VINICARIDÁ
Human-Centered Data & AI

Desbloqueando todo o potencial humano com o poder da

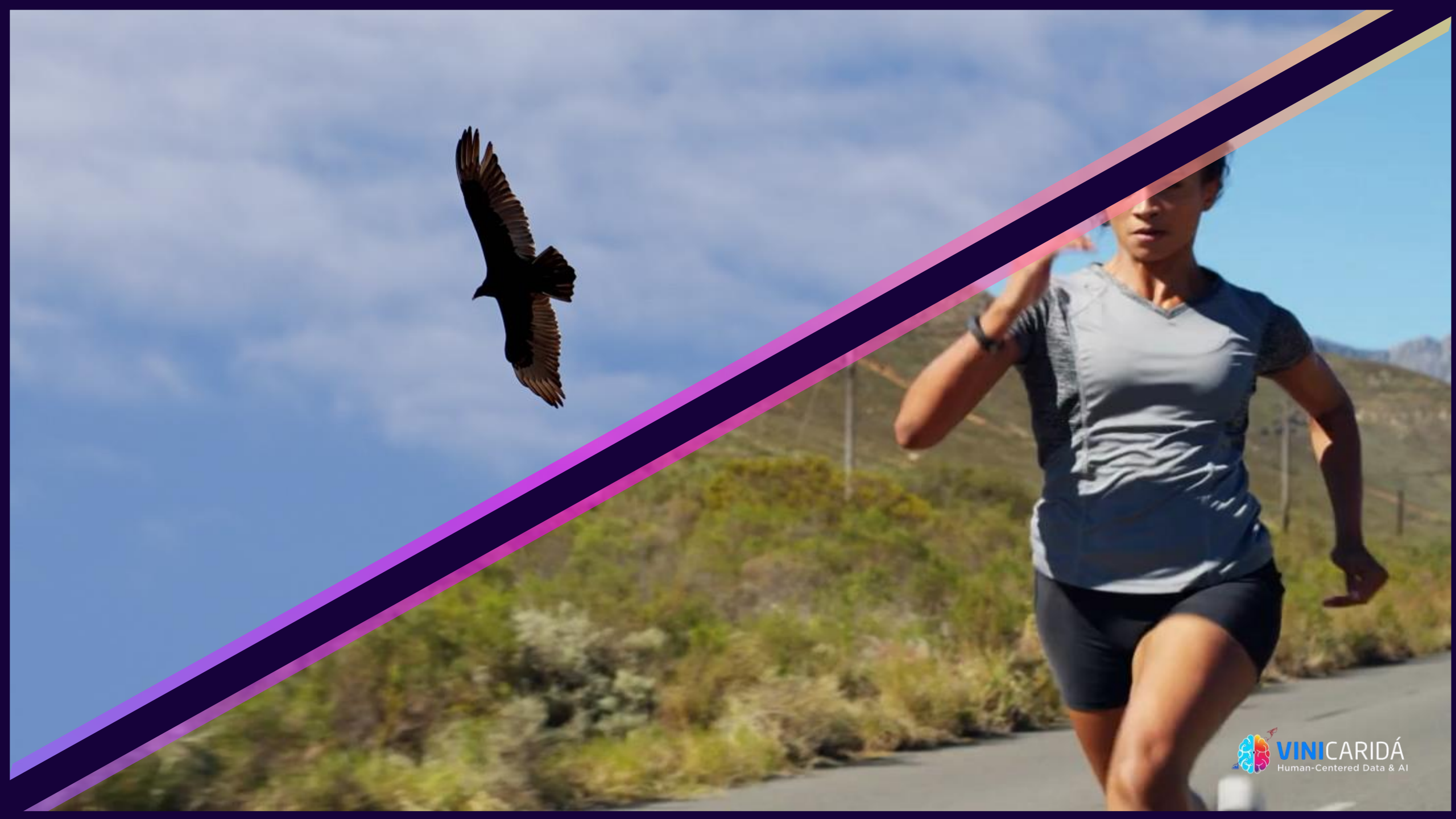
IA generativa

VINICARIDÁ
Human-Centered Data & AI

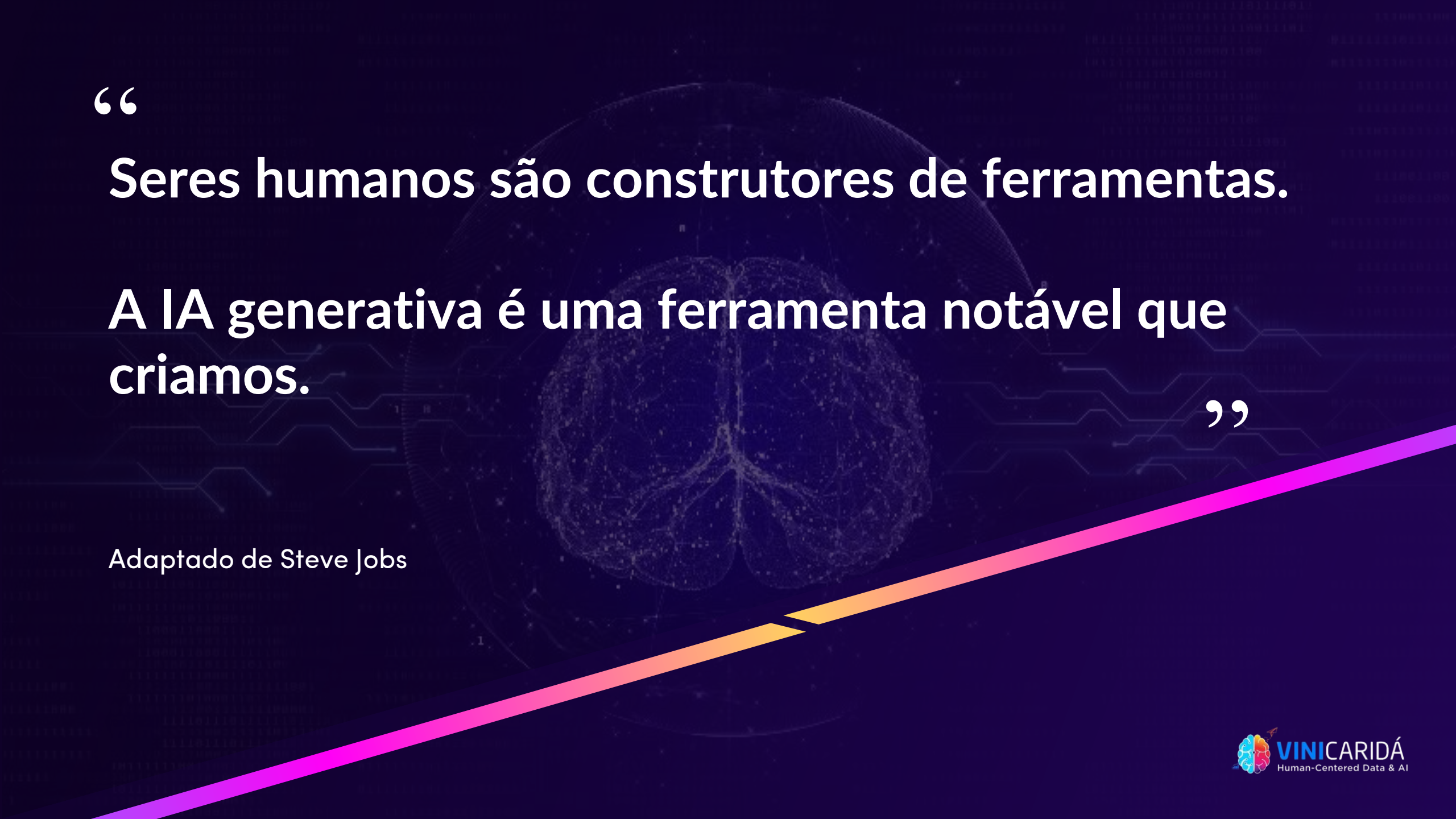
Vinicius Caridá, Ph.D.

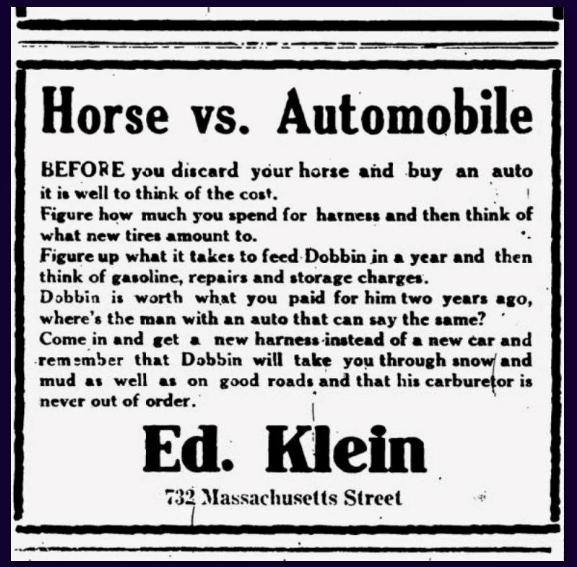
- Head of Digital Customer Service Platforms, PCP, WFM, Data and AI
- MBA Professor – FIAP

@vinicius caridá @vfcarida @vinicius caridá @vfcarida @vinicius caridá @vfcarida


VINICARIDÁ
Human-Centered Data & AI

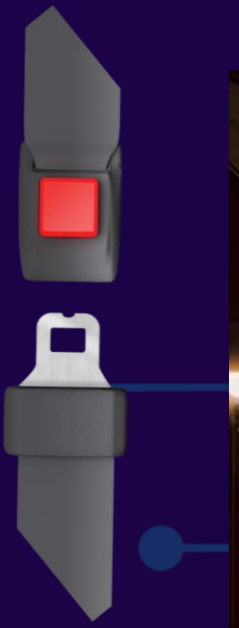
VINICARIDÁ
Human-Centered Data & AI


“
Seres humanos são construtores de ferramentas.
A IA generativa é uma ferramenta notável que criamos.
”
Adaptado de Steve Jobs
VINICARIDÁ
Human-Centered Data & AI



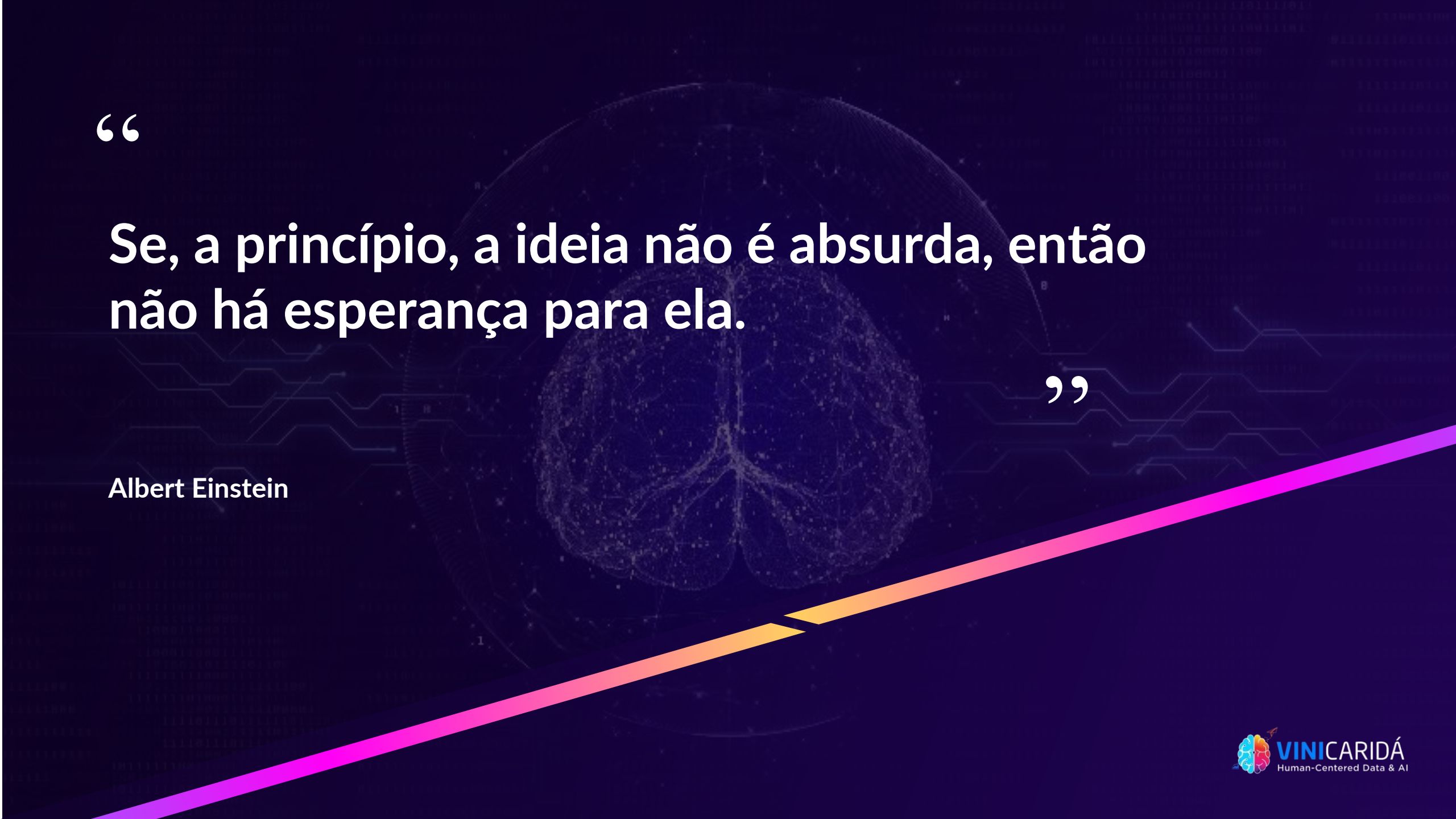
Red Flag Law 1878

Quando a Volvo inventou o cinto de segurança de três pontos, em 1959, ela ofereceu a patente, de graça, a todos os seus concorrentes. De acordo com a empresa, a invenção era tão significativa que tinha mais valor como ferramenta gratuita para salvar vidas do que como algo para se obter lucro.

