

Funções Serverless e AI-APIs

Engineering no Mundo

Grupo Engineering

+ 11.000

Colaboradores

+ 65

Escritórios
no Mundo

Sede Global

Roma (Itália)

O QUE NÓS FAZEMOS?

Apoiamos nossos clientes à transformar os processos de negócio em modelos operacionais alinhados com a Era Digital.

NO MUNDO TODO

Entrega de Projetos

**BASE NA EUROPA
AMÉRICA DO NORTE
AMÉRICA LATINA**

CERTIFICAÇÕES

ISO 9001

Quality Management Systems

ISO 14001

Quality Management Systems

ISO 27001

Information Security Management Systems

ISO 2000

IT Services Management Systems

CMMI – Dev Version 1.3 Level 3

Software Development Systems

NATO AQAP 2110/160

Design, Development and Maintenance of
Application Software for Information
Systems for Military Purposes

CRESCIMENTO CONTÍNUO

USD 1,5 Bilhões

Receita em 2018

+ 39

Anos de crescimento
contínuo

PESQUISA & INOVAÇÃO

USD 50 Milhões

Investimentos

+ 420

Cientistas de Dados
e Pesquisadores

+ 80

Projetos de Pesquisa
em tempo real

160k

Horas de Treinamento
em nossa Academia

+ 200

Innovators



Engineering no Brasil



NOSSA ATUAÇÃO

+ 10 anos

Operação no Brasil

R\$ 203 Milhões

Receita em 2018

Escritórios

São Paulo
Santo André
Belo Horizonte
Rio de Janeiro
Buenos Aires

+ 650

Colaboradores

+ 150

Grandes Cientes

CERTIFICAÇÕES

CMMI-DEV 1.3 Nível 3
DP-FISC SAP TDF HANA
DP-FISC SAP ECC HANA
DP-FISC SAP S/4 HANA
SAP GOLD PARTNER

Somos uma companhia Global de Tecnologia da Informação e Consultoria especializada em Transformação Digital.

Apoiamos nossos clientes à transformar os processos de negócio em modelos operacionais alinhados com a Era Digital.

Agenda

FaaS – Um novo serviço cloud universal?

Virtualização, Containers, e Cloud World

IaaS, PaaS, SaaS e Noix

Just Code!

Nem tudo é festa

AI-APIs – Inteligência via rest

Inteligência on-demand!

Por que eles?

Cores e Sabores

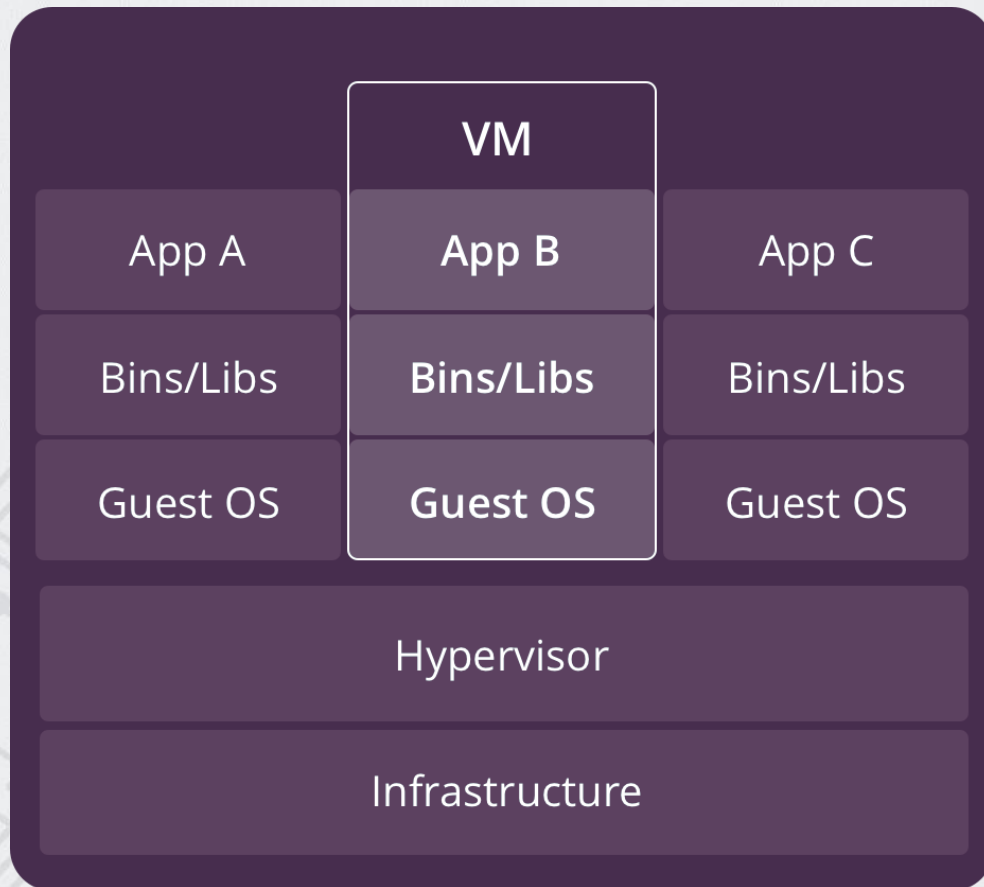
Just Solution!

Aplicação de Inteligência Artificial FaaS em 15 minutos!!!



Virtualização, Containers, e Cloud World

Virtualização



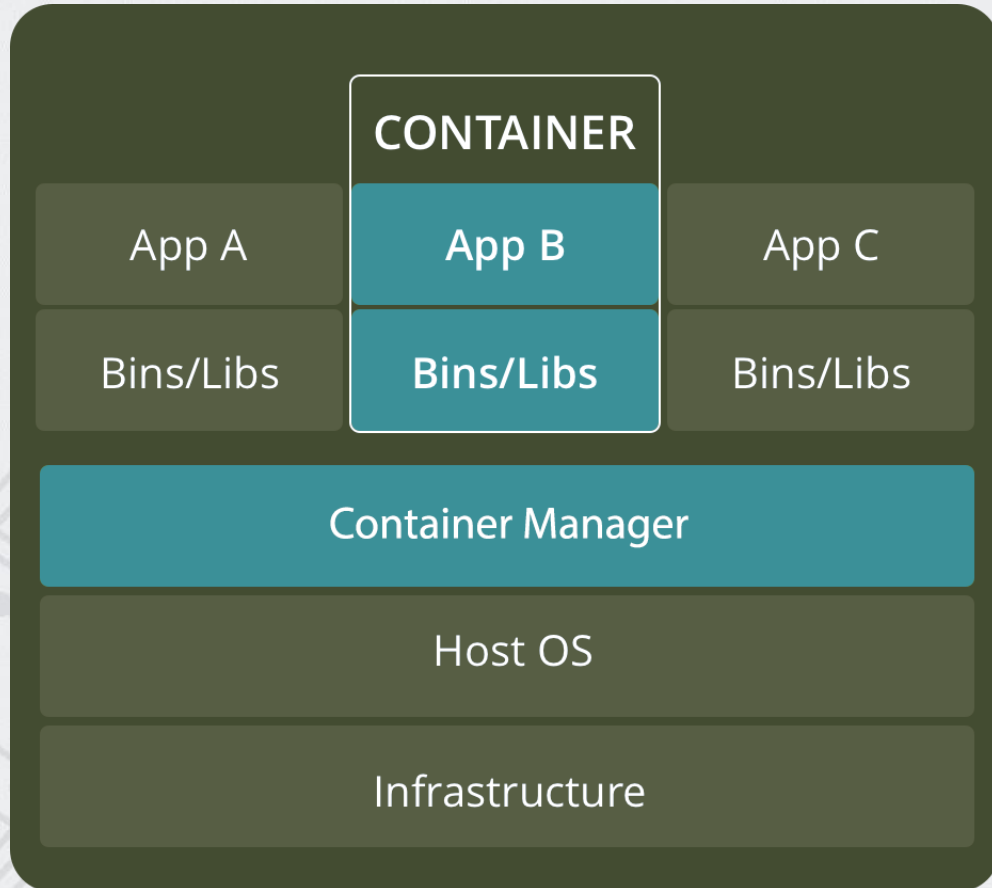
A virtualização de máquina consiste em empacotar determinada aplicação em um modelo padrão com seu próprio ambiente operacional. Uma máquina virtual (VM) é uma emulação de um sistema de computador. Cada VM requer seu próprio SO subjacente e o hardware é virtualizado.

