundingBox': '253,688,6,20'}, ['text': '
大', 'boundingBox': '263,688,18,19'}, ['text': '写', 'boundingBox': '285,688,18,20'}, ['text': '）', 'boundingBox':
'306,688,6,19'}], ['boundingBox': '158,687,154,21'}], ['boundingBox': '453,683,152,26', 'lines': [{'words': [{'text': '8',
'boundingBox': '453,683,27,26'}, ['text': '陆', 'boundingBox': '494,686,19,21'], ['text': '拾', 'boundingBox':
'516,686,21,21'}, ['text': '肆', 'boundingBox': '539,686,21,21'}, ['text': '圆', 'boundingBox': '563,687,19,20'}, ['text': '
整', 'boundingBox': '585,686,20,20'}], ['boundingBox': '453,683,152,26'}], ['boundingBox': '124,736,18,19', 'lines':
```



一个发票识别的例子

北京增1增值税专用发票

发票代码: 011001600H1
发票号码: 30227929
开票日期: 2016年10月22日
校验码: 47396 34079 33946 77720

机器编号: 661616242727

名称: 思杰系统信息技术(北京)有限公司分公司
纳税人识别号:
地址、电话:
开户行及账号:

115 34 6 619 -9 82 2 0 9 < - * 02 * * 3 /
< 31 * 0 < < 418 6 4 12 < 9 - 4 5 9 7 4 < 9 7 < *
7 / 2 - 0 5 > 7 3 * 918796982409 < 4 0 5 < - 000
7 < 35 > 1 < * 0 < < 4186412 1 2 < 9 - 45 > X > +

货物或应税劳务、服务名称	规格型号	单位	数量	单价	金额	税率	税额
餐饮服务			1	60.38	60.38	6%	3.62
合计					¥60.38		¥3.62
价税合计(大写)		陆拾肆圆整		(小写) ¥64.00			

思考一下:

- 怎样在OCR图片中显示中文?



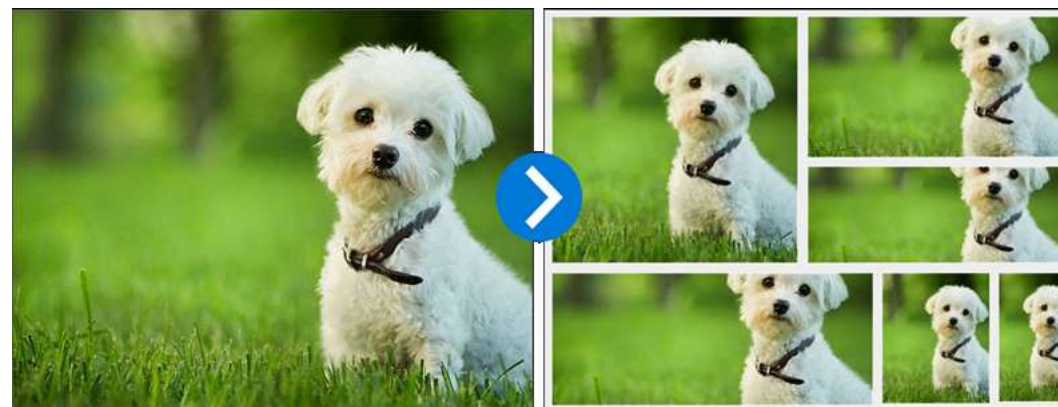
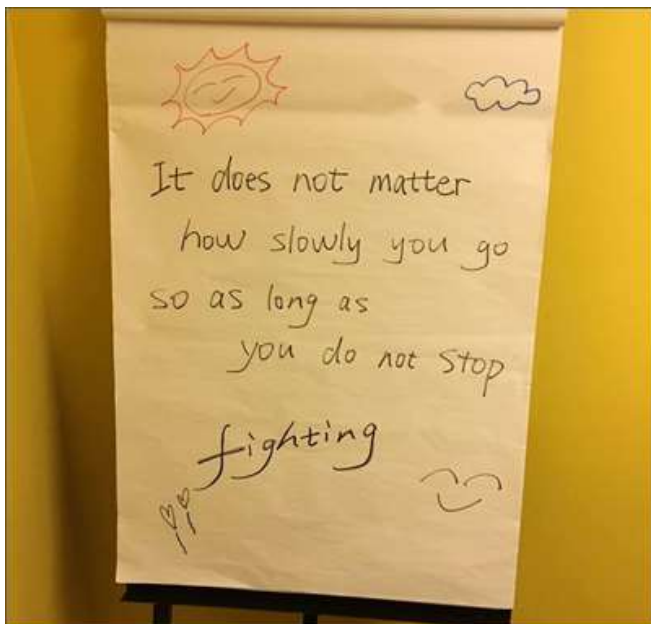
The unanimous Declaration of the thirteen united States of America.