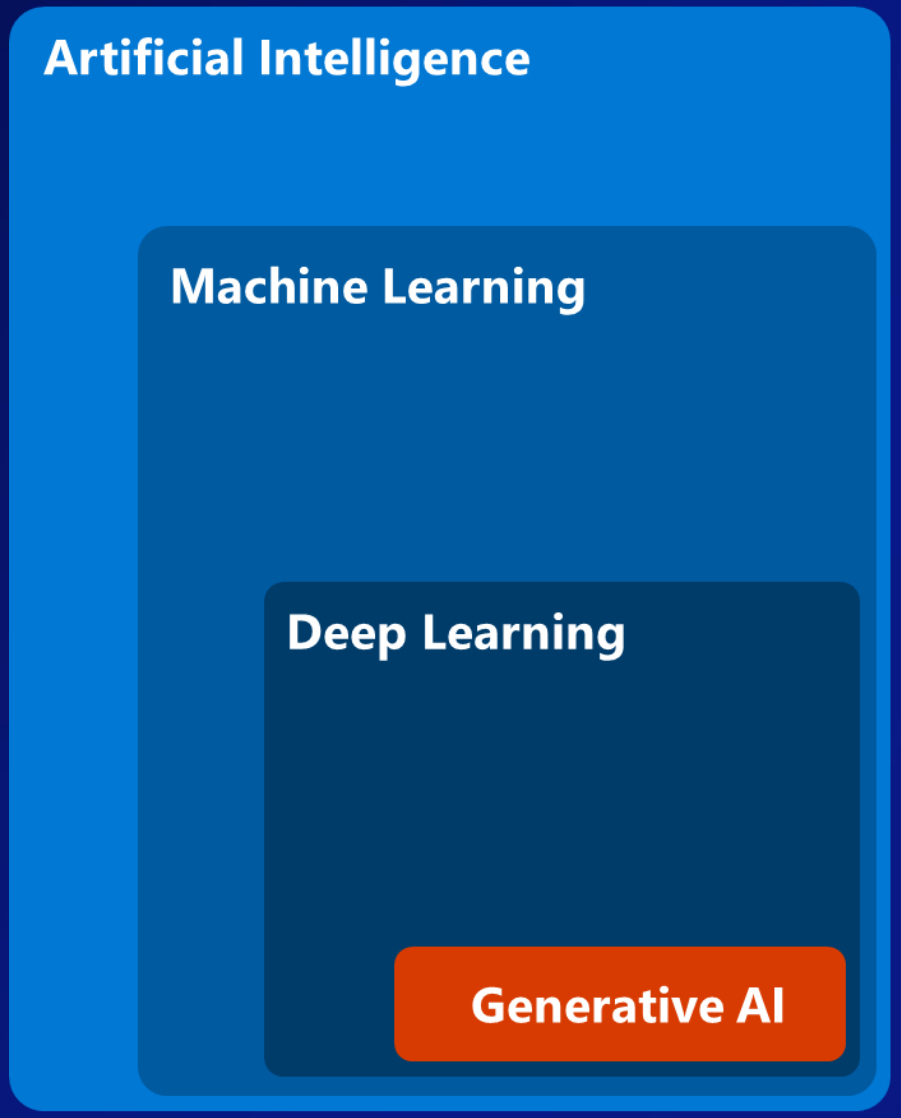


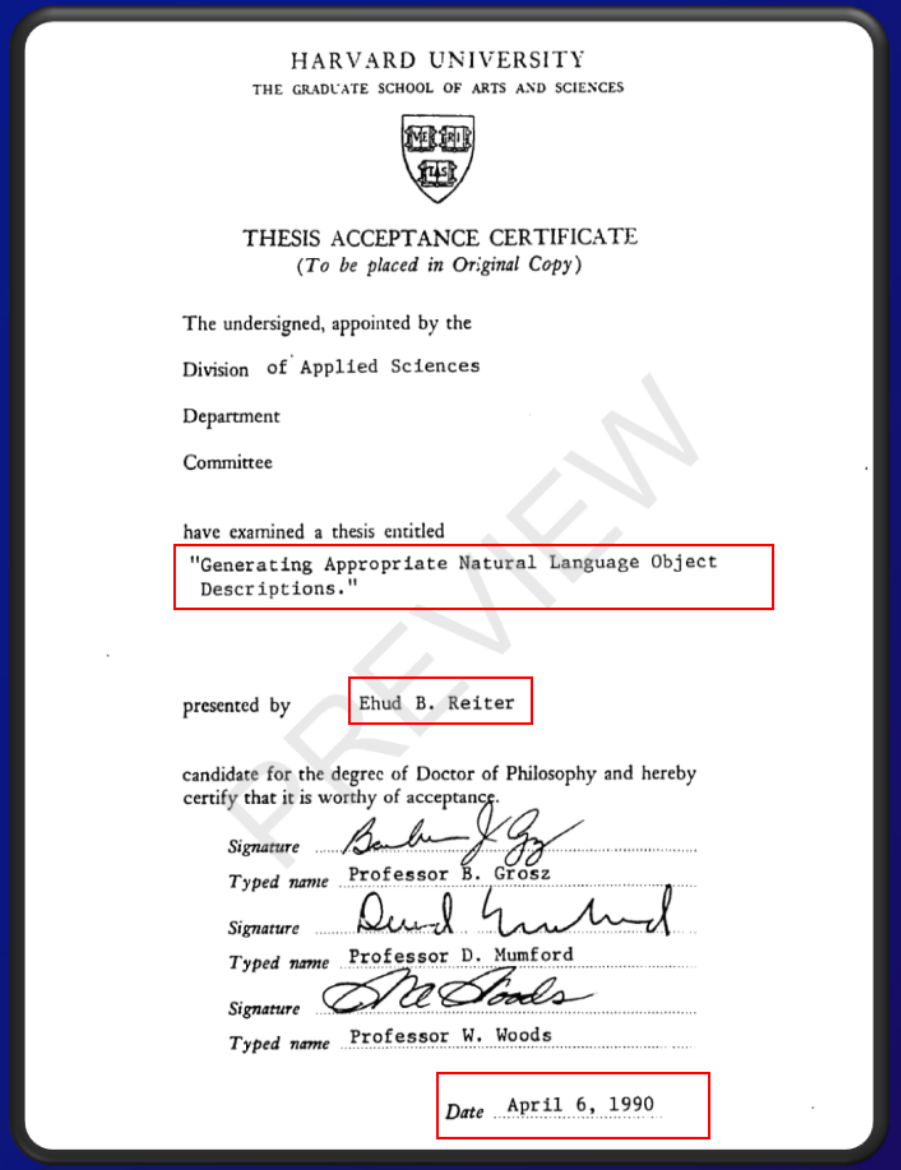
Melhores Regulamentações, sinalizações, infraestrutura e estradas

Ferio ABS 1984


“
Se, a princípio, a ideia não é absurda, então não há esperança para ela.
”
Albert Einstein
VINICARIDÁ
Human-Centered Data & AI

AI Generativa





HARVARD UNIVERSITY
THE GRADUATE SCHOOL OF ARTS AND SCIENCES

THESIS ACCEPTANCE CERTIFICATE
(To be placed in Original Copy)

The undersigned, appointed by the

Division of Applied Sciences

Department

Committee

have examined a thesis entitled

"Generating Appropriate Natural Language Object Descriptions."

presented by Ehud B. Reiter

candidate for the degree of Doctor of Philosophy and hereby certify that it is worthy of acceptance.

Signature .....
Typed name Professor B. Grosz

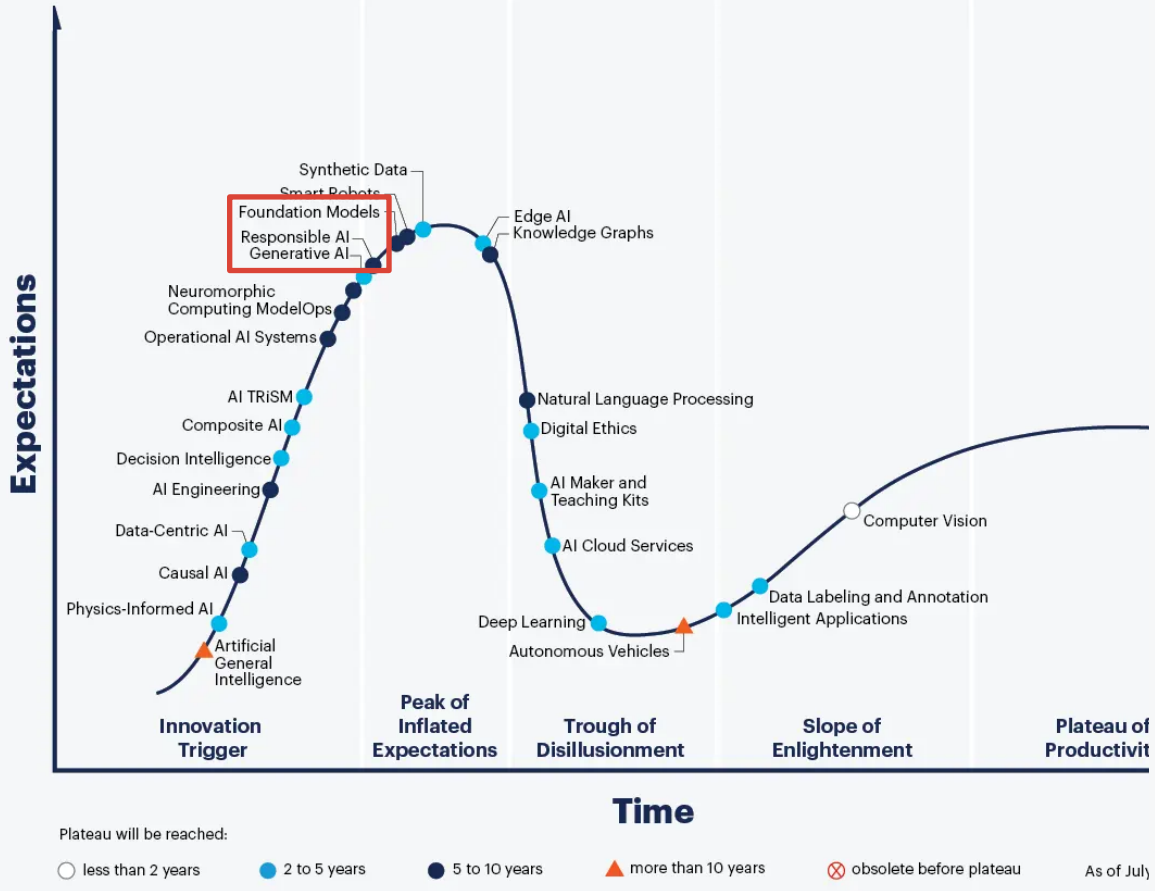
Signature .....
Typed name Professor D. Mumford