A partir da virtualização, não importa a “maquina” onde o aplicação ira “rodar”, suas configurações serão as mesmas desde o nível de SO até as bibliotecas e dependências necessárias



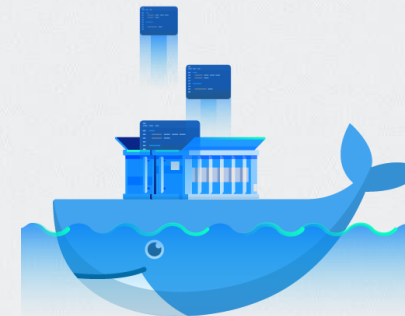
Virtualização, Containers, e Cloud World

Containers

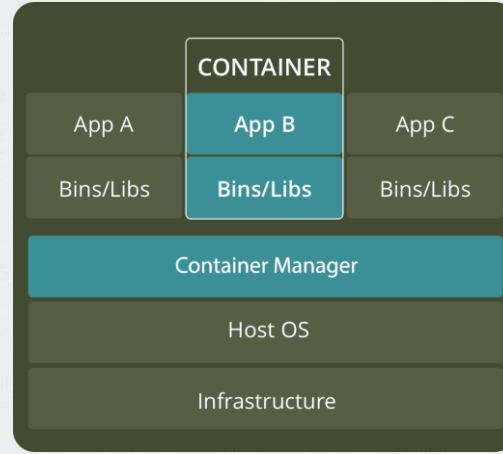
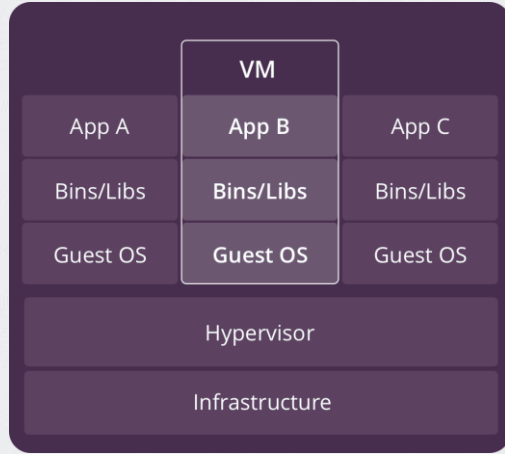


O container trabalha em cima da aplicação e suas dependências criando uma virtualização apenas a nível do SO e não do hardware.

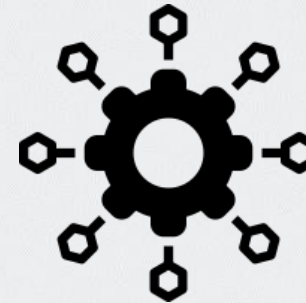
Os contêineres ficam em cima de um servidor físico e de seu SO host - geralmente Linux ou Windows. Cada contêiner compartilha o kernel do sistema operacional host e, geralmente, os binários e bibliotecas também.



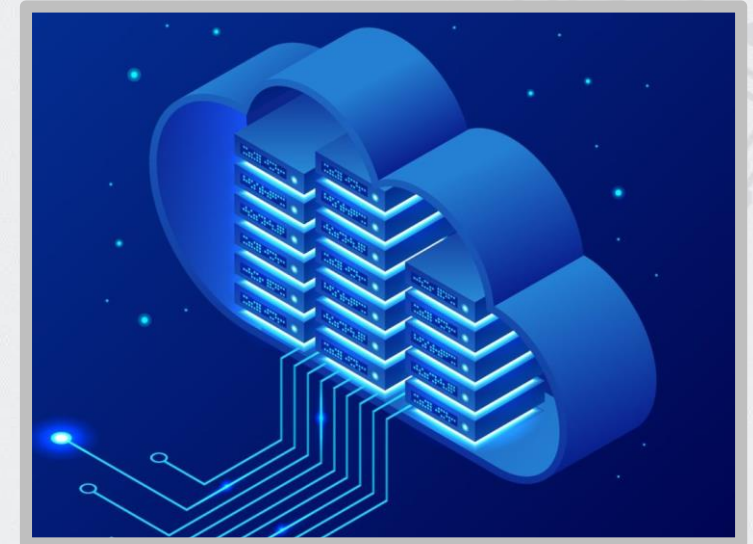
Virtualização, Containers, e Cloud World



kubernetes



MICROSERVIÇOS



Pesado	Leve
Desempenho limitado	Desempenho nativo
Cada VM é executada em seu próprio SO	Todos os contêineres compartilham o SO host
Virtualização em nível de hardware	Virtualização de SO
Tempo de inicialização em minutos	Tempo de inicialização em milissegundos
Aloca a memória necessária	Requer menos espaço de memória
Totalmente isolado e, portanto, mais seguro	Isolamento no nível do processo, possivelmente menos seguro

“Um modelo que permite o acesso onipresente, conveniente e sob demanda da rede a um conjunto compartilhado de recursos de computação configuráveis”.



IaaS, PaaS e SaaS

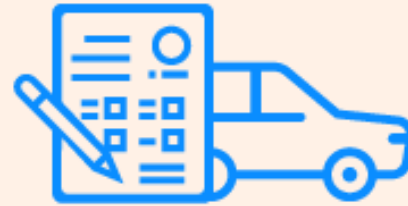
YOUR OWN CAR

On-premises solution



LEASED CAR

IaaS



TAXI

PaaS



BUS

SaaS



Your Own Car: Seu carro, sua responsabilidade. Manter e atualizar.

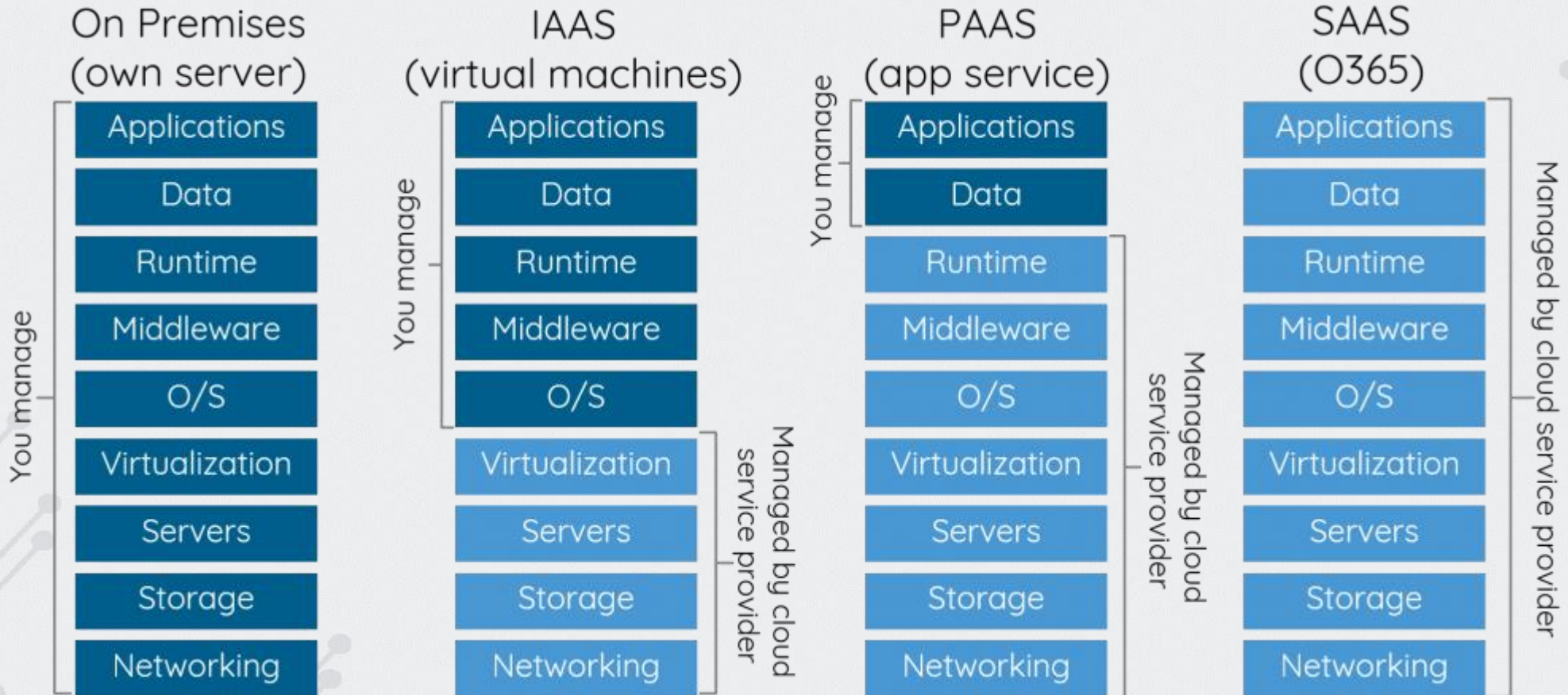
Leased Car: Você dirige, personaliza e controla. Mas o carro não é seu. Você paga por usar um bem de alguém.

Taxi: Você não dirige, mas diz a ele onde quer chegar, e ele te leva ao seu destino. Menos autonomia, mais agilidade.

Bus: Você não escolhe um caminho distinto, mas trafega pelas rotas existentes. Menos esforço, menos autonomia, muito mais agilidade e compartilhamento.