[illegible]

```
https://westus.api.cognitive.microsoft.com/vision/v1.0/textOperations/5a523b1f-026c-4010-920c-384180c4b73e
{'recognitionResult': {'lines': [{'boundingBox': [14, 6, 948, 0, 949, 41, 15, 49],
'text': 'We had these truths to be felt-evident, that all men are created equal',
'words': [{'boundingBox': [33, 5, 85, 5, 56, 48, 4, 48], 'text': 'We'}, {'boundingBox': [90, 5, 164, 4, 135, 48, 61, 48], 'text': 'had'}, {'boundingBox': [161, 4, 235, 4, 206, 47, 132, 48], 'text': 'these'}, {'boundingBox': [235, 4, 327, 3, 298, 47, 206, 47], 'text': 'to be'}, {'boundingBox': [327, 3, 350, 46, 303, 47], 'text': 'felt-evident'}]}]}
```

手写识别和缩略/裁剪

- 识别手写文字
- 智能裁剪图片
- 智能生成缩略图



Custom Vision

<https://www.customvision.ai>



Cognitive Services

Custom Vision




Visual Intelligence Made Easy

Easily customize your own state-of-the-art computer vision models that fit perfectly with your unique use case. Just bring a few examples of labeled images and let Custom Vision do the hard work.

SIGN IN

自己动手 丰衣足食

- **Prediction**
 - uses a **Prediction Key**, supplied as a "Prediction-key" value
- **Training**
 - uses a **Training Key**, supplied as a "Training-key" value



Tag	Probability
Dog	100%
German Shepherd	100%
Large	100%
Small	0%
Chihuahua	0%



Name

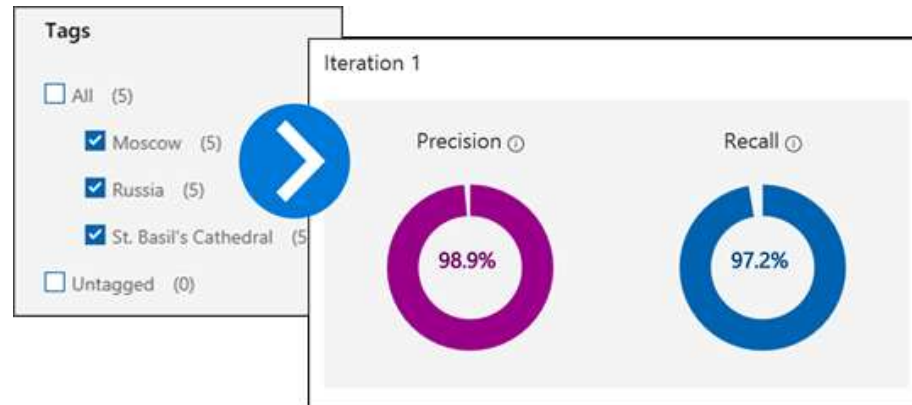
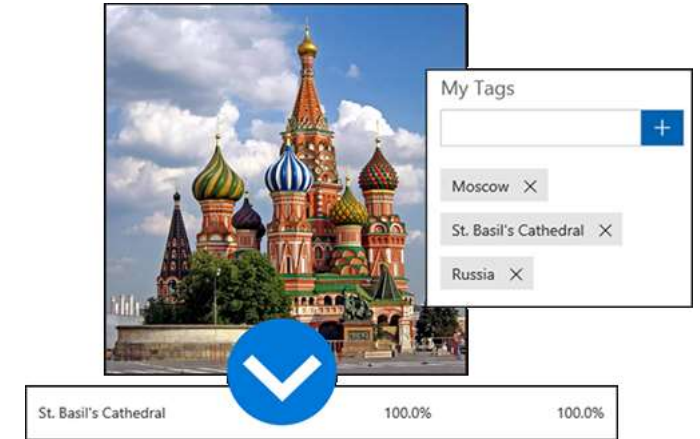
AI Vision Exp

Description

Identify spec

Domains ⓘ

- ☒ General
- ☐ Food
- ☐ Landmarks
- ☐ Retail
- ☐ Adult
- ☐ General (compact) ⓘ
- ☐ Landmarks (compact) ⓘ