Signature .....
Typed name Professor W. Woods

Date April 6, 1990

AI Generativa | Hype Cycle for Artificial Intelligence

2022

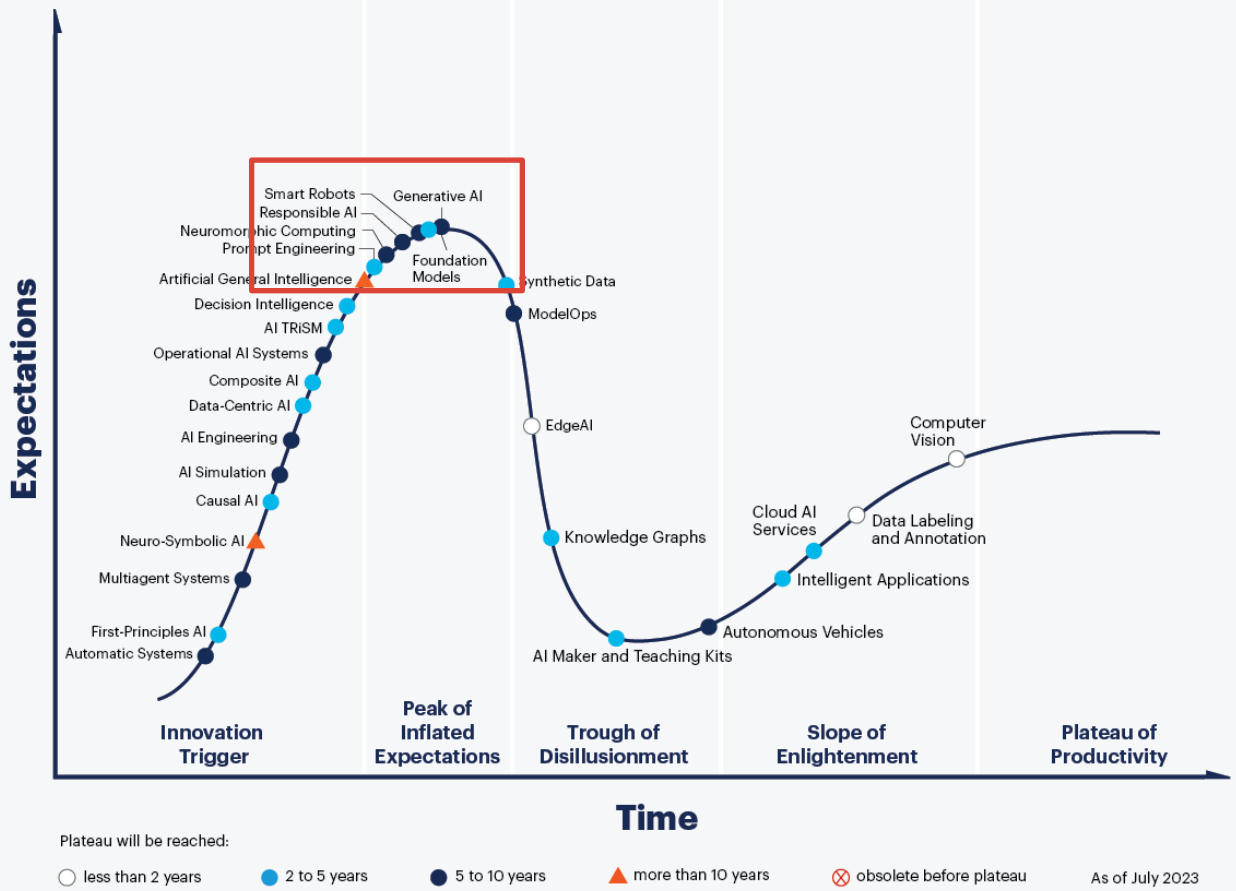


gartner.com

Source: Gartner

Garti

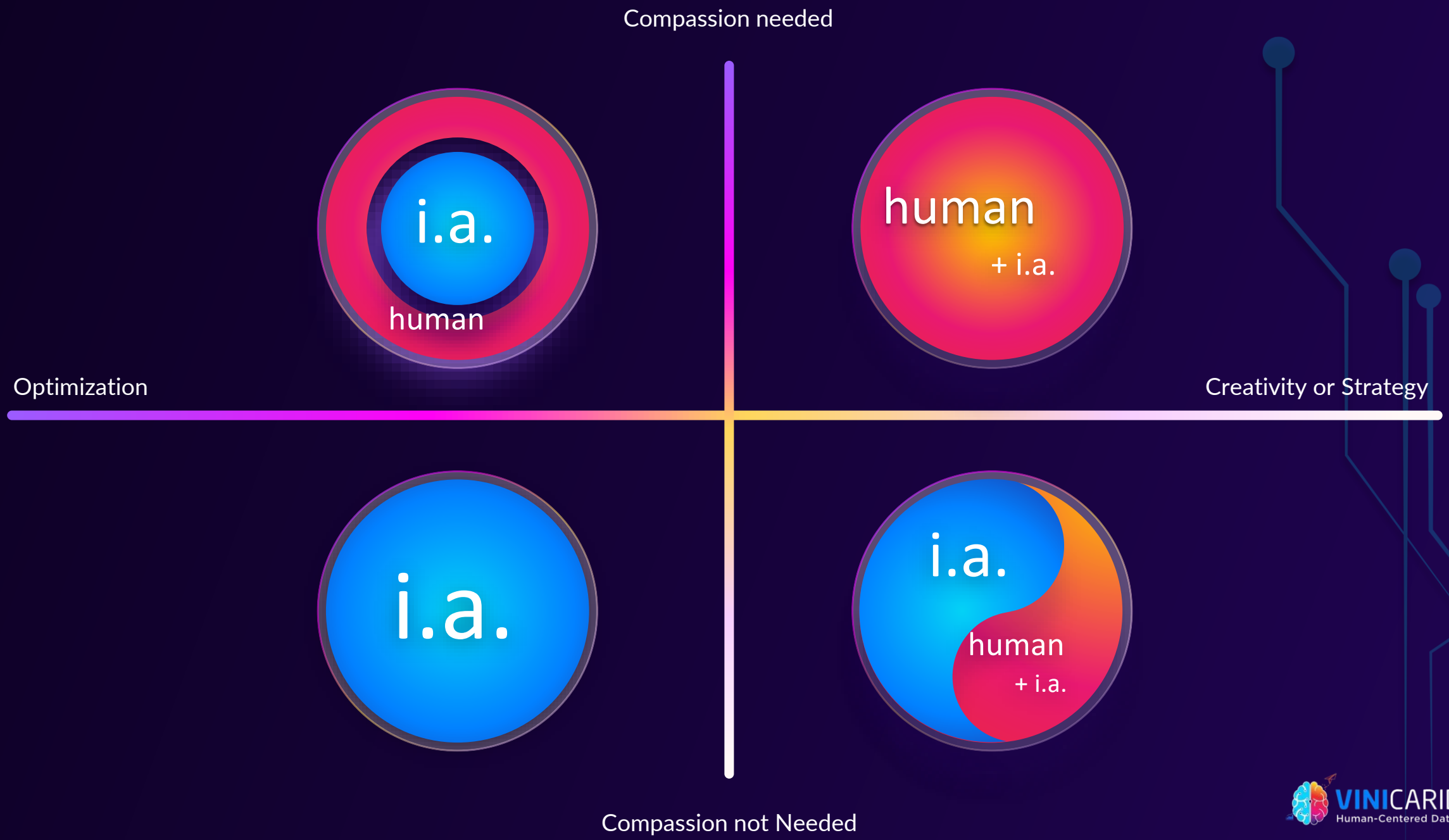
2023



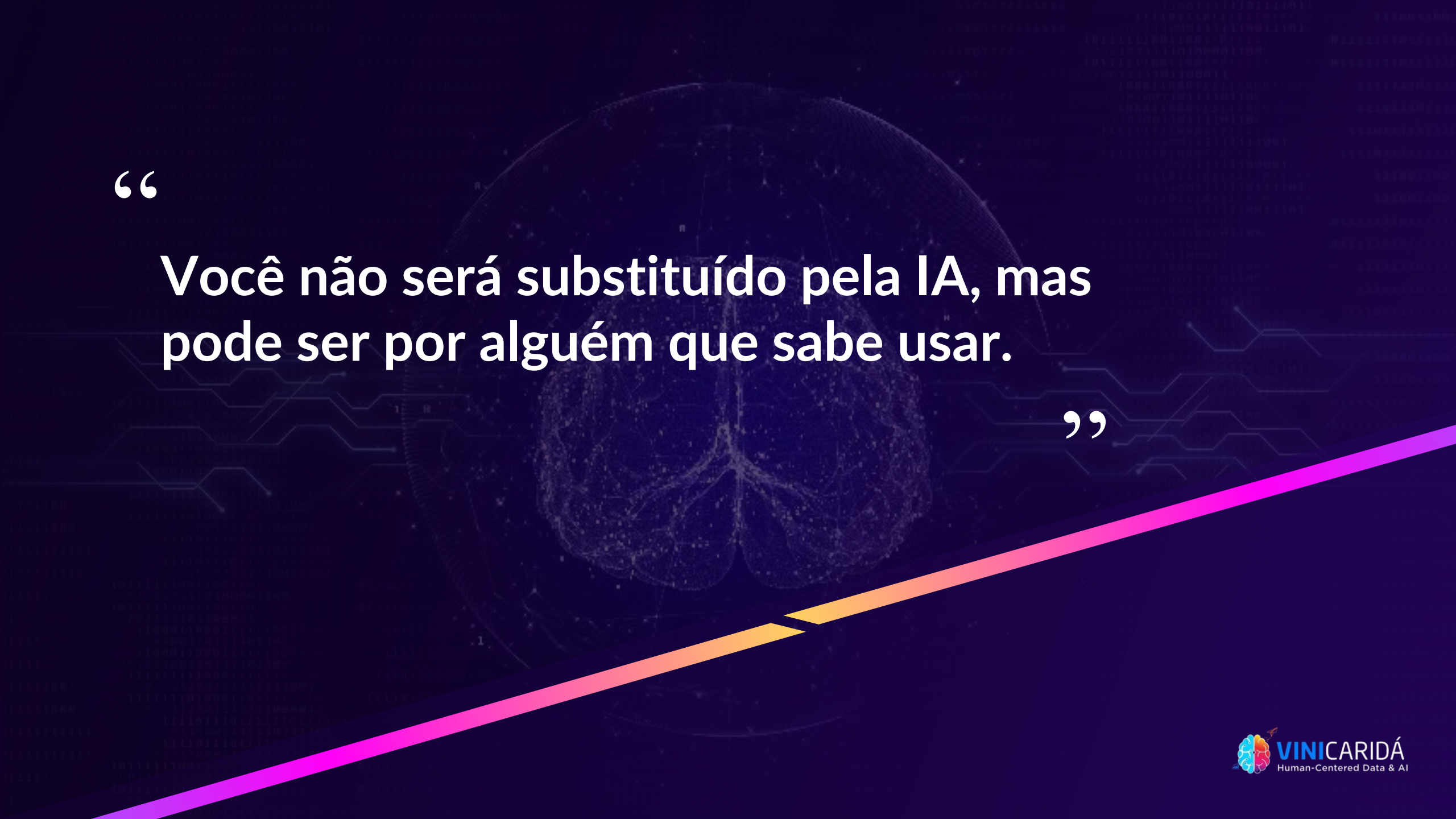
gartner.com

Source: Gartner

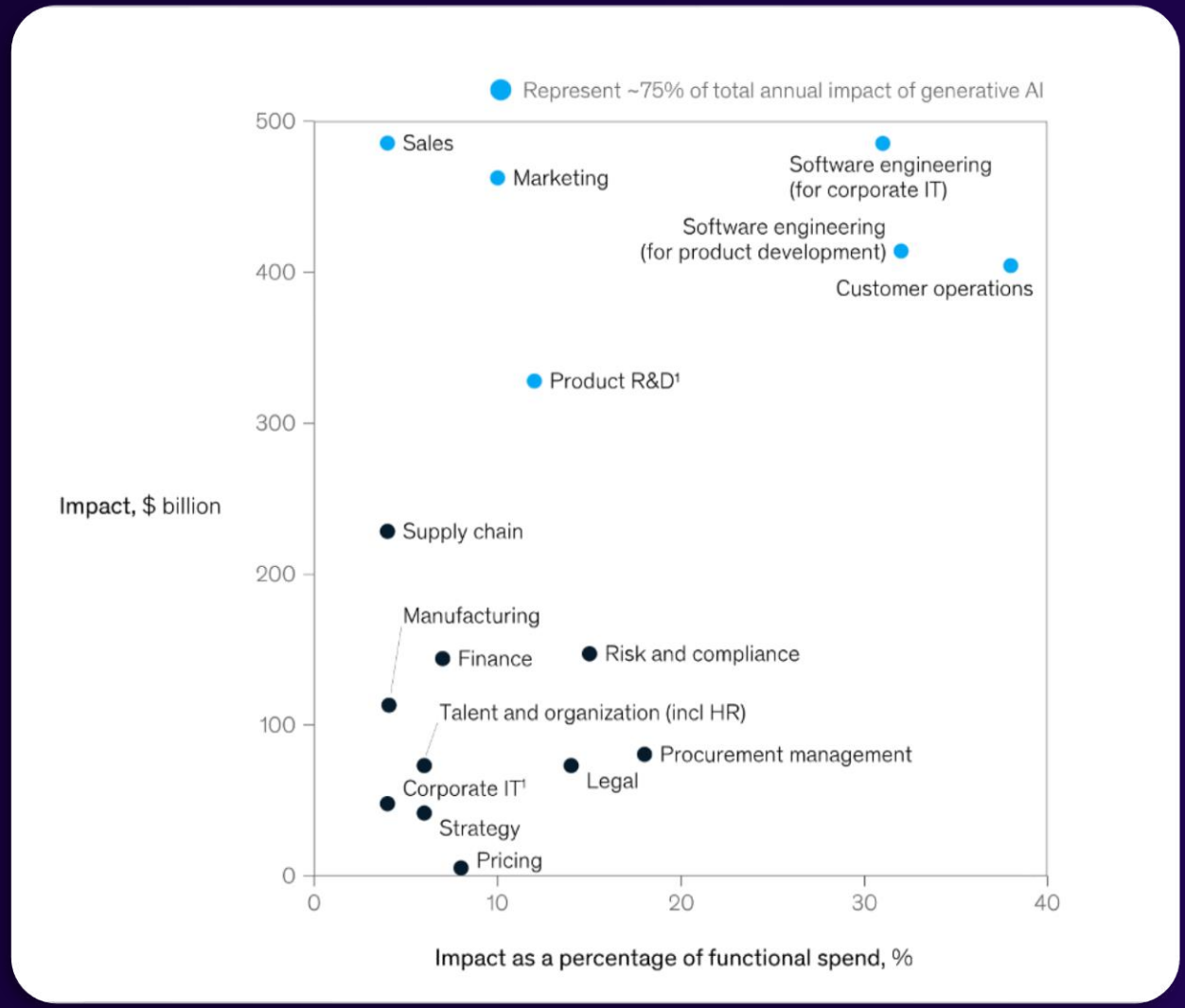
Gartner



VINICARIDÁ
Human-Centered Data & AI

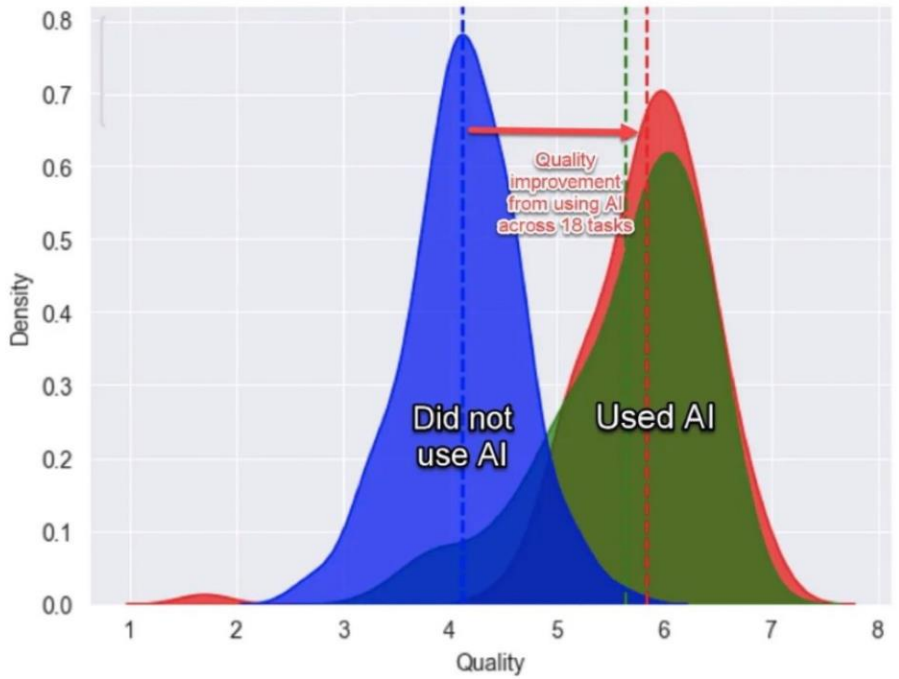

“
Você não será substituído pela IA, mas pode ser por alguém que sabe usar.
”
VINICARIDÁ
Human-Centered Data & AI

Using generative AI in just a few functions could drive most of the technology's impact across potential corporate use cases



https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/The-economic-potential-of-generative-AI-The-next-productivity-frontier

VINICARIDÁ
Human-Centered Data & AI



Distribution of output quality across all the tasks. The blue group did not use AI, the green and red groups used AI, the red group got some additional training on how to use AI.

Os melhores desempenhos na sua empresa serão trabalhadores habilitados por IA.

O **Boston Consulting Group (BCG)**, uma das empresas de consultoria mais prestigiadas, está testando os impactos da IA em seus funcionários.

Consultores que utilizam o GPT-4 concluíram **12,2%** mais tarefas, realizaram tarefas **25,1%** mais rapidamente e produziram resultados com **40%** mais qualidade.

Estamos em um ponto de inflexão massivo para a produtividade dos trabalhadores.

https://www.linkedin.com/posts/alliekmiller_guess-what-the-top-performers-at-your-company-activity-7109183288295968768-B6-M/?utm_source=share&utm_medium=member_android

https://deliverypdf.ssrn.com/delivery.php?ID=044090101066107122071069086006092088034056090036079024023026127111066011108030083075060056032028047044115028102031005085100115112075028080092085020115099016014116048014009071025025113102087092086030024072007080031068115109070026096029028018065101008&EXT=pdf&INDEX=TRUE

VINICARIDÁ
Human-Centered Data & AI

exame.

"IA generativa será a chave para produtiv
no futuro do trabalho"

Fabio Costa, general manager da Salesforce Brasil, fala do papel crucial da inteligê
na eficiência

https://exame.com/inteligencia-artificial/ia-generativa-sera-a-chave-para-produtividade-no-f
trabalho/

Forbes

lysis: How AI
ve Leadership

enerative-analysis-how-ai-enables-effective-

Organização Internacional do Trabalho

Inteligência artificial

A inteligência artificial generativa deve complemer
empregos, em vez de destruí-los

Relatório da OIT avalia o impacto da inteligência artificial generativa na qu
e na qualidade do emprego.

https://www.ilo.org/brasilia/noticias/WCMS_890974/lang--pt/index.htm

ive leaders
to make
sions

15 artigos + Seguir

enerative-analysis-how-ai-enables-

Bloomberg

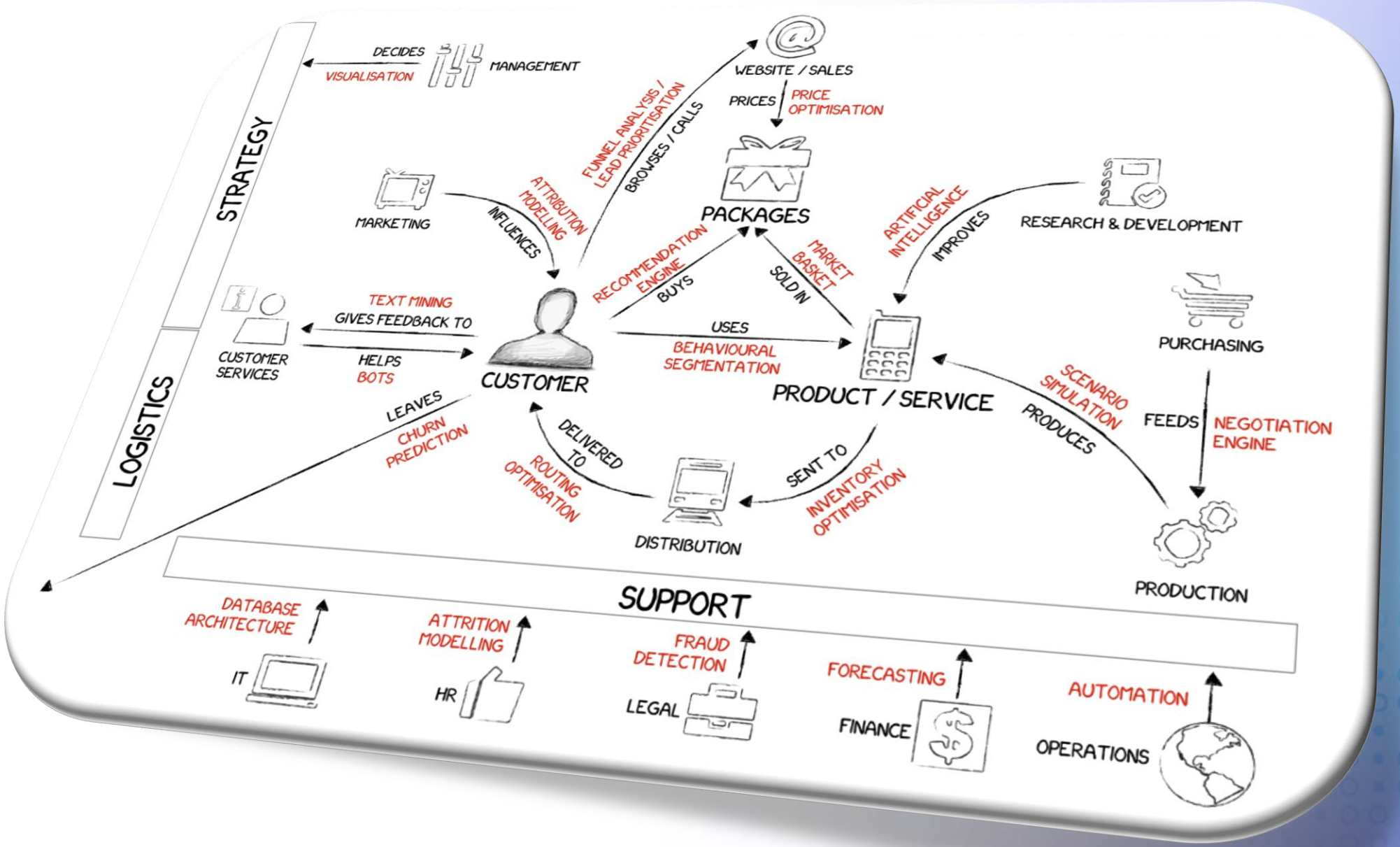
Live Now Markets Economics Industries Technology Politics Wealth

Work Shift
Technology &
Skills

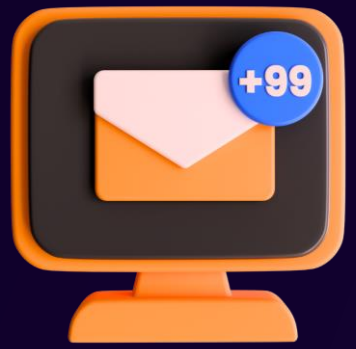
Generative AI Boos
Productivity 14% in
Study

https://www.bloomberg.com/news/articles/2023-04-24/generative-ai-boosts-worker-productivity-

VINICARIDÁ
Human-Centered Data & AI



Fonte: https://applied-data.science



99% accuracy detecting spammy email

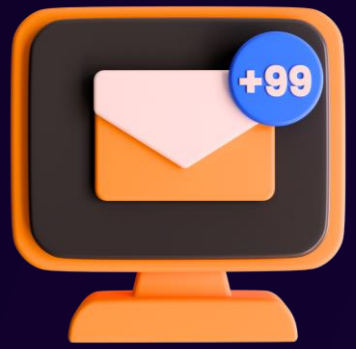
91.7%
Large linear ML classifier

+4,7%
Rule Based System

+3.5%
Deep learning

?
Next gen

VINICARIDÁ
Human-Centered Data & AI



99% accuracy detecting spammy email

91.7%
Large linear ML classifier

+4,7%
Rule Based System

+3.5%
Deep learning

!!
Generative AI

?
Next gen

VINICARIDÁ
Human-Centered Data & AI

IA Generativa | Riscos

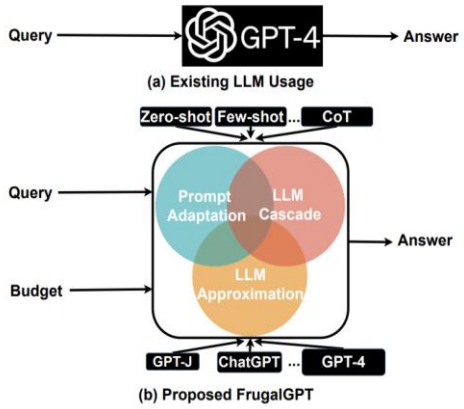
- Vazamento de informações
- Invasão
- “Alucinações”
- Custos
- Trabalhista
- Imagem