IaaS, PaaS e SaaS



SaaS

PaaS

IaaS



Apps/aplicativos
hospedados



Ferramentas de
desenvolvimento,
gerenciamento de
banco de dados,
análise de negócios



Sistemas operacionais



Servidores e
armazenamento

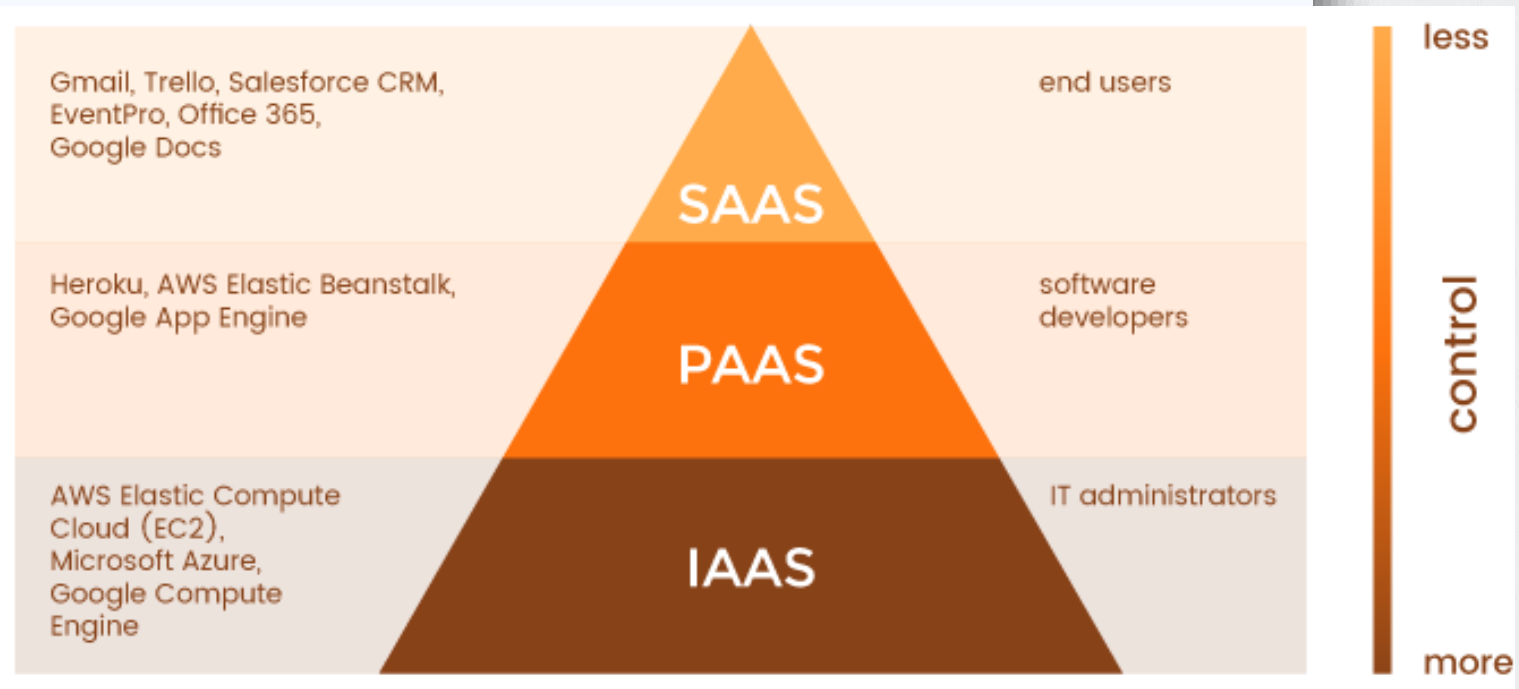


Segurança/firewalls
de rede



Construção/planta
física do datacenter

[S,P,I]aaS



IaaS, PaaS, SaaS e Noix

"Em 30 anos em TI, testemunhei inúmeras tentativas de fornecedores nas últimas décadas para acelerar o processo de desenvolvimento. Passamos do código de máquina para o assembler e idiomas como Cobol, Fortran e RPG nos primeiros dias do mainframe. Cada idioma foi uma tentativa de abstrair as complexidades do idioma anterior para permitir que os desenvolvedores concluíssem as tarefas mais rapidamente."

Mike Kavis

Managing Director at Deloitte Consulting

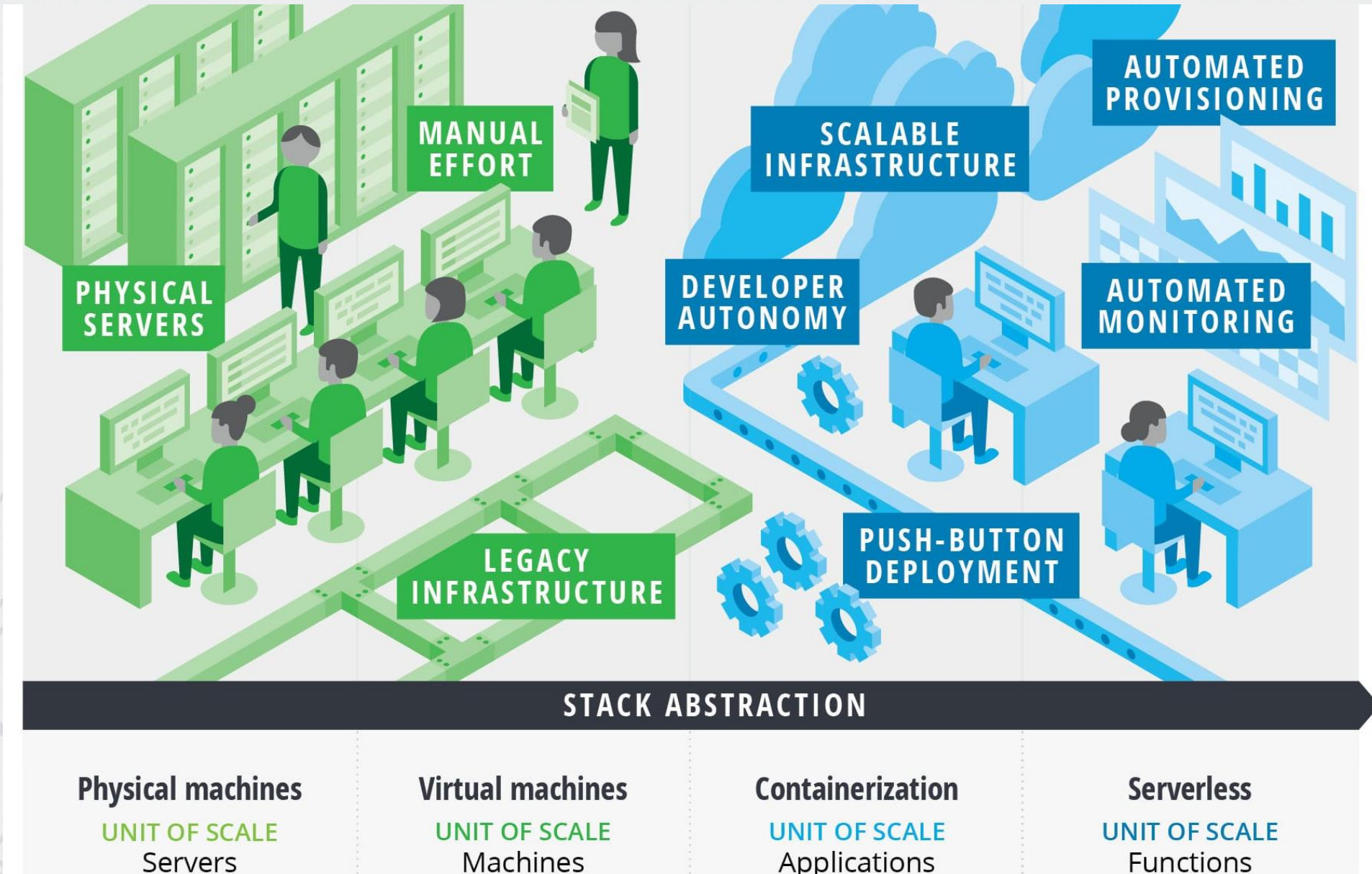
**ABSTRAIR A
COMPLEXIDADE
DE OPERAÇÕES**

**ACELERAR O
DESENVOLVIMENTO**



Going to the core of business

NoOps

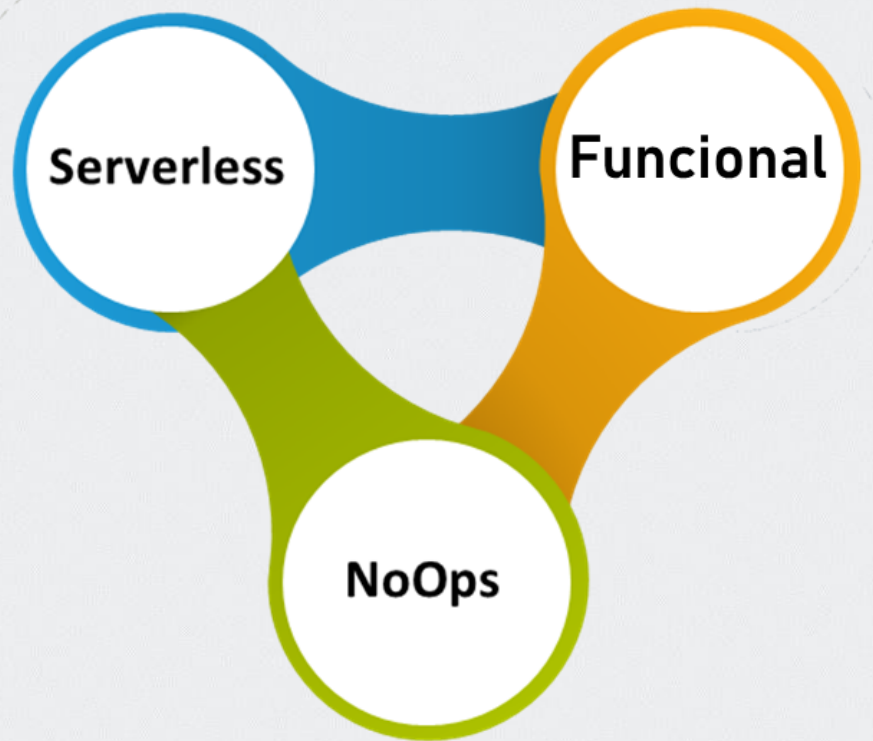


Serverless



FaaS – Function as a Service

Modelo de computação em nuvem que visa abstrair do programador o gerenciamento de servidores, infraestrutura de baixo nível e a alocação de recursos, os quais são gerenciados pelo provedor de nuvem.



