

MODERAÇÃO DE CONTEÚDO COM MACHINE LEARNING

Augusto Savi

Start

O QUE É MODERAÇÃO DE CONTEÚDO?

É a detecção de conteúdo impróprio, indesejado ou ofensivo, podendo ser feito em mídia social, mídia de transmissão, publicidade e situações de comércio eletrônico para criar uma experiência de usuário mais segura, fornecer garantias de segurança de marca aos anunciantes e cumprir as regulamentações locais e globais.

SERVIÇO UTILIZADO: AWS REKOGNITION DETECT MODERATION LABELS

- Muitas empresas dependem inteiramente de moderadores humanos para revisar o conteúdo de terceiros ou gerado pelo usuário, outras simplesmente reagem às reclamações dos usuários para remover imagens, anúncios ou vídeos ofensivos ou inadequados.
- Moderadores humanos sozinhos não podem escalar para atender a essas necessidades com qualidade ou velocidade suficiente, o que leva a uma experiência do usuário ruim, altos custos para alcançar escala ou até mesmo uma perda da reputação da marca.
- Os moderadores humanos podem revisar um conjunto muito menor de conteúdo, normalmente de 1 a 5% do volume total,

É POSSÍVEL CRIAR SUA PRÓPRIA CLASSIFICAÇÃO ?

Sim, a AWS disponibiliza o serviço Rekognition Custom Labels onde você pode criar a classificação que seu projeto precisa...

[AWS rekognition, classificação de imagens](#)

REGIÕES DISPONÍVEIS

US East (Ohio)

US West (N. California)

Asia Pacific (Mumbai)

Asia Pacific (Singapore)

Asia Pacific (Tokyo)

Europe (Frankfurt)

Europe (London)

US East (N. Virginia)

US West (Oregon)

Asia Pacific (Seoul)

Asia Pacific (Sydney)

Canada (Central)

Europe (Ireland)

AWS GovCloud (US-West)

[Back](#)

CUSTO

Região: Leste dos EUA (Ohio) ▾

Volume (Imagens/mês)	Imagens processadas por APIs do Grupo 1	Imagens processadas por APIs do Grupo 2
Primeiro 1 milhão de imagens	0,001 USD por imagem	0,001 USD por imagem
Próximos 4 milhões de imagens	0,0008 USD por imagem	0,0008 USD por imagem
Próximos 30 milhões de imagens	0,0006 USD por imagem	0,0006 USD por imagem
Mais de 35 milhões de imagens	0,0004 USD por imagem	0,00025 USD por imagem

*Cada API que aceita uma ou mais imagens de entrada conta como uma imagem processada. [Saiba mais »](#)



ALGUNS CLIENTES E SUA UTILIZAÇÃO



Pinterest usa a detecção de texto para moderar as imagens enviadas pelo usuário »



O PBS usa a detecção de objetos para simplificar as operações de conteúdo de mídia »



SDG&E usa etiquetas personalizadas do Rekognition para detectar danos causados por transformadores de imagens de drones »

DESENVOLVIMENTO DA APLICAÇÃO!

Go!

TECNOLOGIAS USADAS!

- docker-compose
- RabbitMQ
- Mongodb
- s3
- Spring Boot
- Lombok
- Swagger
- Aws-java-sdk
- CI/CD



EXEMPLO



```
{
  "moderationLabels": [
    {
      "confidence": 99.9941,
      "name": "Drug Products",
      "parentName": "Drugs"
    },
    {
      "confidence": 99.9941,
      "name": "Drugs",
      "parentName": ""
    }
  ]
}
```



```
{
  "ModerationLabels": [
    {
      "Confidence": 86.06100463867188,
      "Name": "Suggestive",
      "ParentName": ""
    },
    {
      "Confidence": 86.06100463867188,
      "Name": "Female Swimwear Or Underwear",
      "ParentName": "Suggestive"
    },
    {
      "Confidence": 72.55979919433594,
      "Name": "Revealing Clothes",
      "ParentName": "Suggestive"
    }
  ],
  "ModerationModelVersion": "5.0"
}
```



```
"ModerationLabels": [],  
"ModerationModelVersion": "5.0"
```

EXPLICIT NUDITY

Nudity, Sexual Activity, Adult Toys

SUGGESTIVE

Swimwear Or Underwear, Partial Nudity, Revealing Clothes, Sexual Situations

VIOLENCE

Graphic Violence Or Gore, Physical Violence, Weapon Violence, Weapons, Self Injury

VISUALLY DISTURBING

Emaciated Bodies, Corpses, Hanging, Air Crash, Explosions And Blasts

RUDE GESTURES

Middle Finger

DRUGS

Drug Products, Drug Use, Pills, Drug Paraphernalia

TOBACCO

Tobacco Products, Smoking

ALCOHOL

Drinking, Alcoholic Beverages

HATE SYMBOLS

Nazi Party, White Supremacy, Extremist

REFERENCIAS

- <https://docs.aws.amazon.com/rekognition/latest/dg/procedure-moderate-images.html>
- <https://github.com/lean1dev/classificador-imagem-publisher>
- <https://github.com/AugustoSavi/spring-boot-aws-rekognition>
- <https://docs.aws.amazon.com/rekognition/latest/dg/procedure-moderate-images.html>
- <https://docs.aws.amazon.com/rekognition/latest/dg/programming.html>
- <https://docs.aws.amazon.com/rekognition/latest/dg/moderation.html>
- <https://docs.aws.amazon.com/rekognition/latest/dg/procedure-moderate-videos.html>
- <https://docs.aws.amazon.com/rekognition/latest/dg/labels.html>
- https://docs.aws.amazon.com/rekognition/latest/dg/API_DetectModerationLabels.html#API_DetectModerationLabels_SeeAlso
- https://aws.amazon.com/pt/rekognition/?nc2=h_gl_prod_ml_rek
- <https://aws.amazon.com/pt/rekognition/customers/?nc=sn&loc=8>
- <https://aws.amazon.com/pt/rekognition/resources/?nc=sn&loc=6>
- <https://aws.amazon.com/pt/machine-learning/?hp=tile&tile=solutions>
- <https://us-east-2.console.aws.amazon.com/rekognition/home?region=us-east-2#/image-moderation>

DESENVOLVEDORES



Jean1dev



AugustoSavi



FIM...