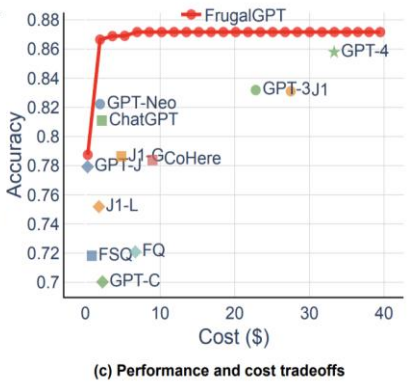
FrugalGPT: How to Use Large Language Models While Reducing Cost and Improving Performance

Lingjiao Chen, Matei Zaharia, James Zou

Stanford University

https://arxiv.org/pdf/2305.05176.pdf





versão GPT

_Para a release 3, foram avaliadas ambas as versões de GPT (3.5 e 4)

Resultado	3.3	vs	3.5
Taxa de Erros	24%	vs	16%
oto (ano)	$3k	vs	$34k
j. (ano)	$2MM	vs	$22MM
e resp	13s	vs	33s

ão do grupo: versão 3.5

VINICARIDÁ
Human-Centered Data & AI

The AWS AI/ML stack

Consumer

AI SERVICES

Generative AI

SCALING GENERATIVE AI
- Amazon Bedrock

AMAZON BUILT FOUNDATIONAL MODEL
- Amazon Titan

Specialized AI Services

BUSINESS PROCESSES
- Amazon Personalize
- Amazon Forecast
- Amazon Fraud Detector
- Amazon Lookout for Metrics

SEARCH
- Amazon Kendra

CONVERSATION
- Amazon Lex
- Contact Lens
- Voice ID
- Amazon Transcribe Call Analytics

CODE + DEVOPS
- Amazon CodeGuru
- Amazon CodeWhisperer
- Amazon DevOps Guru

INDUSTRIAL
- Amazon Monitron
- Amazon Lookout for Equipment
- Amazon Lookout for Vision

HEALTH
- AWS HealthLake
- Amazon Comprehend Medical
- Amazon Transcribe Medical
- AWS HealthOmics

Core

TEXT
- Amazon Translate
- Amazon Comprehend

SPEECH
- Amazon Polly
- Amazon Transcribe

VISION
- Amazon Textract
- Amazon Rekognition
- AWS Panorama

Tuner

AMAZON SAGEMAKER

JUMPSTART
Use pre-built models in SageMaker

CANVAS
No-code ML for business analysts

STUDIO LAB
Learn ML

GROUND TRUTH
Label data

STUDIO IDE
- Prepare data
- Store features
- Geospatial ML
- Build with notebooks
- Train models
- Tune parameters
- Deploy in production
- Manage and monitor

CI/CD | GOVERNANCE | RESPONSIBLE ML

Provider

ML FRAMEWORKS & INFRASTRUCTURE

PyTorch, TensorFlow | Amazon EC2 | CPUs | GPUs | AWS Inferentia | AWS Trainium | Habana Gaudi | FPGA

The AWS AI/ML stack

Consumer

Tuner

Provider

AI SERVICES

Generative AI

SCALING GENERATIVE AI
- Amazon Bedrock NEW

AMAZON BUILT FOUNDATIONAL MODEL
- Amazon Titan NEW

Specialized AI Services

BUSINESS PROCESSES
- Amazon Personalize
- Amazon Forecast
- Amazon Fraud Detector
- Amazon Lookout for Metrics

SEARCH
- Amazon Kendra

CONVERSATION
- Amazon Lex
- Contact Lens
- Voice ID
- Amazon Transcribe Call Analytics

CODE + DEVOPS
- Amazon CodeGuru
- NEW Amazon CodeWhisperer
- Amazon DevOps Guru

INDUSTRIAL
- Amazon Monitron
- Amazon Lookout for Equipment
- Amazon Lookout for Vision

HEALTH
- AWS HealthLake
- Amazon Comprehend Medical
- Amazon Transcribe Medical
- AWS HealthOmics

Core

TEXT
- Amazon Translate
- Amazon Comprehend

SPEECH
- Amazon Polly
- Amazon Transcribe

VISION
- Amazon Textract
- Amazon Rekognition
- AWS Panorama

AMAZON SAGEMAKER

JUMPSTART NEW
Use pre-built models in SageMaker

CANVAS
No-code ML for business analysts

STUDIO LAB
Learn ML

GROUND TRUTH
Label data

STUDIO IDE
- Prepare data
- Store features
- Geospatial ML
- Build with notebooks
- Train models
- Tune parameters
- Deploy in production
- Manage and monitor

CI/CD | GOVERNANCE | RESPONSIBLE ML

ML FRAMEWORKS & INFRASTRUCTURE

PyTorch, TensorFlow | Amazon EC2 | CPUs | GPUs | NEW AWS Inferentia | NEW AWS Trainium | Habana Gaudi | FPGA

Itaú VINICARIDÁ Human-Centered Data & AI

Como funciona?

Como representar contexto/significado das palavras

Você sabe qual o significado da palavra **tezgüino**?

exemplo apresentado em: https://github.com/jacobeisenstein/gt-nlp-class/blob/master/notes/eisenstein-nlp-notes.pdf - cap 14

Como representar contexto/significado das palavras

Observe a palavra **tezgüino** em diferentes contextos:

Como representar contexto/significado das palavras

Observe a palavra **tezgüino** em diferentes contextos:

- Uma garrafa de **tezgüino** está sobre a mesa.

Como representar contexto/significado das palavras

Observe a palavra **tezgüino** em diferentes contextos:

- Uma garrafa de **tezgüino** está sobre a mesa.
- Todo mundo gosta de beber **tezgüino**.

Como representar contexto/significado das palavras

Observe a palavra **tezgüino** em diferentes contextos:

- Uma garrafa de **tezgüino** está sobre a mesa.
- Todo mundo gosta de beber **tezgüino**.
- Você pode ficar bêbado com **tezgüino**.

Como representar contexto/significado das palavras

Observe a palavra **tezgüino** em diferentes contextos:

- Uma garrafa de **tezgüino** está sobre a mesa.
- Todo mundo gosta de beber **tezgüino**.
- Você pode ficar bêbado com **tezgüino**.
- **Tezgüino** é feito de milho.

Como representar contexto/significado das palavras

Observe a palavra **tezgüino** em diferentes contextos:

- Uma garrafa de **tezgüino** está sobre a mesa.
- Todo mundo gosta de beber **tezgüino**.
- Você pode ficar bêbado com **tezgüino**.
- **Tezgüino** é feito de milho.

Consegue entender o que é tezgüino?

Como representar contexto/significado das palavras

Observe a palavra **tezgüino** em diferentes contextos:

- Uma garrafa de **tezgüino** está sobre a mesa.
- Todo mundo gosta de beber **tezgüino**.
- Você pode ficar bêbado com **tezgüino**.
- **Tezgüino** é feito de milho.

Com o contexto, conseguimos identificar do que se refere a palavra **tezgüino**.

Tezgüino:= é uma bebida alcoólica feita a base de milho.

Como representar contexto/significado das palavras

Como o cérebro faz isso?

Como representar contexto/significado das palavras

Quais outras palavras se “***encaixam***” nos slots das perguntas 1 até 4?

1. Uma garrafa de _____ está sobre a mesa.

Como representar contexto/significado das palavras

Quais outras palavras se “***encaixam***” nos slots das perguntas 1 até 4?

1. Uma garrafa de _____ está sobre a mesa.
2. Todo mundo gosta de beber _____.

Como representar contexto/significado das palavras

Quais outras palavras se “***encaixam***” nos slots das perguntas 1 até 4?

1. Uma garrafa de _____ está sobre a mesa.
2. Todo mundo gosta de beber _____.
3. Você pode ficar bêbado com _____.

Como representar contexto/significado das palavras

Quais outras palavras se “***encaixam***” nos slots das perguntas 1 até 4?

1. Uma garrafa de _____ está sobre a mesa.
2. Todo mundo gosta de beber _____.
3. Você pode ficar bêbado com _____.
4. _____ é feito de milho.

Como representar contexto/significado das palavras

Inserindo contexto de forma manual...

1. Uma garrafa de _____ está sobre a mesa.
2. Todo mundo gosta de beber _____.
3. Você pode ficar bêbado com _____.
4. _____ é feito de milho.

	(1)	(2)	(3)	(4) ← contextos
tezgüino	1	1	1	1
som	0	0	0	0
suco de laranja	1	1	0	0
vinho	1	1	1	0

Como representar contexto/significado das palavras

Inserindo contexto de forma manual...

1. Uma garrafa de _____ está sobre a mesa.
2. Todo mundo gosta de beber _____.
3. Você pode ficar bêbado com _____.
4. _____ é feito de milho.

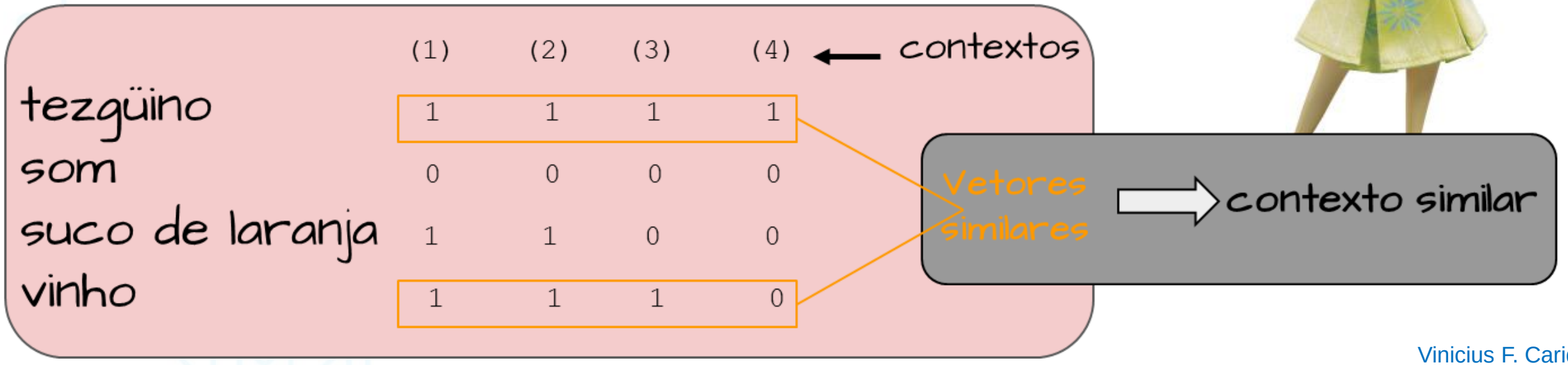
	(1)	(2)	(3)	(4) ← contextos
tezgüino	1	1	1	1
som	0	0	0	0
suco de laranja	1	1	0	0
vinho	1	1	1	0

Vetores similares

Como representar contexto/significado das palavras

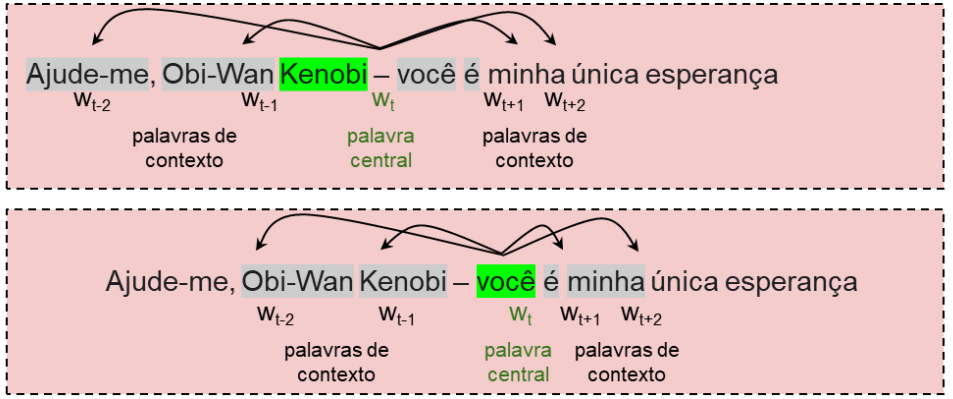
Inserindo contexto de forma manual...

1. Uma garrafa de _____ está sobre a mesa.
2. Todo mundo gosta de beber _____.
3. Você pode ficar bêbado com _____.
4. _____ é feito de milho.



Processamento de Linguagem Natural

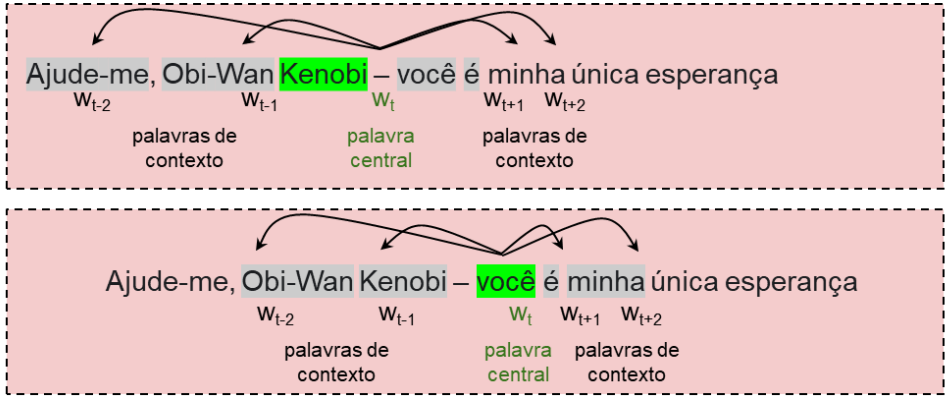
Natural Language Processing (NLP)

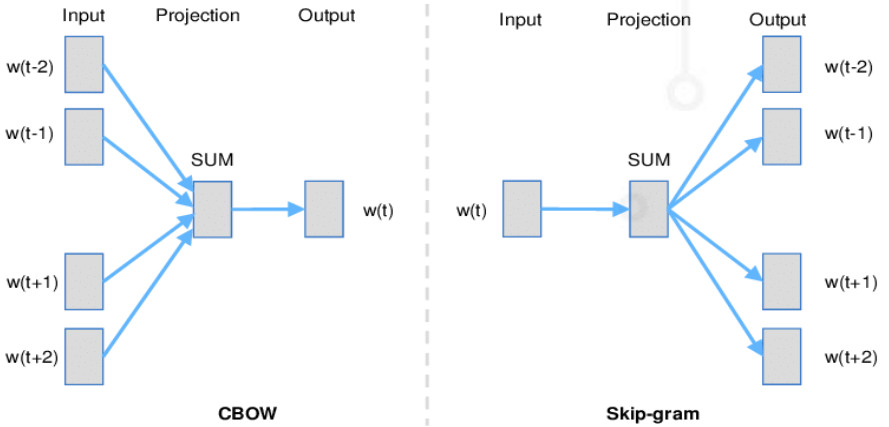


Processamento de Linguagem Natural

Natural Language Processing (NLP)