Paradigma de programação baseado nas funções matemáticas onde o valor de saída depende apenas dos argumentos de entrada. Expressões (Métodos e funções) como objetos sem dependência externa.

“Adoção generalizada de uma estratégia de desenvolvimento de software onde os desenvolvedores não dependem da área de operações de TI.”

<https://www.infoq.com/br/news/2012/03/NoOps/>



FaaS – Vendors



Node.js, Java, Python, .NET, Go.

Com adaptadores construídos
pode-se usar quase todas as
linguagens



Node.js, Python, e Go.



Com o Cloud Run, podemos
usar quase todas as linguagens



.NET (C #, F #), Node.js, Java.
E está experimentando
Python, PHP, TypeScript,
Batch, Bash e Powershell.

FaaS – Use cases



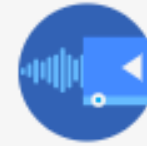
Back-ends de aplicativos sem servidor

- Integração com serviços e APIs de terceiros
- Back-ends para dispositivos móveis
- Back-ends para Internet das Coisas (IoT, na sigla em inglês)



Processamento de dados em tempo real

- Processamento de arquivos
- Processamento de stream
- Extração, transformação e carregamento (ETL) orientados a eventos.



Aplicativos inteligentes

- Assistentes virtuais e chatbots
- Análise de vídeo e imagem
- Análise de sentimento



FaaS – Just code!



Quantidade de memória para execução da function

MEMORY

Tempo de existência da instância da function

TEMPO DE PERSISTÊNCIA

Trigger

DEPENDENCY

MAIN

Acionadores:
APIs
Triggers
nativas

Dependência dos pacotes da aplicação.

Function que recebe a requisição rest

Funções do Google Cloud | < Criar função

Memória alocada
512 MB ← Memory

Acionador
HTTP ← Trigger

URL
https://us-central1-digital-integration.cloudfunctions.net/function-4 ← Host

Autenticação
☒ Permitir invocações não autenticadas
Marque se você estiver criando uma API ou um site público.

Tempo de execução
Python 3.7 ← Linguagem

main.py requirements.txt ← Dependência

Função a ser executada
hello_world ← Main

Opções avançadas
Região
us-central1

Tempo limite
60 segundos ← Tempo de existência da Instância



FaaS – Nem tudo é festa

Gerenciamento

Muitas funções implementadas, pode ser difícil gerencia-las.

Tempos de inicialização

Funções nem sempre estão ativadas, às vezes o sistema fica mais lento para responder porque precisa instanciar a função.

STATELESS

Tempos de execução

São proprietários dos provedores. AWS, GCP e Azure têm sua própria implementação interna separada dos tempos de execução e ninguém sabe realmente o que está fazendo.

Potencialmente difícil de depurar.

Em muitos casos a depuração precisa ser feita externamente.



Inteligência Artificial

“Intelligence that is not biological”

Dr. Tegmark is a professor at the Massachusetts Institute of Technology and the scientific director of the Foundational Questions Institute.

MEMORANDUM
RM-3731-CC
JUNE 1963

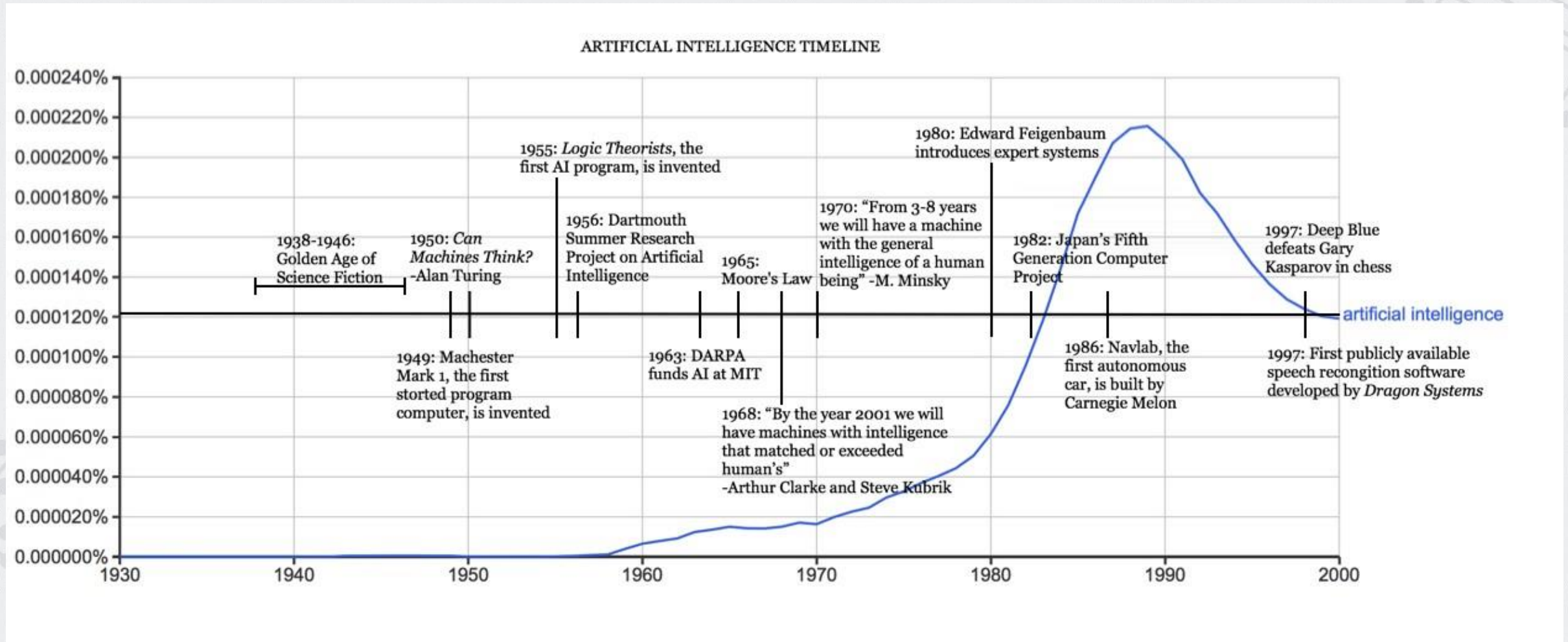
**THE LOGIC THEORY MACHINE:
A MODEL HEURISTIC PROGRAM**

Einar Stefferud

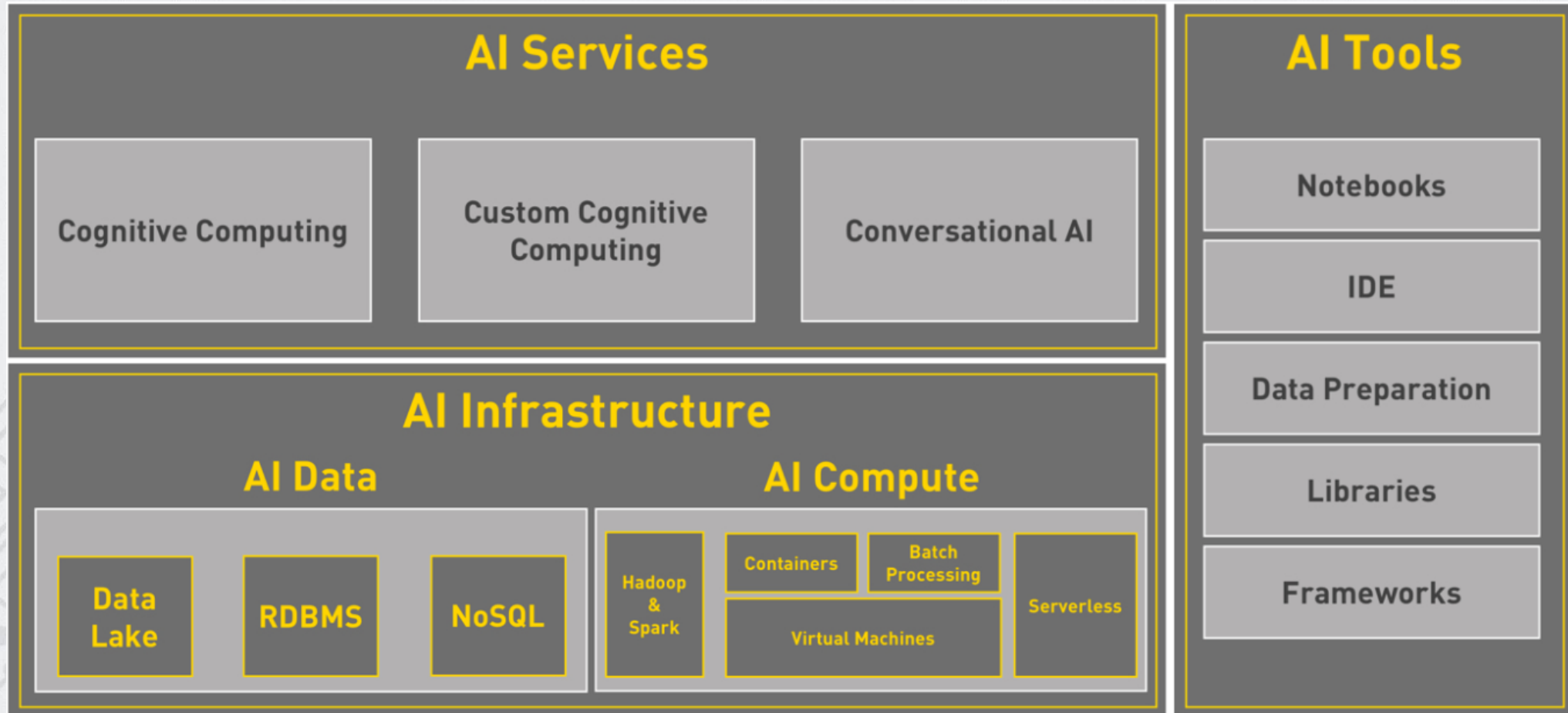
“No verão de 1956, John McCarthy , Marvin Minsky e Claude Shannon organizaram uma conferência no Dartmouth College sobre o que chamavam de "inteligência artificial" (termo cunhado por McCarthy para a ocasião). Simon e Newell orgulhosamente apresentaram ao grupo seu Teórico da Lógica”



Inteligência Artificial



AI APIs - Inteligência on-demand!