- ☐ Retail (compact) ⓘ



A maximum of 20 active projects
A maximum of 1,000 image predictions per day



Content Moderator

<https://contentmoderator.cognitive.microsoft.com>



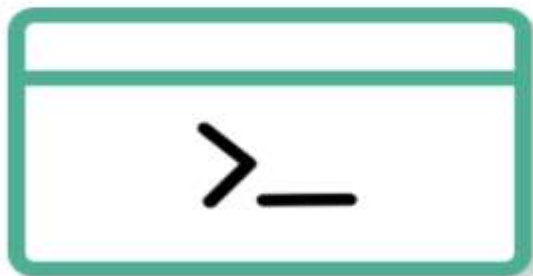
[Sign in](#)

Content Moderator

Test drive Content Moderator to learn how we enable a complete, configurable content moderation lifecycle by combining automated moderation with human reviews and workflows.

[Sign Up](#)

Moderate



Automated results reduce time and costs

Collaborate



Manage teams for increased productivity

Review



Approve or reject content for better filtering

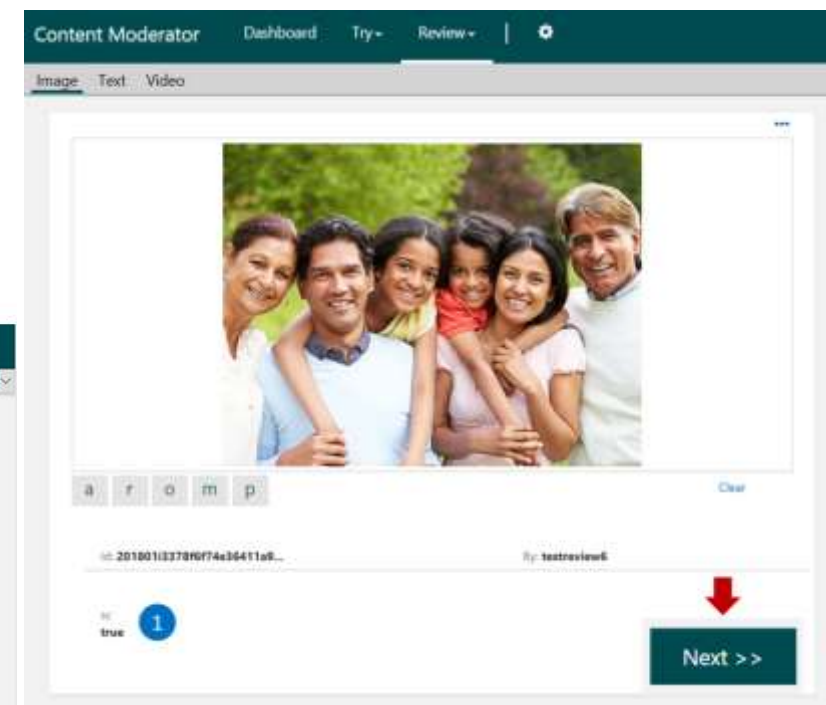
内容管控



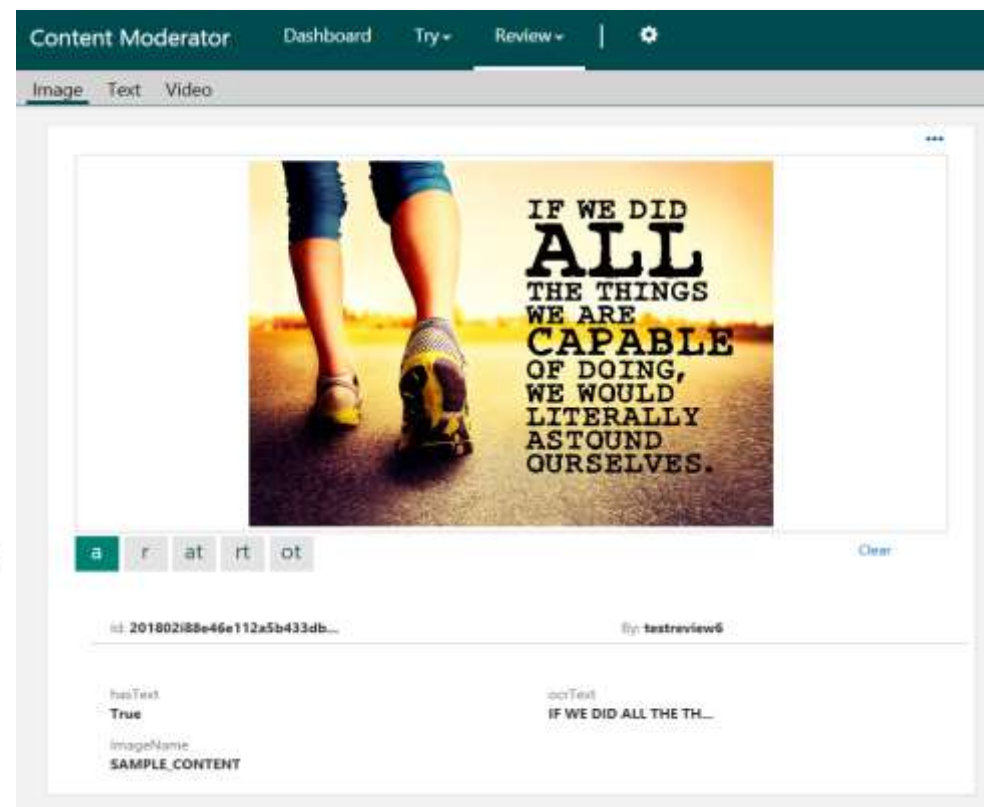
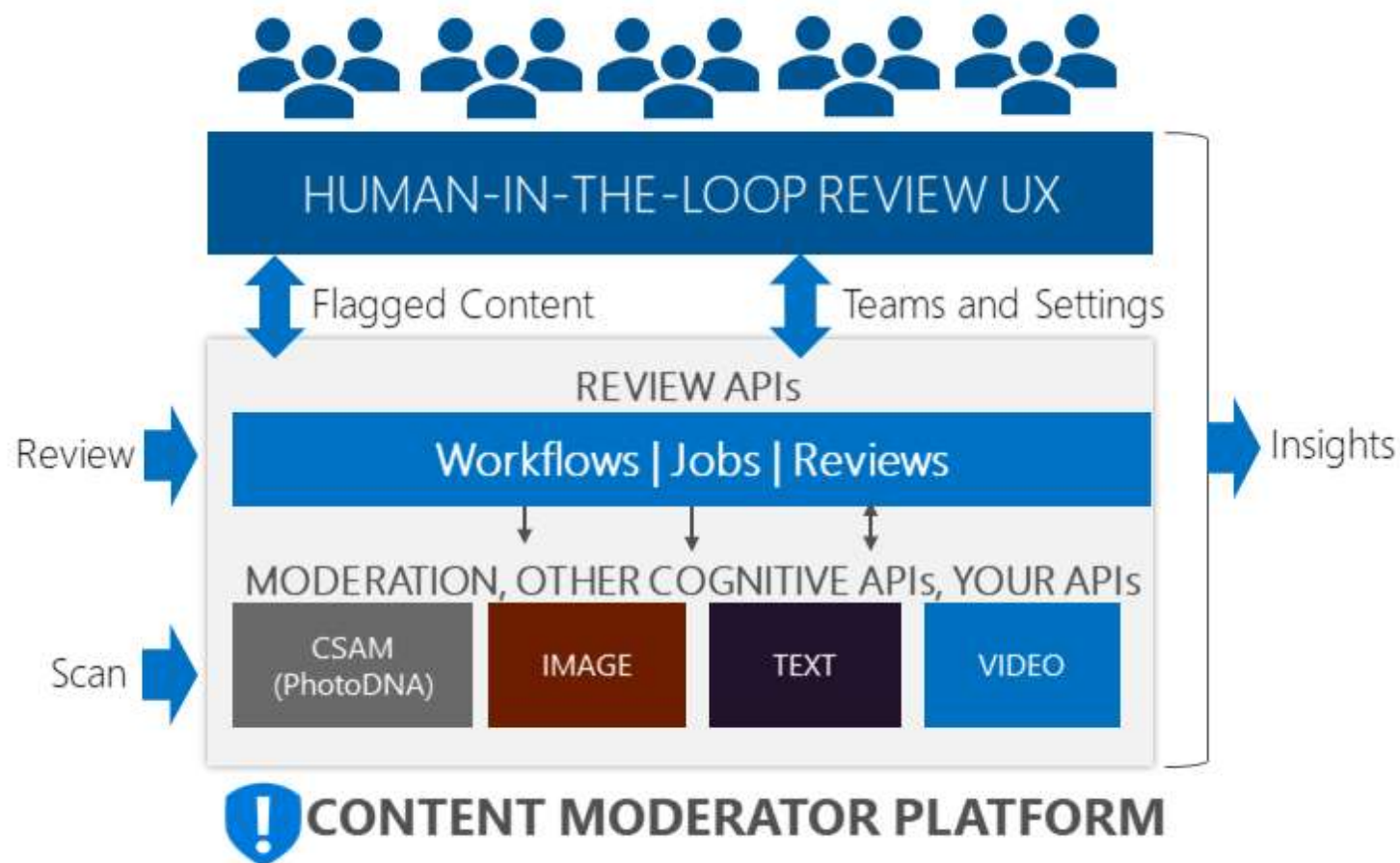
图像、视频、文字