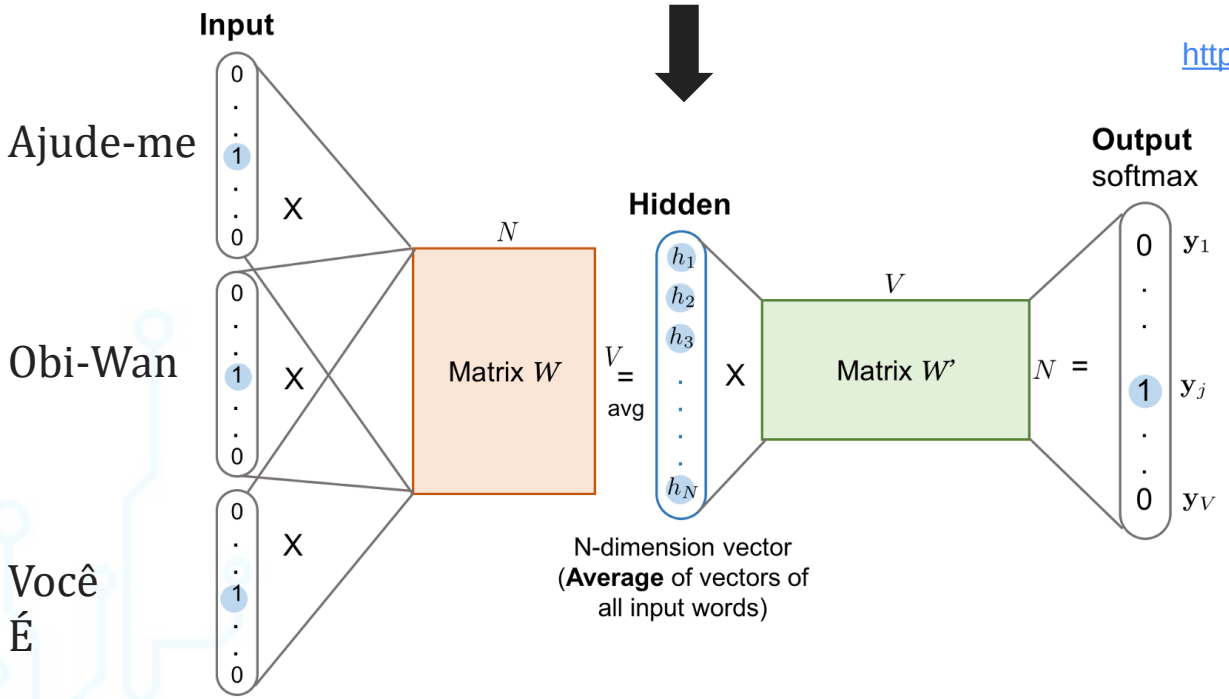
VINI CARIDÁ
Human-Centered Data & AI





https://arxiv.org/abs/1301.3781

https://arxiv.org/pdf/1310.4546.pdf



MultiLayer Perceptron MLP

https://dl.acm.org/doi/10.5555/1639537.1639542

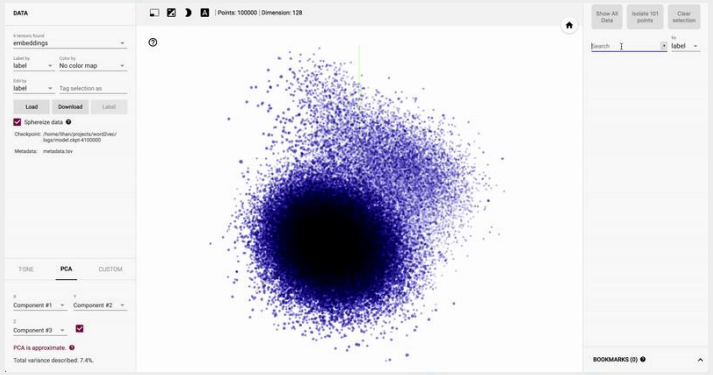
Self Supervised Learning (SSL)

https://arxiv.org/abs/2110.09327

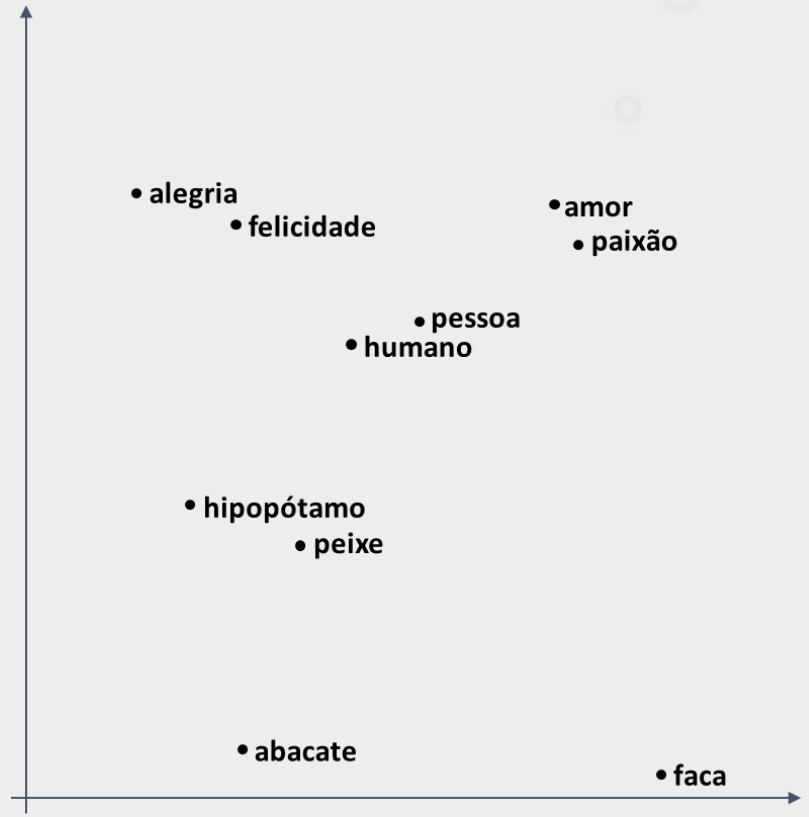
Word Embedding

Representação densa

Palavras com significado similar próximas (semântica)

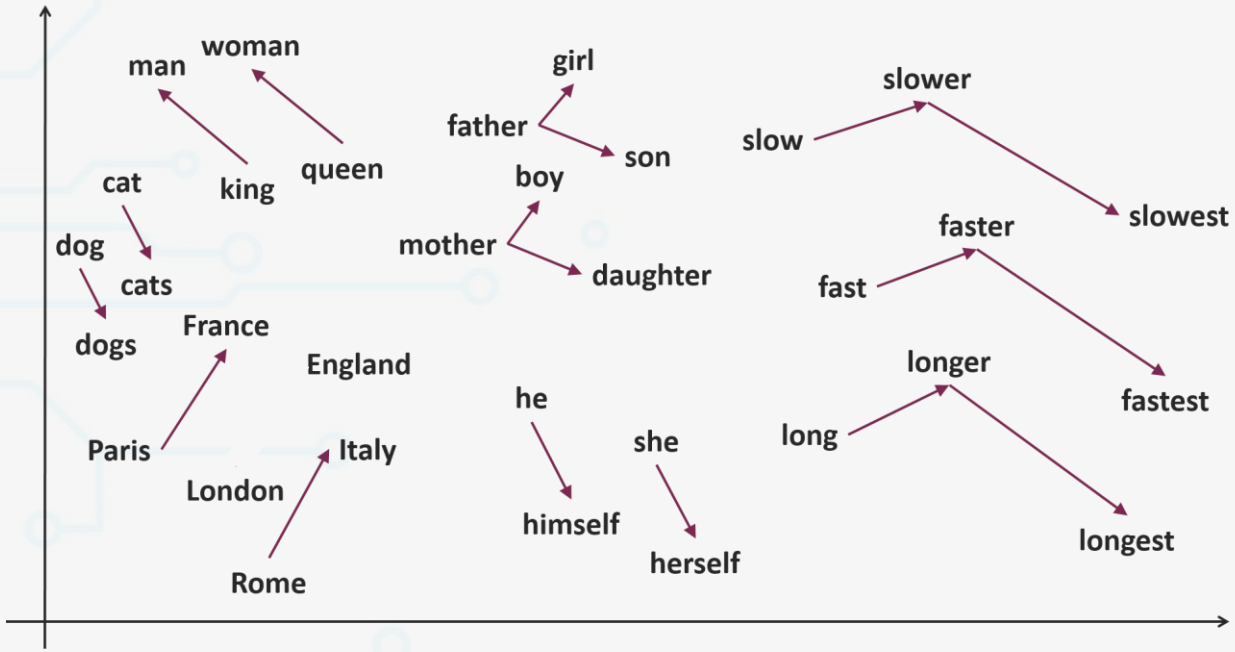


	tamanho do embedding				
a	.1234	.4763	.0948	.9384	.7363
abafar	.6463	.2347	.9434	.3232	.5832
acampar	.5546	.7293	.1039	.3220	.1342
acordo	.9482	.3293	.4201	.4403	.8494
adeus	.9584	.7564	.9924	.2134	.1123
aéreo	.1111	.4949	.2847	.2856	.1345
afiado	.6754	.4342	.9683	.3219	.1103
água	.9831	.3456	.1234	.9873	.3435
águia	.4426	.8464	.9868	.4458	.3244
.	.				
.	.				
.	.				
zimbábue	.9857	.7746	.2234	.3342	.9948



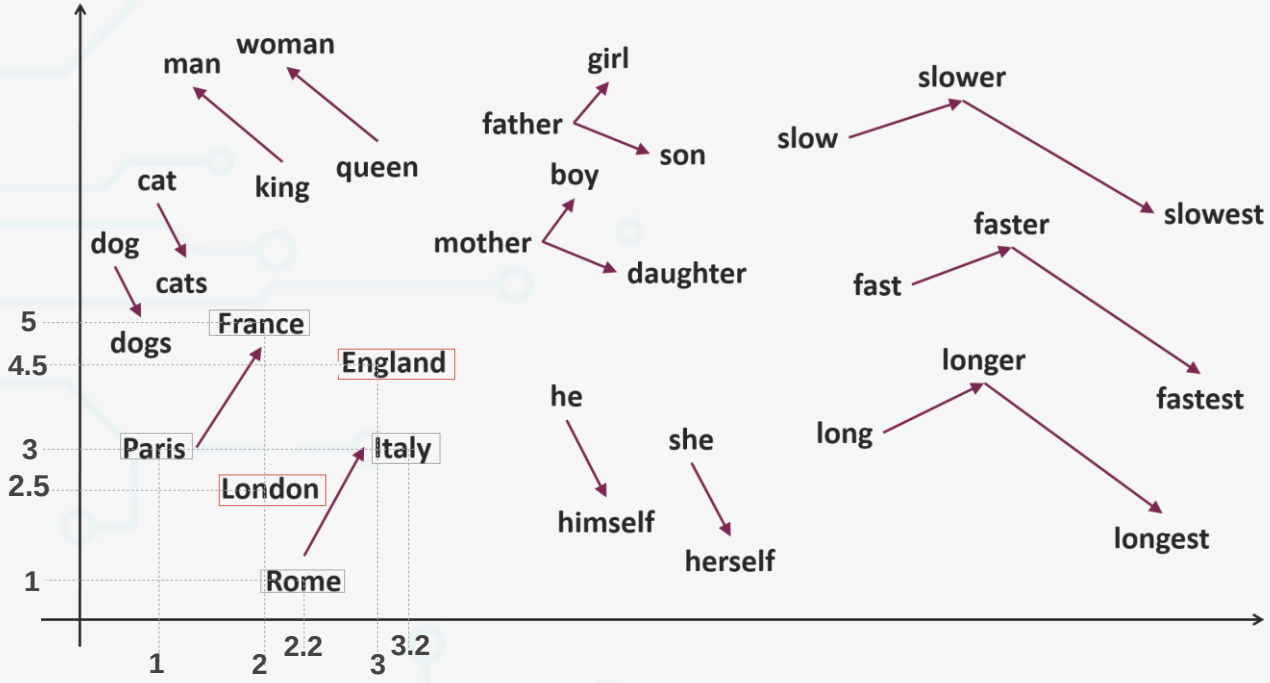
Processamento de Linguagem Natural

Natural Language Processing (NLP)



Processamento de Linguagem Natural

Natural Language Processing (NLP)



Paris [1, 3]

France [2, 5]

London [2, 2.5]

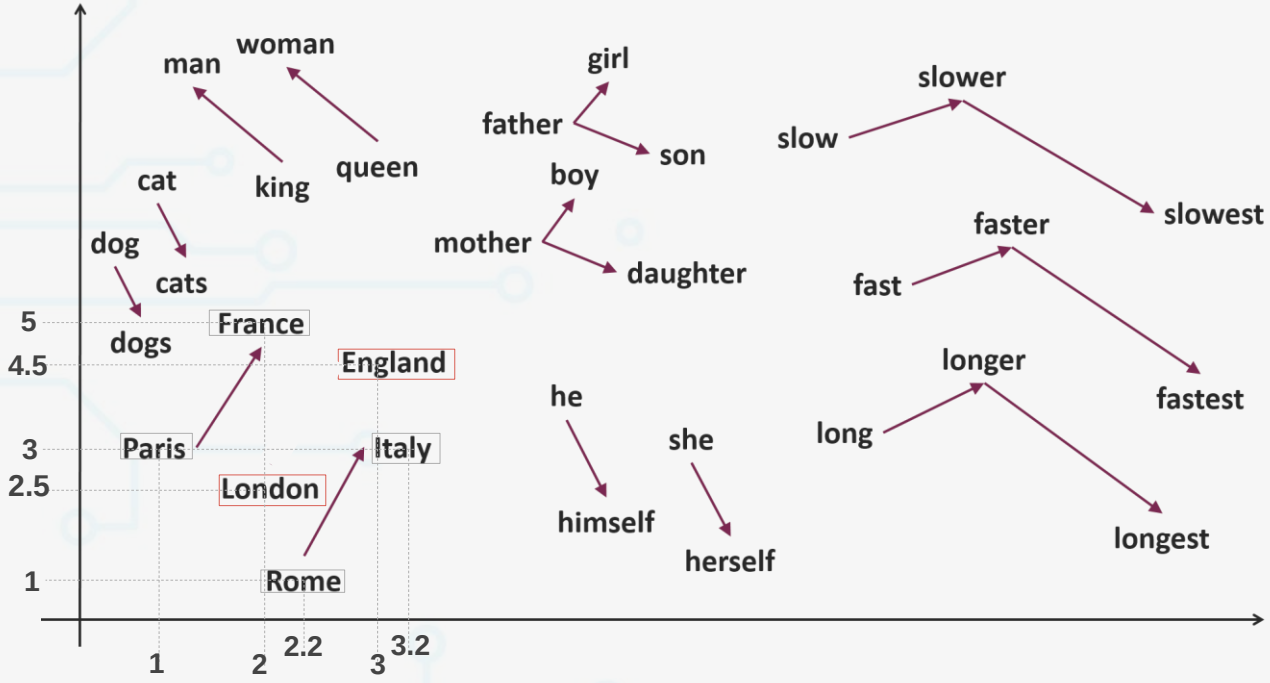
England [3, 4.5]

Rome [2.2, 1]

Italy [3.2, 3]

Processamento de Linguagem Natural

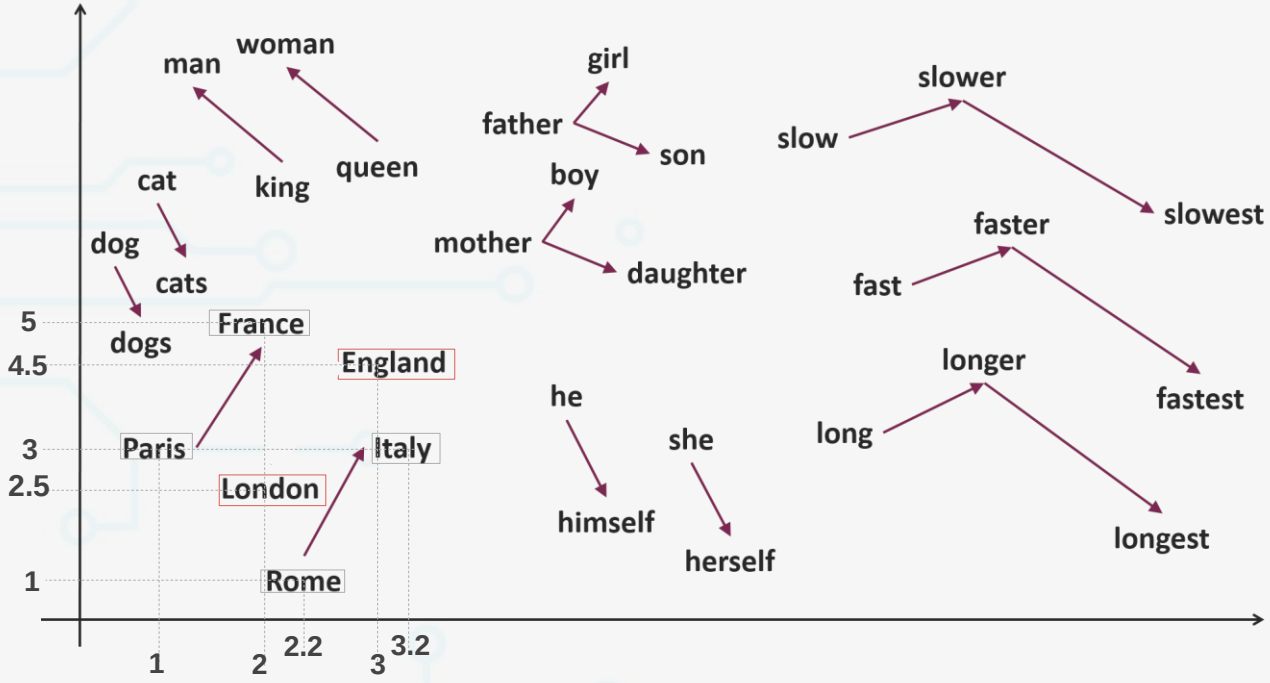
Natural Language Processing (NLP)



Qual a capital da Inglaterra?