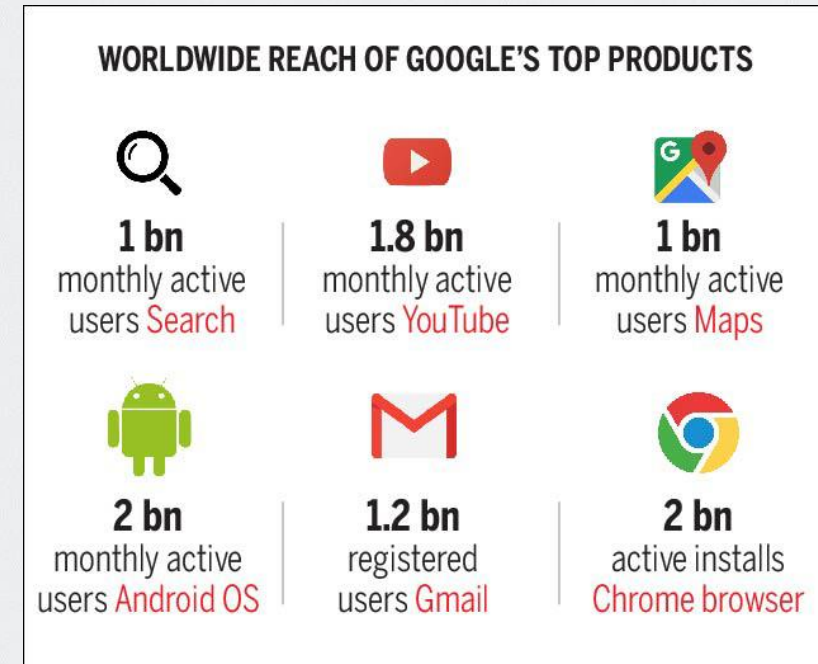


AI APIs – Por que Eles?

“Os dois pilares críticos da inteligência artificial e do aprendizado de máquina são dados e computação. O aprendizado profundo e as redes neurais - técnicas avançadas de aprendizado de máquina - executam cálculos complexos que exigem uma combinação de CPUs, GPUs e até TPUs.”



ESCALA



DADOS



AI APIs – Por que Eles?

“APIs de computação cognitiva, como visão, fala, tradução, análise de texto e pesquisa, estão disponíveis como pontos de extremidade REST para os desenvolvedores”

“Para permitir que os clientes aproveitem os benefícios da computação cognitiva com base em conjuntos de dados personalizados, os fornecedores de nuvem estão adotando a computação cognitiva personalizada. Nesse modelo, os clientes trazem seus próprios dados para treinar serviços cognitivos para fornecer serviços especializados e de nicho. Essa abordagem elimina o ônus de escolher os algoritmos certos e treinar modelos personalizados.”



Cloud AutoML

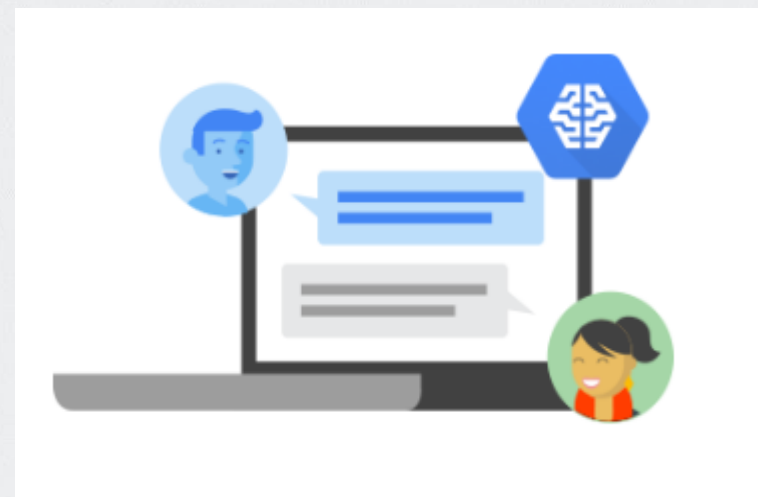
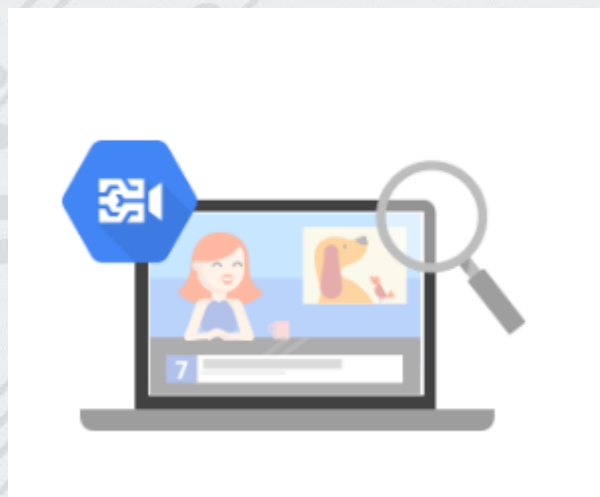
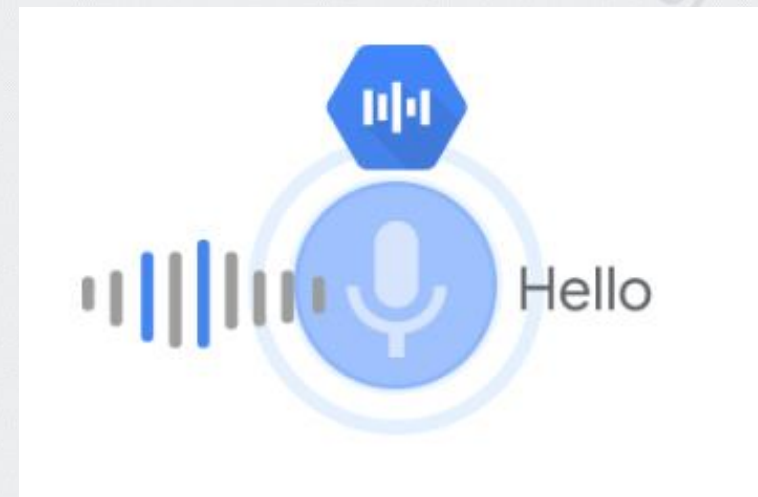
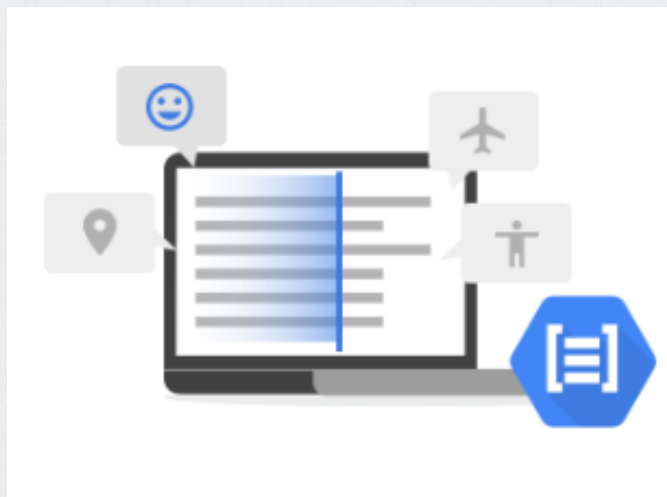
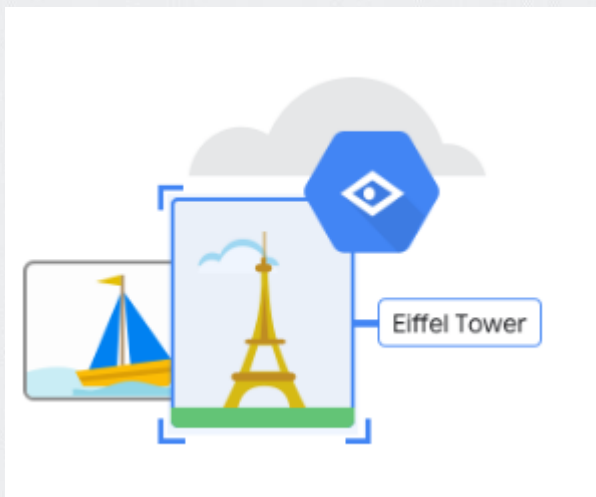
Treine modelos ML personalizados de maneira rápida e fácil

O Cloud AutoML é um conjunto de produtos de aprendizado de máquina que permite que desenvolvedores com conhecimento limitado em ML treinem modelos de alta qualidade específicos para suas necessidades. O Cloud AutoML aproveita

mais de dez anos da tecnologia proprietária do Google Research para ajudar seus modelos de aprendizado de máquina a obter desempenho mais rápido e previsões mais precisas.