Content Moderator interface with annotations:

- B&W and Blurring
- Cinema mode (expand player view)
- Transcript
- Add notes
- Expand Frames view
- Tagged only
- Sync Frames
- Frames
- Transcript
- Details
- Notes
- Next >>
- Tags: a r o m p
- Mute by default
- Highest scoring frame tag
- Player control bar
- Flagged frames
- Keyboard controls
- Select multiple frames for tagging
- Click to toggle selection
- Submit your decision

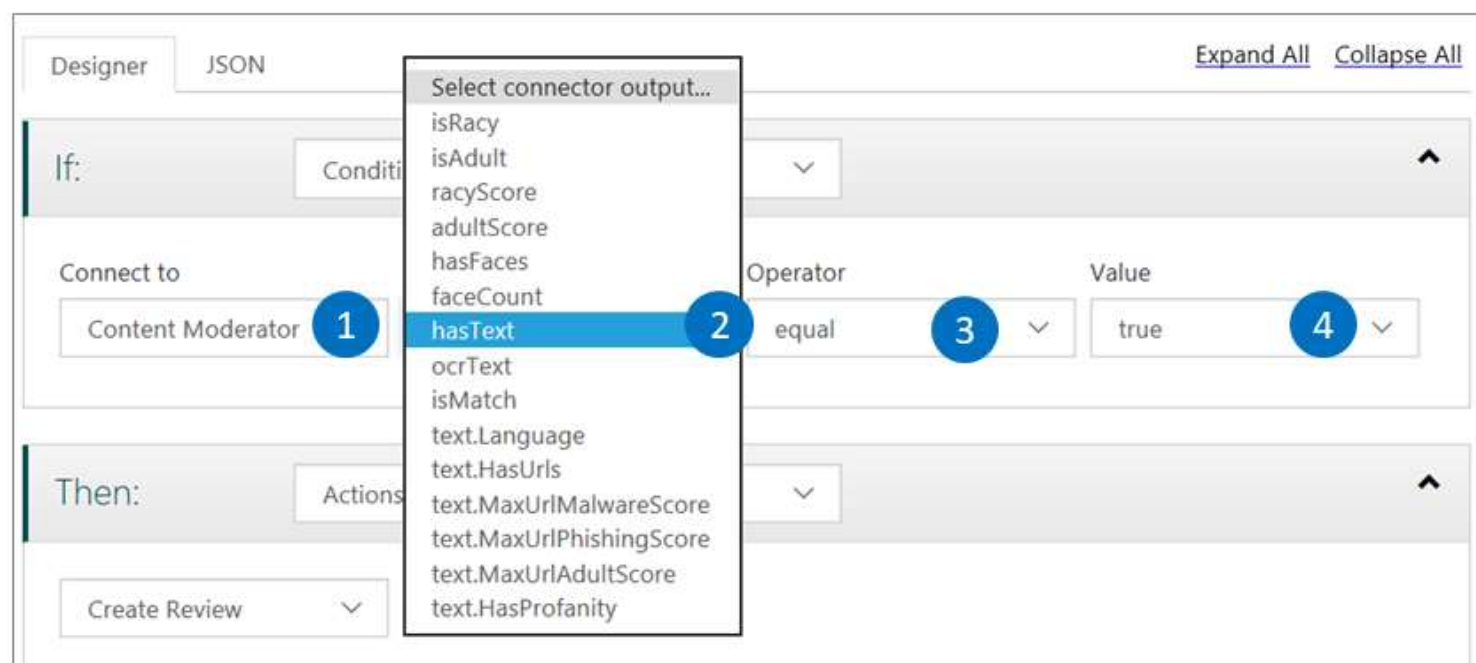


支持 workflow



workflow 实现

```
{
  "Type": "Logic",
  "If": {
    "ConnectorName": "moderator",
    "OutputName": "hasText",
    "Operator": "eq",
    "Value": "true",
    "Type": "Condition"
  },
  "Then": {
    "Perform": [
      {
        "Name": "createreview",
        "CallbackEndpoint": null,
        "Tags": [
          {
            "Tag": "a",
            "IfCondition": {
              "ConnectorName": "moderator",
              "OutputName": "hasText",
              "Operator": "eq",
              "Value": "true",
              "Type": "Condition"
            }
          }
        ]
      }
    ]
  },
  "Type": "Actions"
}
```





Community
Meetup
Shanghai

vă mulțumesc

Salamat sa iyo

Grazie

Спасибо

Teşekkür ederiz

Jamädi

Gracias

Merci

Aitäh Dziękuje Maururu ia oe

Danke

Obrigado

diolch

Terima kasih

Cảm ơn bạn

תודה

Vinaka vakalevu

Thank you

谢谢

Tak

Благодаря

köszönöm

Ευχαριστούμε

Hvala tí

ขอบคุณ

謝謝

Dėkuju

Dankie

Gràcies

köszönöm

감사합니다

hvala

Dank u

धन्यवाद

شكرا

ありがとう

Дякую