Processamento de Linguagem Natural

Natural Language Processing (NLP)

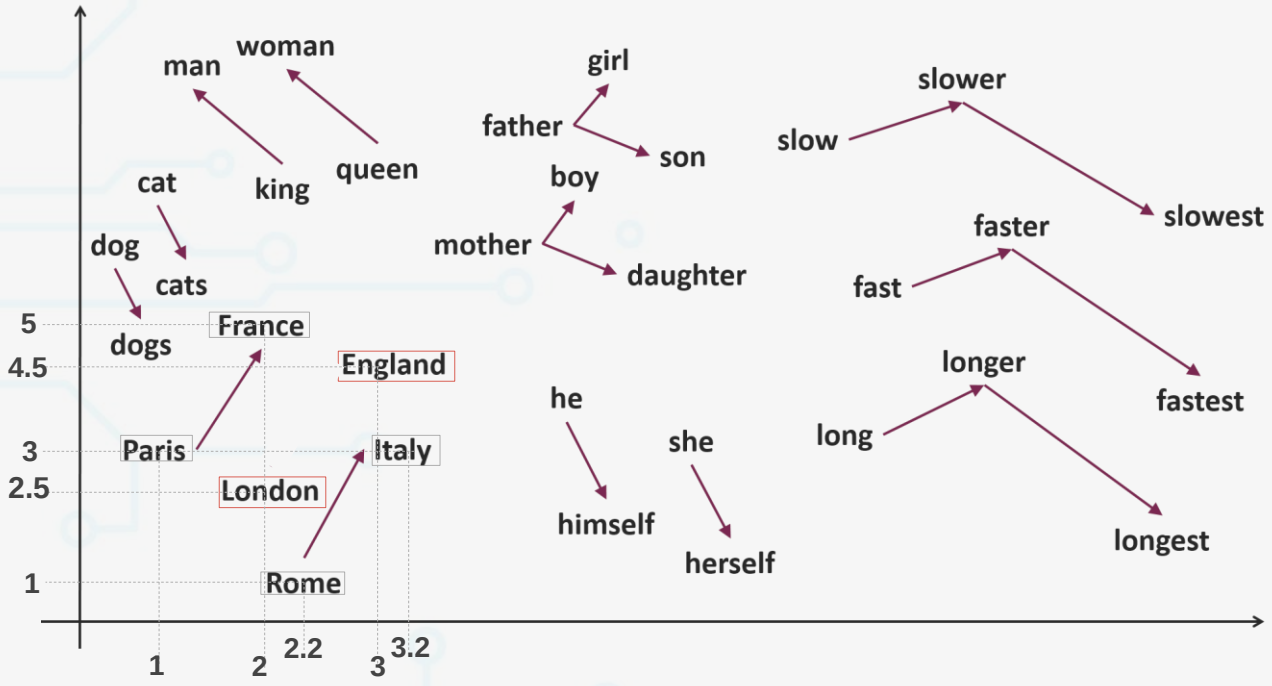


Qual a capital da Inglaterra?

Paris – France + England = ?

Processamento de Linguagem Natural

Natural Language Processing (NLP)



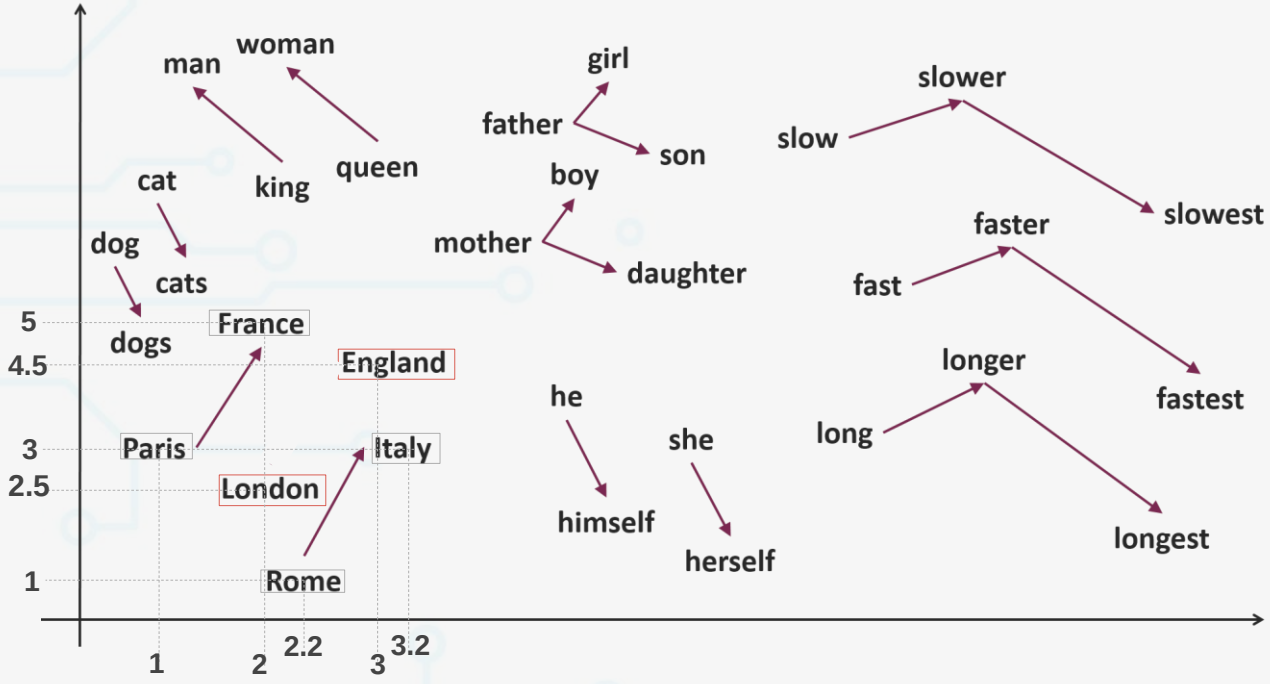
Qual a capital da Inglaterra?

Paris – France + England = ?

Paris [1, 3]
France [2, 5]
=
Result. [-1, -2]
+
England [3, 4.5]
=
Result. [2, 2.5]
==
London [2, 2.5]

Processamento de Linguagem Natural

Natural Language Processing (NLP)



Qual a capital da Inglaterra?

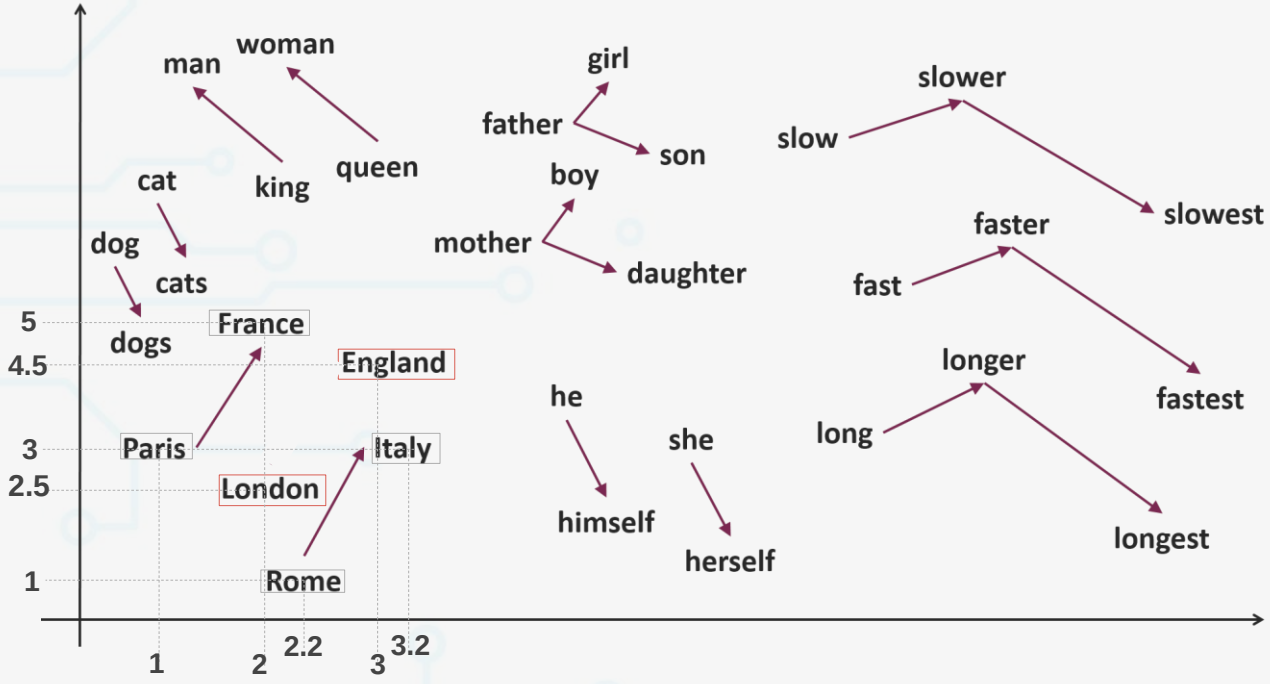
Paris – France + England = ?

Paris [1, 3]
France [2, 5]
=
Result. [-1, -2]
+
England [3, 4.5]
=
Result. [2, 2.5]
==
London [2, 2.5]

Paris – France + England = London

Processamento de Linguagem Natural

Natural Language Processing (NLP)



Qual a capital da Inglaterra?

Paris – France + England = ?

Paris [1, 3]
France [2, 5]
=
Result. [-1, -2]
+
England [3, 4.5]
=
Result. [2, 2.5]
==
London [2, 2.5]

Rome – Italy + England = ?

Rome [2.2, 1]
Italy [3.2, 3]
=
Result. [-1, -2]
+
England [3, 4.5]
=
Result. [2, 2.5]
==
London [2, 2.5]

Paris – France + England = London

Rome – Italy + England = London

Processamento de Linguagem Natural

Natural Language Processing (NLP)

Bag of Words

TF-IDF
Akiko Aizaka, 2003

Word2Vec
Mikolov et al, 2013
Google

Glove
Pennington et al., 2014

Seq2Seq
Sutskever et al, 2014
Google

CoVe
McCann et al., 2017

FastText
Bojanowski et al, 2017
Meta

ULMFiT
Howard et al.. 2018

ELMo
Peters et al, 2018
AI2

Attention Is All You Need

Ashish Vaswani* Google Brain avaswani@google.com
Noam Shazeer* Google Brain noam@google.com
Niki Parmar* Google Research nikip@google.com
Jakob Uszkoreit* Google Research usz@google.com
Llion Jones* Google Research llion@google.com
Aidan N. Gomez* † University of Toronto aidan@cs.toronto.edu
Łukasz Kaiser* Google Brain lukaszkaiser@google.com
Illia Polosukhin* ‡ illia.polosukhin@gmail.com
Google

https://arxiv.org/abs/1706.03762

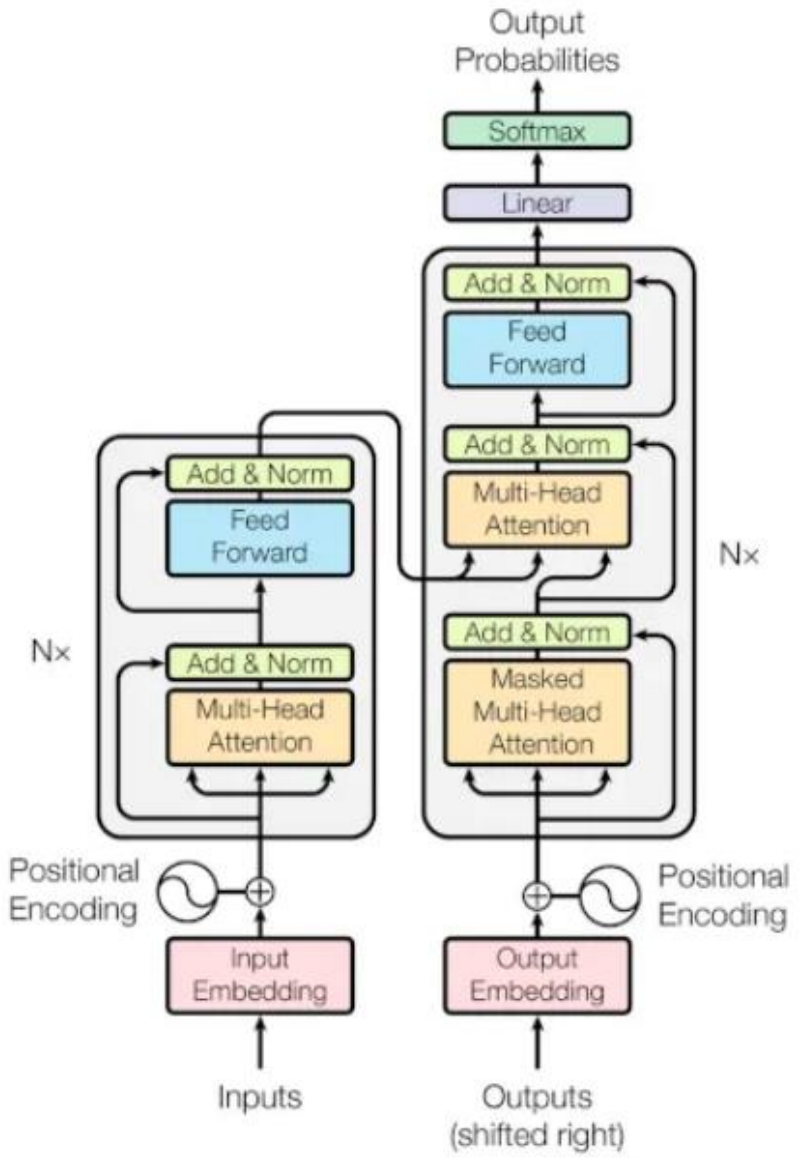
Attention Is All You Need

Ashish Vaswani* Google Brain avaswani@google.com
Noam Shazeer* Google Brain noam@google.com
Niki Parmar* Google Research nikip@google.com
Jakob Uszkoreit* Google Research usz@google.com
Llion Jones* Google Research llion@google.com
Aidan N. Gomez* † University of Toronto aidan@cs.toronto.edu
Łukasz Kaiser* Google Brain lukaszkaiser@google.com
Illia Polosukhin* ‡ illia.polosukhin@gmail.com