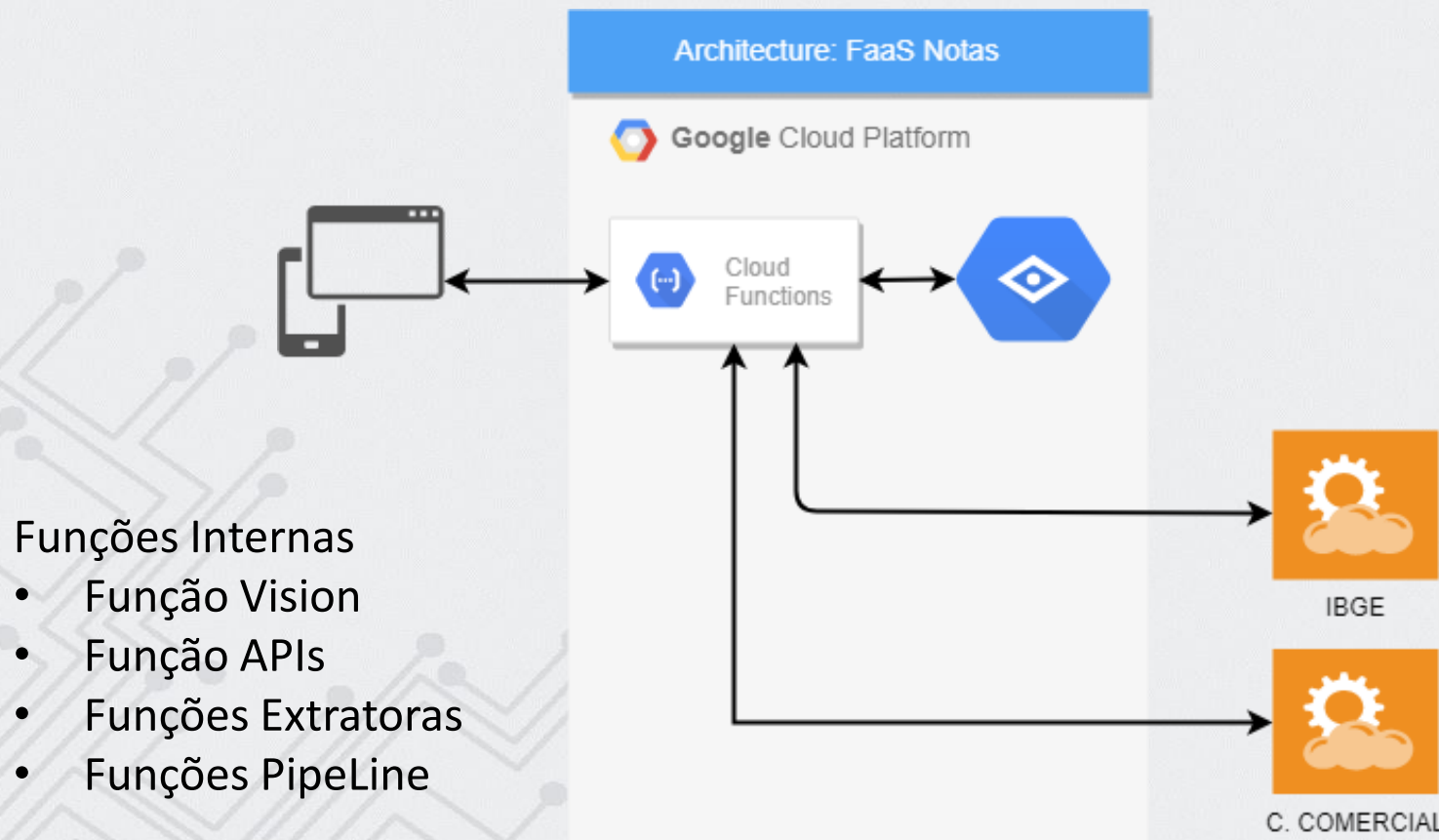
AI APIs – Cores e Sabores



AI APIs – Just Solution!

Uma solução FaaS para transformação de recibos em papel, em dados observáveis às aplicações, identificação do emissor do recibo, do tipo de estabelecimento, valor e outros atributos.

Nosso exemplo serão recibos de estacionamento.



Componentes

- GCP Cloud Storage (Bucket para as fotos)
- GCP Cloud Functions (250MB, 60 s, Python)
- GCP Vision API (document_text_detection)
- Chamada Rest API IBGE
- Chamada Rest API receiptaws



AI APIs e FaaS em 15 minutos – 1



- Criar uma conta no GCP ou entrar com a sua <https://console.cloud.google.com>.
- Ir até “Intervalos”
- Criar um intervalo “meet_up” e uma pasta “fotos” no Cloud Storage e subir as fotos nesta pasta, como abaixo.

meet_up

[Objetos](#) [Visão geral](#) [Permissões](#) [Bloqueio do intervalo](#)

[Enviar arquivos](#) [Enviar pasta](#) [Criar pasta](#) [Gerenciar retenções](#) [Excluir](#)

[Intervalos](#) / [meet_up](#) / fotos

<input type="checkbox"/>	Nome	Tamanho	Tipo	Classe de armazenamento	Última modificação	Acesso público ?	Criptografia ?
<input type="checkbox"/>	 estacionamento.jpeg	1,87 MB	image/jpeg	Standard	10/09/2019, 18:21:21 UTC-3	Não público	Chave gerenciada pelo Google
<input type="checkbox"/>	 testeRASautomation.jpeg	2,83 MB	image/jpeg	Standard	10/09/2019, 15:08:40 UTC-3	Não público	Chave gerenciada pelo Google



AI APIs e FaaS em 15 minutos - 2

- Ir até Cloud Functions. Clicar em Criar Função e alterar:
 - Nome da função
 - Runtime: python 3.7
 - Manter os demais padrões como estão

The screenshot shows the 'Create Function' form in the Google Cloud console. A red arrow points from the text 'No exemplo usamos o nome meet_up' to the 'Nome' field, which contains 'nome_da_funcao'. Another red arrow points from the text 'Altere para Python 3.7' to the 'Tempo de execução' field, which contains 'Python 3.7'. The form includes sections for 'Memória alocada' (256 MB), 'Acionador' (HTTP), 'URL', 'Autenticação' (checked), and 'Código-fonte' (Editor in-line selected).

Nome ?
nome_da_funcao

Memória alocada
256 MB

Acionador
HTTP

URL
https://us-central1-digital-integration.cloudfunctions.net/nome_da_funcao

Autenticação
☒ Permitir invocações não autenticadas
Marque se você estiver criando uma API ou um site público.

Atalho para atribuir o papel de invocador do IAM ao identificador especial allUsers.
Use o IAM para editar essa configuração depois de criar a função.

Código-fonte
☒ Editor in-line
☐ Upload do ZIP
☐ ZIP do Cloud Storage
☐ Repositório do Cloud Source

Tempo de execução
Python 3.7

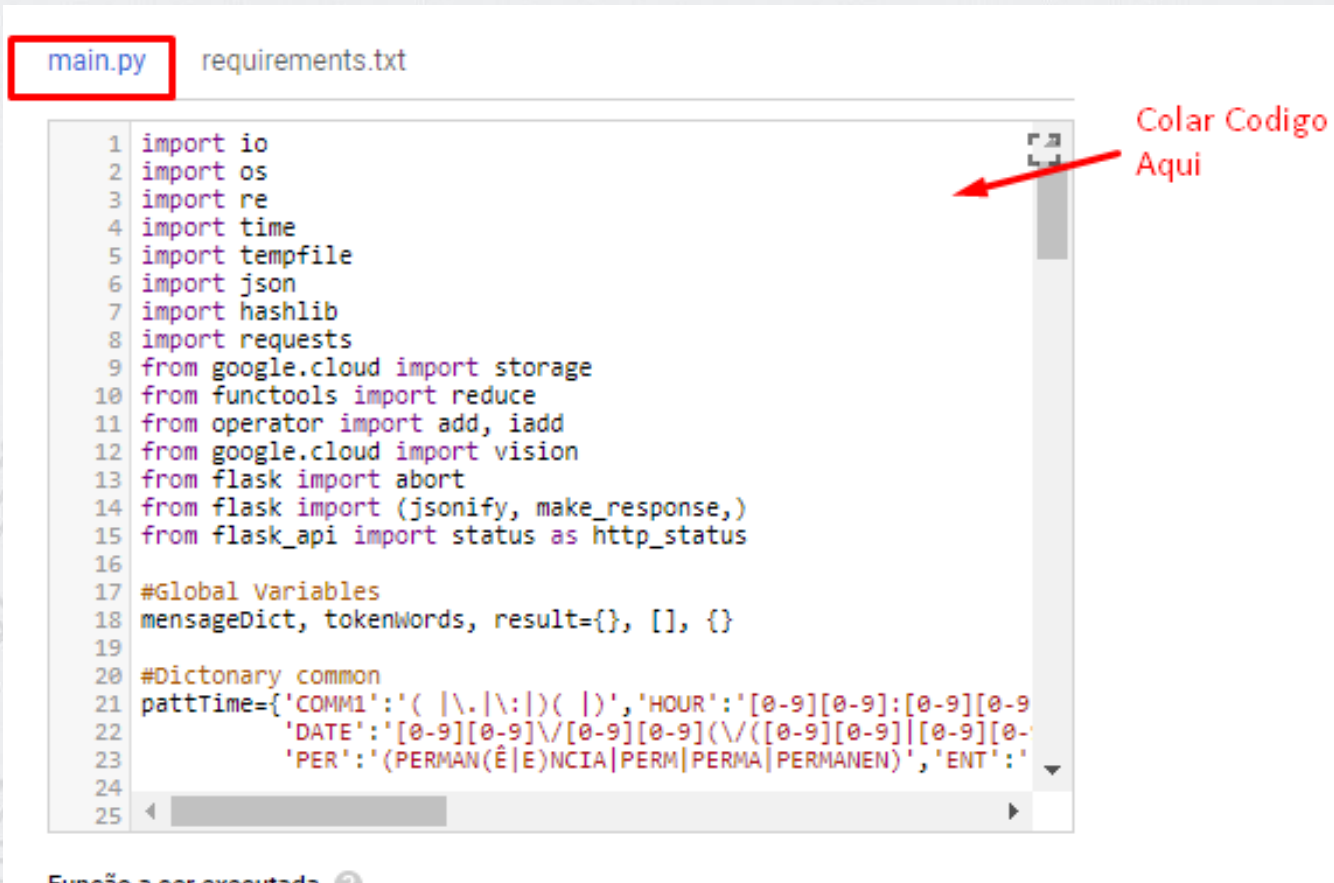
No exemplo usamos o nome meet_up

Altere para Python 3.7



AI APIs e FaaS em 15 minutos - 3

- Baixe o código da função no github git@github.com:THIAGONOMA/meet_up.git, o código está no arquivo meetupFaas.py



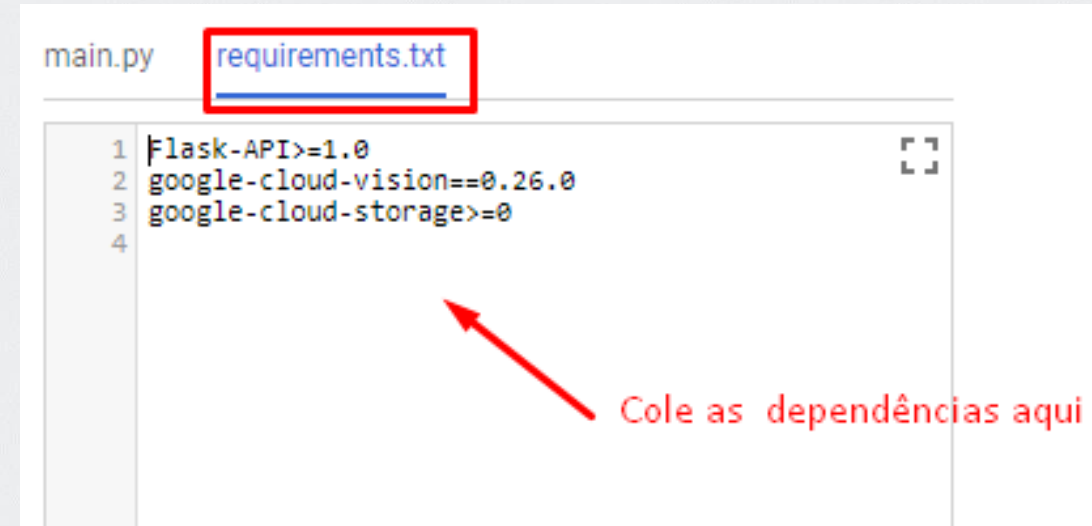
```
1 import io
2 import os
3 import re
4 import time
5 import tempfile
6 import json
7 import hashlib
8 import requests
9 from google.cloud import storage
10 from functools import reduce
11 from operator import add, iadd
12 from google.cloud import vision
13 from flask import abort
14 from flask import (jsonify, make_response,)
15 from flask_api import status as http_status
16
17 #Global Variables
18 messageDict, tokenWords, result={}, [], {}
19
20 #Dictionary common
21 pattTime={'COMM1': '( |\.\:|)( |)', 'HOUR': '[0-9][0-9]:[0-9][0-9]
22           'DATE': '[0-9][0-9]\/[0-9][0-9](\/|([0-9][0-9]|([0-9][0-9]
23           'PER': '(PERMAN(Ê|E)NCIA|PERM|PERMA|PERMANEN)', 'ENT': '
24
25
```

main.py requirements.txt

Colar Código Aqui

Função a ser executada

- No mesmo GIT as dependências necessárias estão no arquivo requirements.txt



```
1 Flask-API>=1.0
2 google-cloud-vision==0.26.0
3 google-cloud-storage>=0
4
```

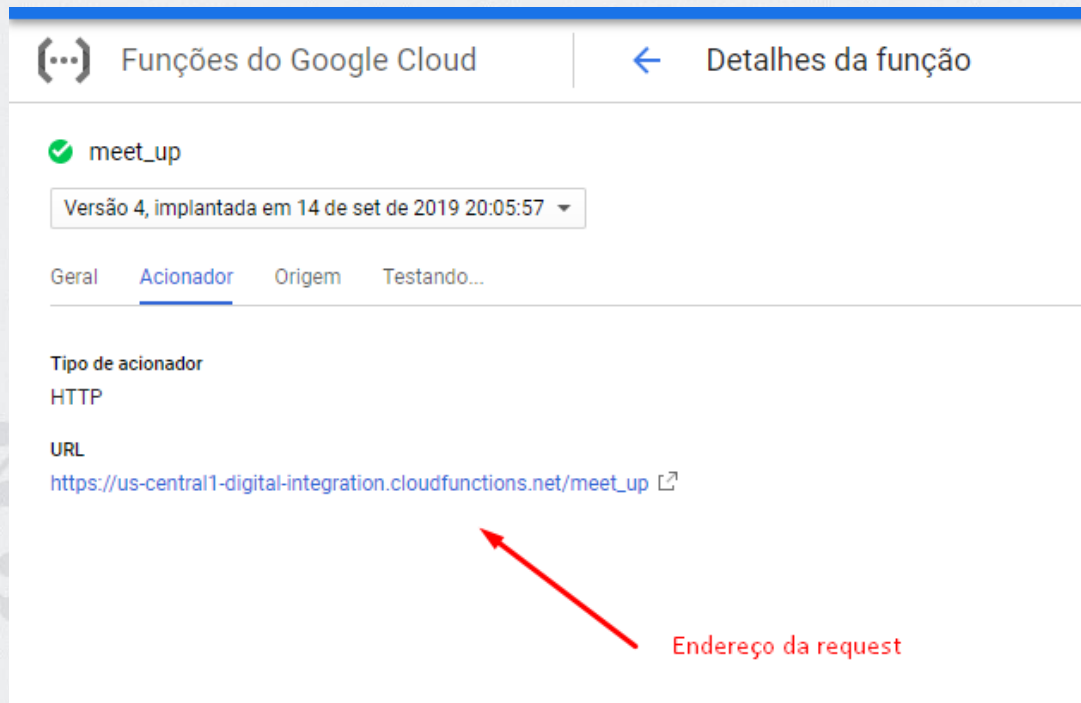
main.py requirements.txt

Cole as dependências aqui



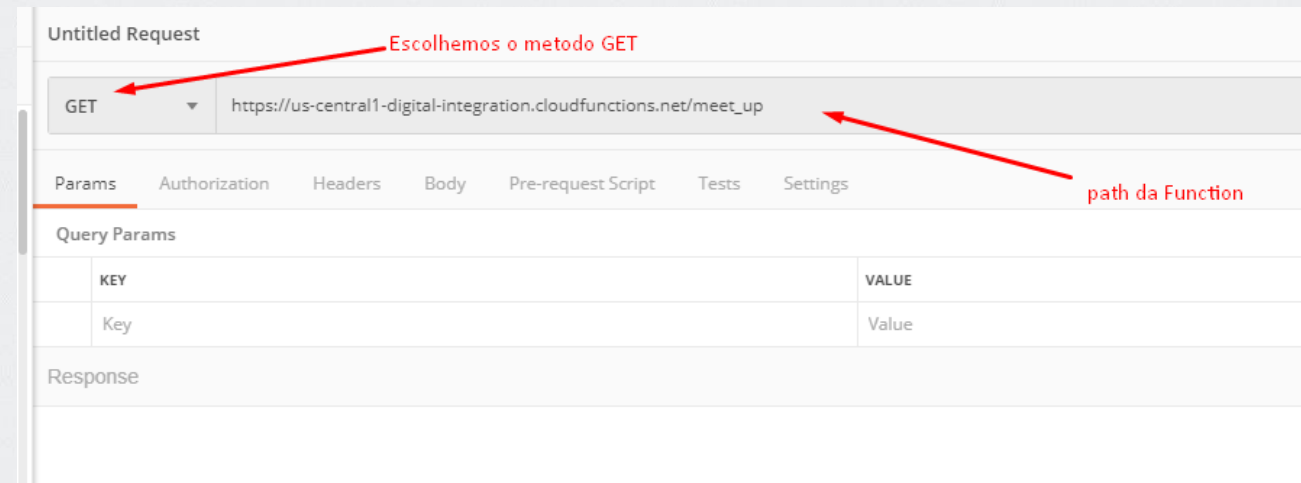
AI APIs e FaaS em 15 minutos – Testando a Function

- Copie o endereço do caminho da função e cole em uma app para testar APIs, como o Postman por exemplo.