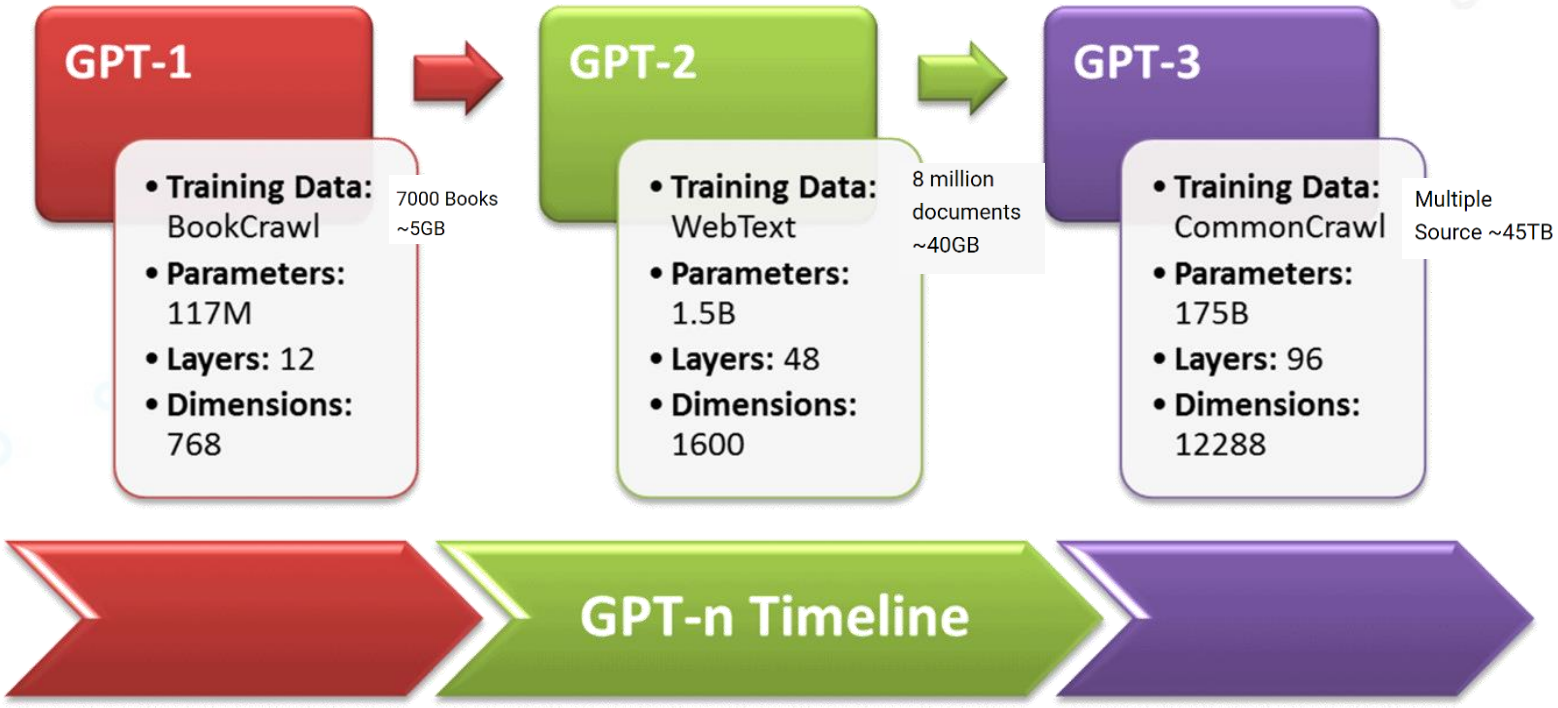
Abstract

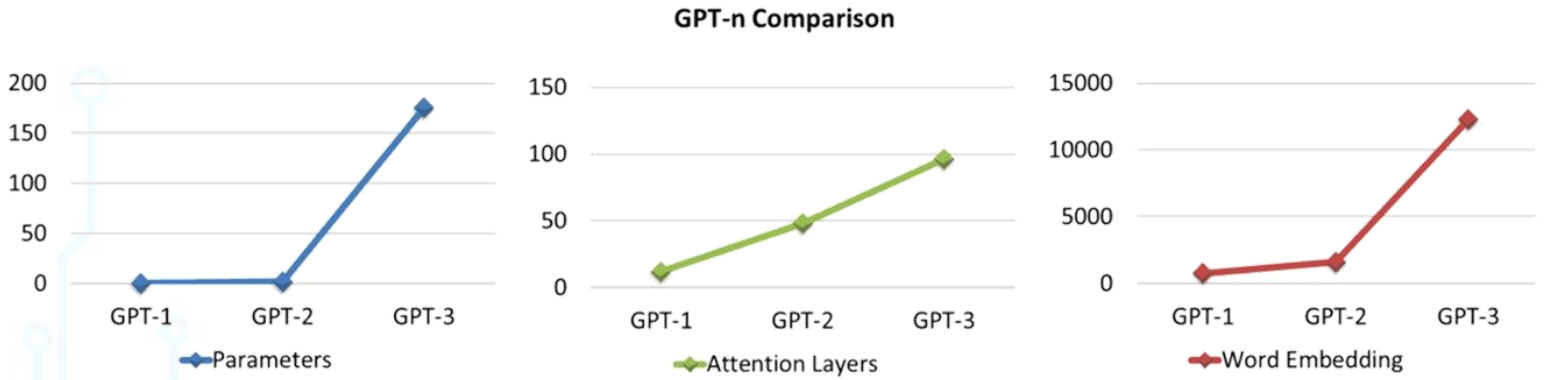
The dominant sequence transduction models are based on complex recurrent or convolutional neural networks that include an encoder and a decoder. The best performing models also connect the encoder and decoder through an attention mechanism. We propose a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely. Experiments on two machine translation tasks show these models to be superior in quality while being more parallelizable and requiring significantly less time to train. Our model achieves 28.4 BLEU on the WMT 2014 English-to-German translation task, improving over the existing best results, including ensembles, by over 2 BLEU. On the WMT 2014 English-to-French translation task,

https://arxiv.org/abs/1706.03762



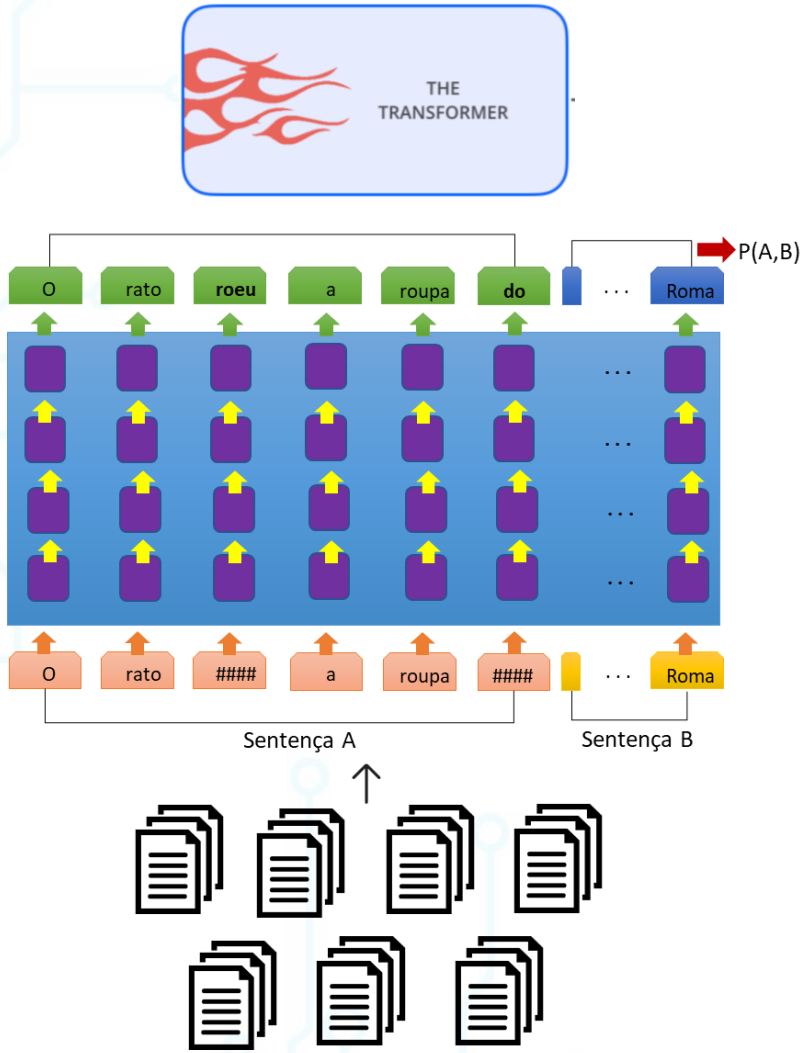
Família GPT



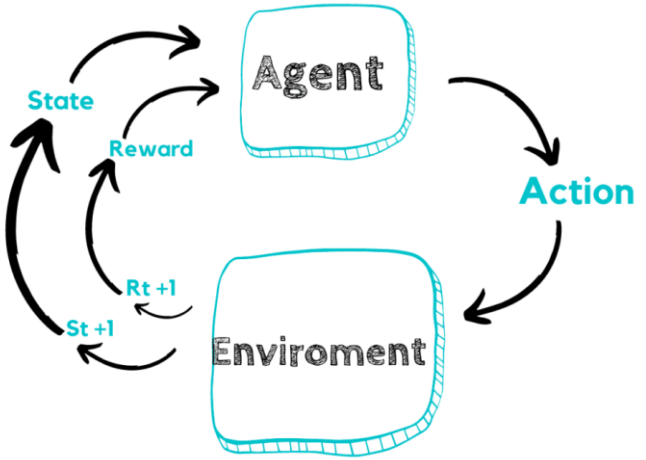


https://businessolution.org/gpt-3-statistics/

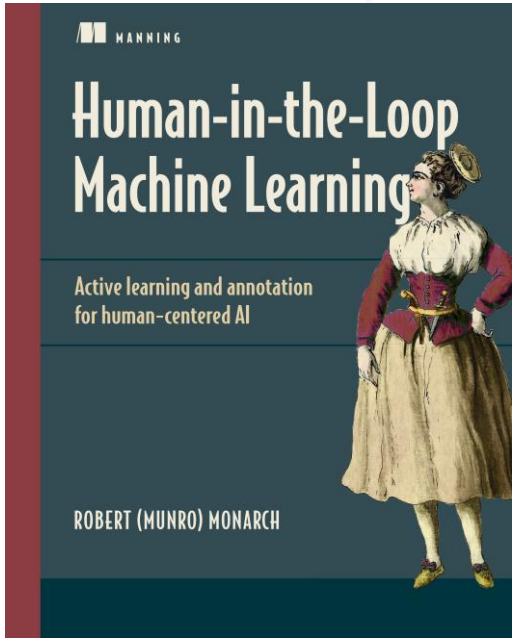
ChatGPT



+



Reinforcement Learning



Reinforcement Learning from Human Feedback (RLHF)

Quanto Custa os Modelos GPTs

LLM Training Costs on MosaicML Cloud			
Model	Billions of Tokens (Compute-optimal)	Days to Train on MosaicML Cloud	Approx. Cost on MosaicML Cloud
GPT-1.3B	26B	0.14	$2,000
GPT-2.7B	54B	0.48	$6,000
GPT-6.7B	134B	2.32	$30,000
GPT-13B	260B	7.43	$100,000
GPT-30B *	**610B**	**35.98**	**$450,000**
GPT-70B **	1400B	176.55	$2,500,000

original GPT-3, which had more parameters (175 billion params) but was trained on less data (300 billion tokens).

https://www.mosaicml.com/blog/gpt-3-quality-for-500k#:~:text=The%20bottom%20line%3A%20it%20costs,10x%20less%20than%20people%20think.

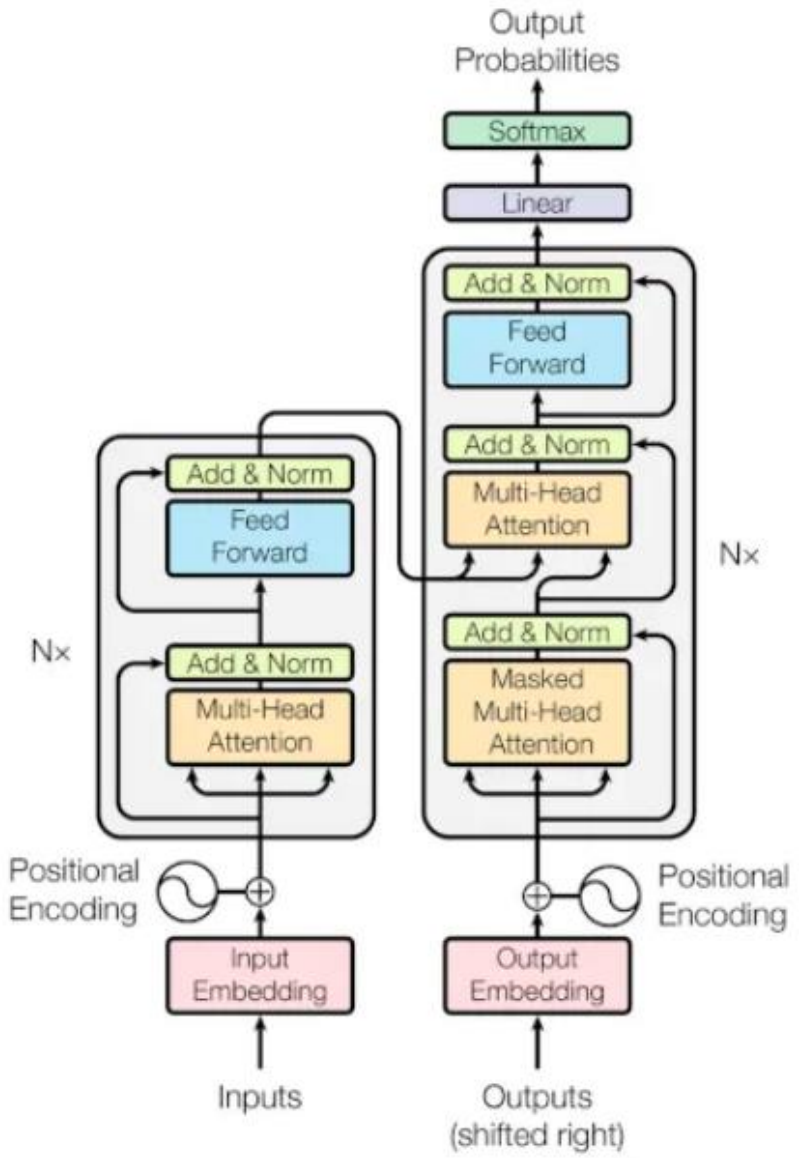
Attention Is All You Need

Ashish Vaswani* Google Brain avaswani@google.com
Noam Shazeer* Google Brain noam@google.com
Niki Parmar* Google Research nikip@google.com
Jakob Uszkoreit* Google Research usz@google.com
Llion Jones* Google Research llion@google.com
Aidan N. Gomez* † University of Toronto aidan@cs.toronto.edu
Łukasz Kaiser* Google Brain lukaszkaiser@google.com
Illia Polosukhin* ‡ illia.polosukhin@gmail.com

Abstract

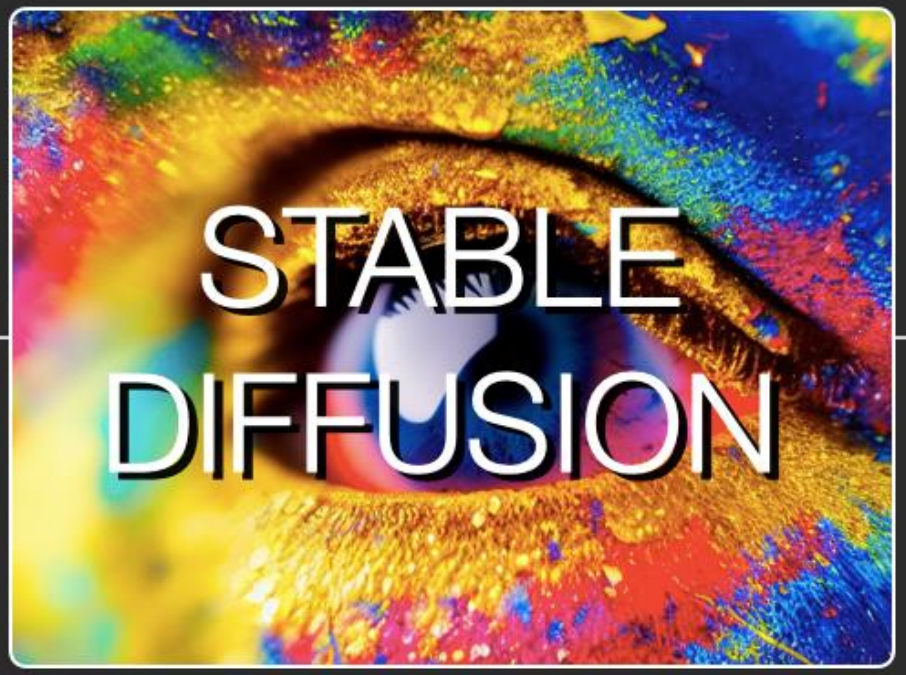
The dominant sequence transduction models are based on complex recurrent or convolutional neural networks that include an encoder and a decoder. The best performing models also connect the encoder and decoder through an attention mechanism. We propose a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely. Experiments on two machine translation tasks show these models to be superior in quality while being more parallelizable and requiring significantly less time to train. Our model achieves 28.4 BLEU on the WMT 2014 English-to-German translation task, improving over the existing best results, including ensembles, by over 2 BLEU. On the WMT 2014 English-to-French translation task,