The screenshot shows the Google Cloud Functions console. At the top, there's a breadcrumb navigation: 'Funções do Google Cloud' > 'Detalhes da função'. Below this, the function 'meet_up' is listed with a green checkmark. A dropdown menu shows 'Versão 4, implantada em 14 de set de 2019 20:05:57'. There are four tabs: 'Geral', 'Acionador' (selected), 'Origem', and 'Testando...'. Under the 'Acionador' tab, it shows 'Tipo de acionador: HTTP' and 'URL: https://us-central1-digital-integration.cloudfunctions.net/meet_up'. A red arrow points from the text 'Endereço da request' to the URL.

- Escolha o método GET e cole o path da Function.



The screenshot shows the Postman 'Untitled Request' editor. The 'GET' method is selected from a dropdown menu, with a red arrow pointing to it from the text 'Escolhemos o metodo GET'. The URL 'https://us-central1-digital-integration.cloudfunctions.net/meet_up' is entered in the address bar, with a red arrow pointing to it from the text 'path da Function'. Below the URL bar, there are tabs for 'Params', 'Authorization', 'Headers', 'Body', 'Pre-request Script', 'Tests', and 'Settings'. The 'Params' tab is selected, showing a table for 'Query Params' with columns 'KEY' and 'VALUE'. The table has one row with 'Key' and 'Value'. Below this is a 'Response' section.

KEY	VALUE
Key	Value



AI APIs e FaaS em 15 minutos – Testando a Function

- Podemos ter dois parâmetros “path” e “tokenwords”.
- “path” endereço da imagem no Cloud Storage. (obrigatório)

GET https://us-central1-digital-integration.cloudfunctions.net/meet_up?path=gs://meet_up/fotos/estacionamento.jpeg Send

Params Authorization Headers (7) Body Pre-request Script Tests Settings Cookies Cod

Query Params

KEY	VALUE	DESCRIPTION
<input checked="" type="checkbox"/> path	gs://meet_up/fotos/estacionamento.jpeg	
Key	Value	Description

Body Cookies Headers (7) Test Results

Status: 200 OK Time: 12.08s Size: 997 B Sa

Pretty Raw Preview Visualize BETA JSON

```
1 {
2   "message": {
3     "DATATRANS": "09/09/19: 15:54",
4     "ENT": "09/09 12:54",
5     "ESTAB": {
6       "BAIRRO": "BROOKLIN NOVO",
7       "CEP": "04.571-020",
8       "CNPJ": "07.097.831/0002-53",
9       "COMPLE": "",
10      "FANTAS": "TECNO GARAGE E ESTACIONAMENTOS",
11      "LOGRA": "RUA DR. GERALDO DE CAMPOS MOREIRA",
12      "MUN": "SAO PAULO",
13      "NAME": "TECNO GARAGE E ESTACIONAMENTOS EIRELI",
14      "NUM": "375",
15      "OBSACT": [
16        "Esta classe compreende - a exploração de edifícios-garagem e parques de estacionamento para veículos, por curta duração",
17        "Esta classe NÃO compreende - a lavagem, lubrificação e polimento de veículos automotores (45.20-0)"
18      ],
19      "TYPEACT": "ESTACIONAMENTO DE VEÍCULOS",
20      "UF": "SP"
21    },
22    "PERM": "03:00",
23    "SAI": null,
24    "VALOR": "20,00"
25  },
26   "status": "OK"
27 }
```

nome do argumento

valor do argumento: sua imagem

Note o primeiro tempo de execução. Que é na verdade o tempo de inicialização

Retorno: objeto JSON com 2 atributos:

"message": resultado efetivo da nossa análise.

"Status": Estado do processamento interno da function: se execução sem nenhum problema, "OK" senão "ERRO" + Log de exception



AI APIs e FaaS em 15 minutos – Testando a Function

- Delta entre o primeiro tempo de execução e os demais

The image displays two screenshots of the Google Cloud Functions test interface, illustrating the difference in execution time between the first and subsequent calls.

Top Screenshot (First Execution):

- Method: GET
- URL: `https://us-central1-digital-integration.cloudfunctions.net/meet_up?path=gs://meet_up/fotos/estacionamento.jpeg`
- Query Params:

KEY	VALUE	DESCRIPTION
<input checked="" type="checkbox"/> path	<code>gs://meet_up/fotos/estacionamento.jpeg</code>	
Key	Value	Description

Annotations: "nome do argumento" points to the 'path' key; "valor do argumento: sua imagem" points to the 'gs://meet_up/fotos/estacionamento.jpeg' value.

- Status: 200 OK, Time: 12.06s, Size: 997 B
- Body (JSON):

```
1 {
2   "message": {
3     "DATATRANS": "09/09/19: 15:54",
4     "ENT": "09/09 12:54",
5     "ESTAB": {
6       "BAIRRO": "BROOKLIN NOVO",
7       "CEP": "04.571-020",

```

Annotation: "Note o primeiro tempo de execução. Que é na verdade o tempo de inicialização" points to the 12.06s time.

Bottom Screenshot (Subsequent Execution):

- Method: GET
- URL: `https://us-central1-digital-integration.cloudfunctions.net/meet_up?path=gs://meet_up/fotos/estacionamento.jpeg`
- Query Params:

KEY	VALUE	DESCRIPTION
<input checked="" type="checkbox"/> path	<code>gs://meet_up/fotos/estacionamento.jpeg</code>	
Key	Value	Description

- Status: 200 OK, Time: 2.84s, Size: 997 B
- Body (JSON):

```
1 {
2   "message": {

```

Annotation: "note que a segunda chamada demorou quase 10s a menos que a primeira" points to the 2.84s time.

A large blue arrow points from the 12.06s time in the top screenshot to the 2.84s time in the bottom screenshot, highlighting the significant reduction in execution time.



AI APIs e FaaS em 15 minutos – Testando a Function

- “tokenwords” token das palavras contida no texto. (opcional)

GET https://us-central1-digital-integration.cloudfunctions.net/meet_up?path=gs://meet_up/fotos/estacionamento.jpeg&tokenwords=1

Params Authorization Headers (7) Body Pre-request Script Tests Settings

Query Params

KEY	VALUE	DESCRIPTION
<input checked="" type="checkbox"/> path	gs://meet_up/fotos/estacionamento.jpeg	
<input checked="" type="checkbox"/> tokenwords	1	
Key	Value	Description

Body Cookies Headers (7) Test Results Status: 200 OK

Pretty Raw Preview Visualize BETA JSON

```
1 {
2   "message": {
3     "DATATRANS": "09/09/19: 15:54",
4     "ENT": "09/09 12:54",
5     "ESTAB": {
6       "BAIRRO": "BROOKLIN NOVO",
7       "CEP": "04.571-020",
8       "CNPJ": "07.097.831/0002-53",
9       "COMPLE": "",
10      "FANTAS": "TECNO GARAGE E ESTACIONAMENTOS",
11      "LOGRA": "RUA DR. GERALDO DE CAMPOS MOREIRA",
12      "MUN": "SAO PAULO",
13      "NAME": "TECNO GARAGE E ESTACIONAMENTOS EIRELI",
14      "NUM": "375",
15      "OBSACT": [
16        "Esta classe compreende - a exploração de edifícios-garagem e parques de estacionamento para veículos, por curta duração",
17        "Esta classe NÃO compreende - a lavagem, lubrificação e polimento de veículos automotores (45.20-0)"
18      ],
19      "TYPEACT": "ESTACIONAMENTO DE VEÍCULOS",
20      "UF": "SP"
21    },
22    "PERM": "03:00",
23    "SAI": null,
24    "VALOR": "20,00",
25    "tokenwords": [
26      "d43aa582a4ae20906333dbab629d1409",
27      1,
28      0,
29      "E",
30      [
31        "(2885,1171)",
32        "(2886,1181)",
33        "(2765,1192)",
34        "(2764,1182)"
35      ]
36    ]
37  }
38 }
```

Novo objeto de retorno, agora com 3 atributos: "message" e "status" que são os atributos de retorno padrão, mais o atributo tokenwords, contendo uma tokenização das palavras extraídas do ocr da imagem.

- Array “tokenwords”

```
"tokenwords": [
  [
    "d43aa582a4ae20906333dbab629d1409",
    1,
    0,
    "E",
    [
      "(2885,1171)",
      "(2886,1181)",
      "(2765,1192)",
      "(2764,1182)"
    ]
  ]
]
```

hash única da requisição

Numero do documento

Numero da palavra em ordem no texto

Palavra

Posição geografica da palavra no texto



A decorative background featuring a light gray circuit board pattern with various lines and dots, primarily visible in the corners of the slide.

Obrigado !!!



SEU PARCEIRO NA TRANSFORMAÇÃO DIGITAL