https://arxiv.org/abs/1706.03762



Output Probabilities

Pirate ship

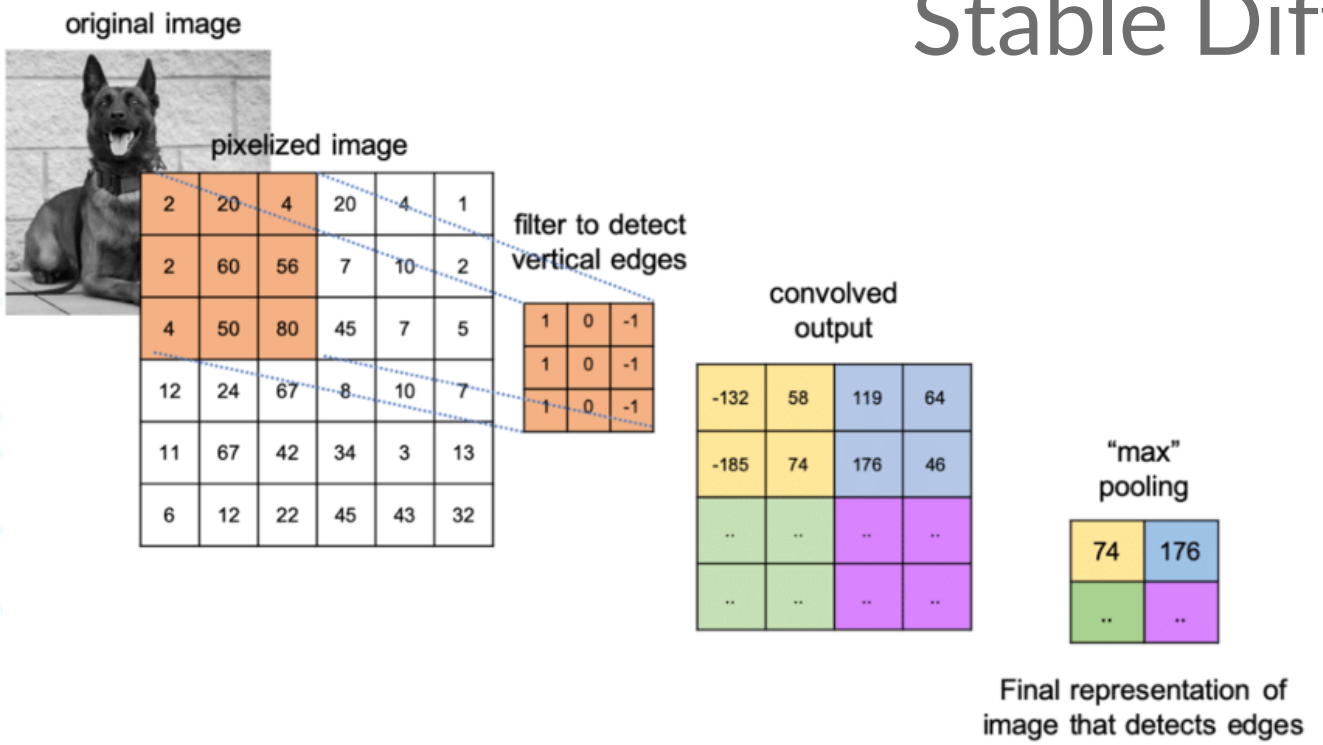


Inputs

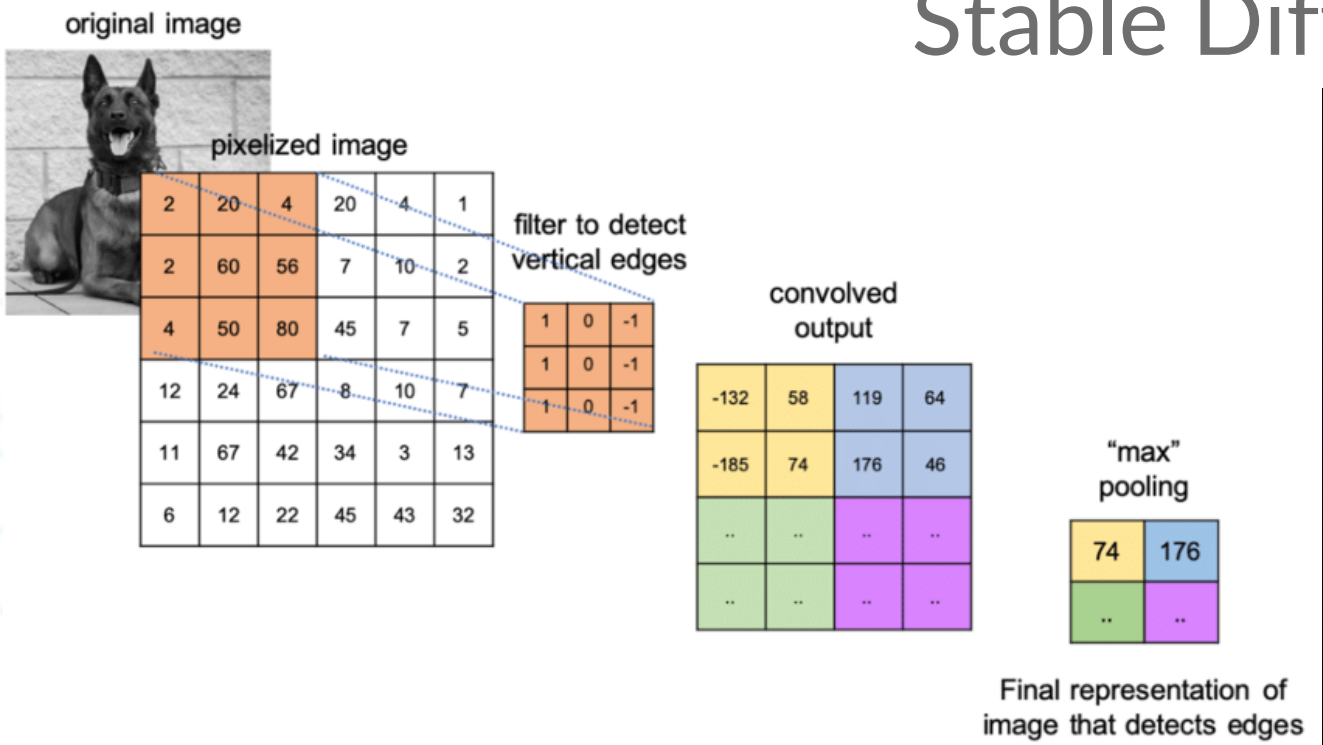
Outputs (shifted right)

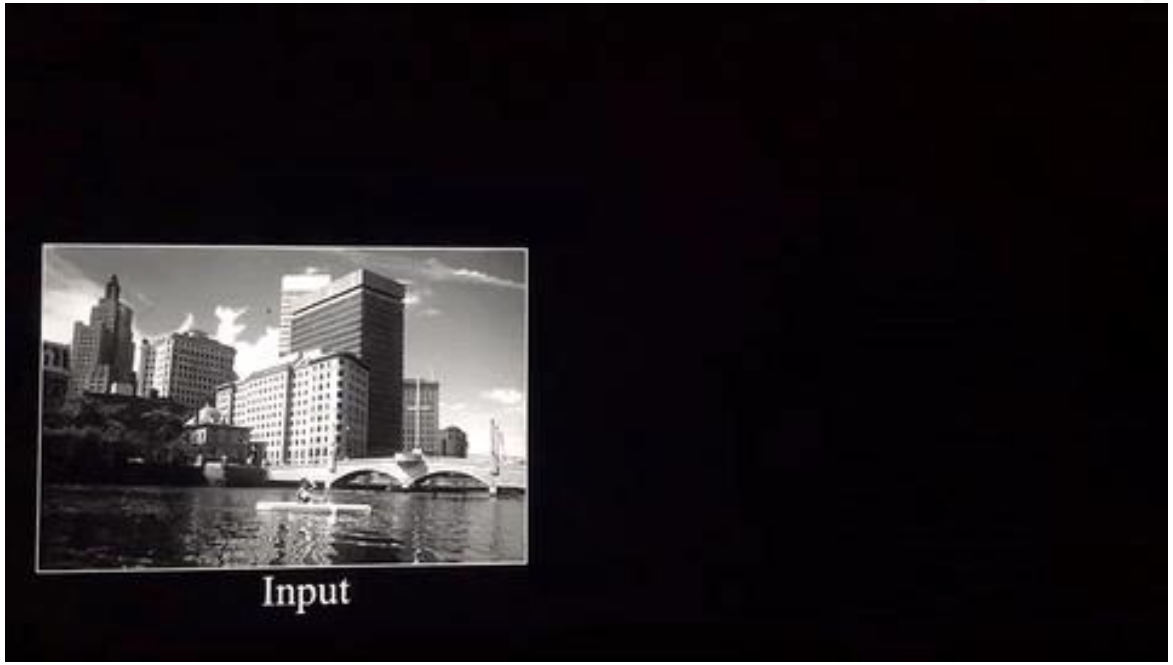
https://arxiv.org/abs/1706.03762

Stable Diffusion

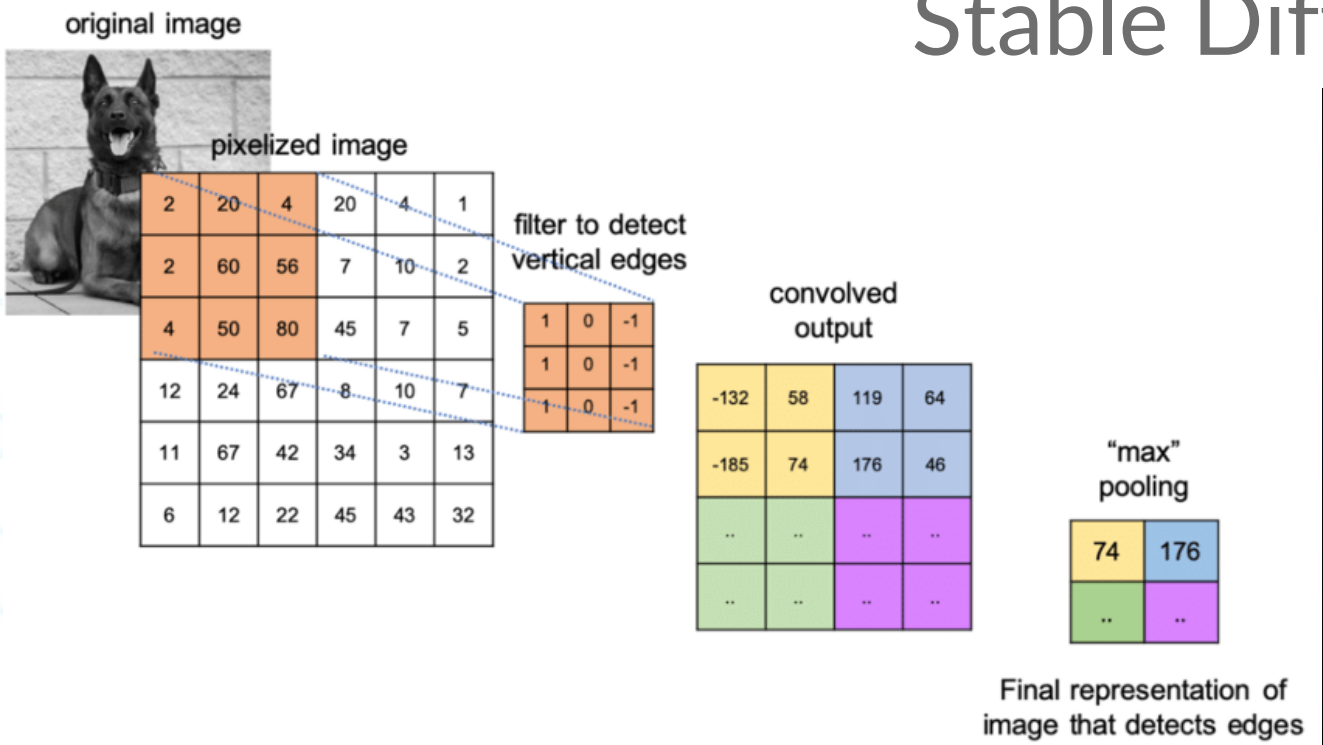


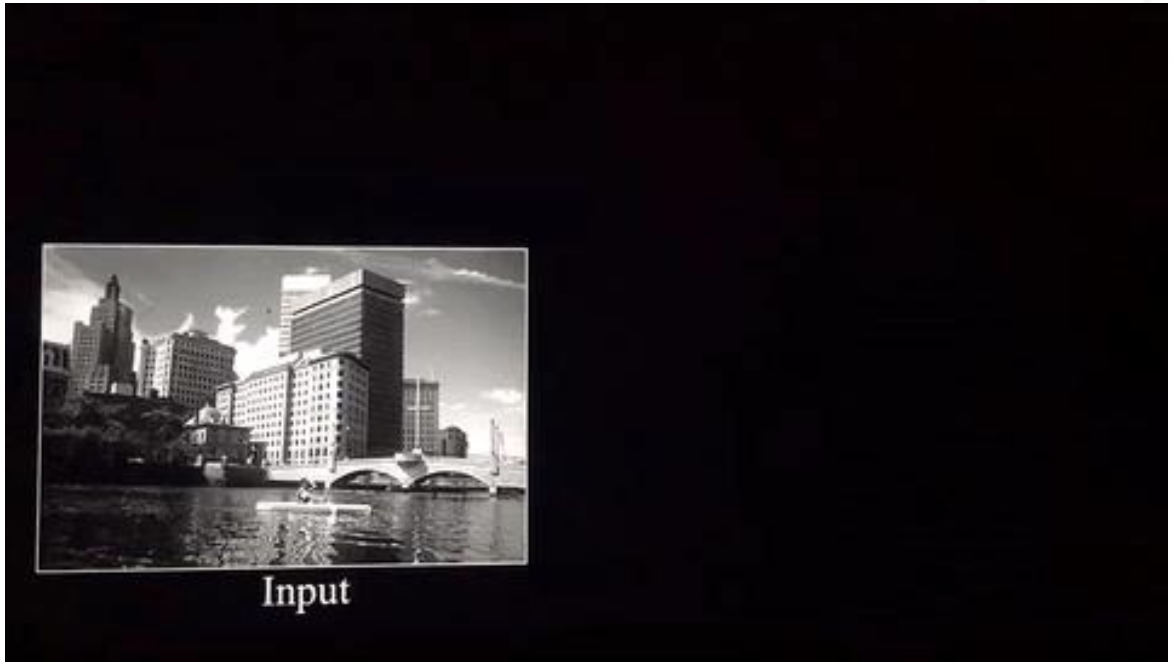
Stable Diffusion





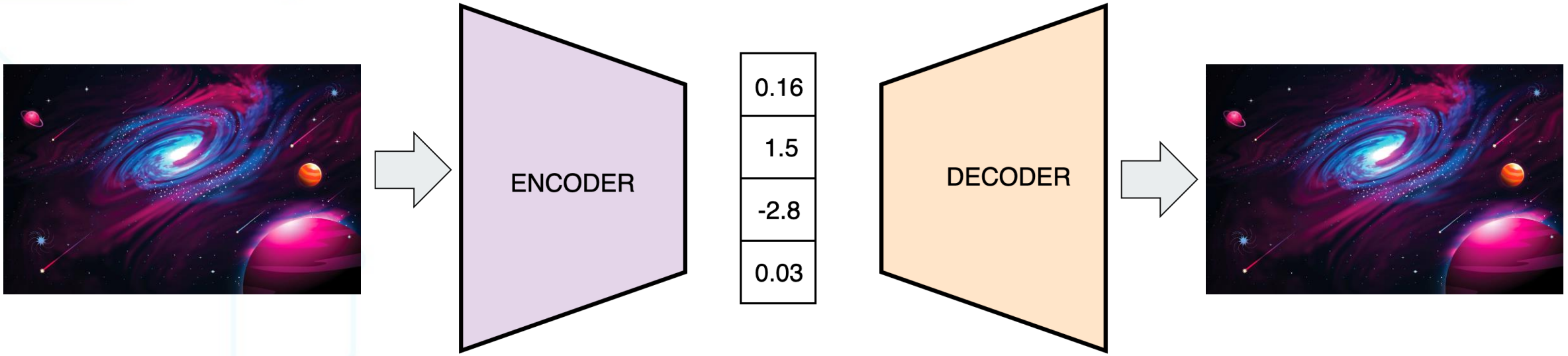
Stable Diffusion





Stable Diffusion

Espaço latente



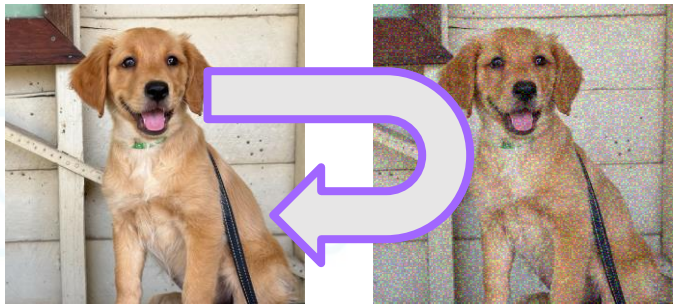
Modelos de difusão

Diffusion model

Adds noise and learns how to work backwards to the original image.

Modelos de difusão

Diffusion model



Adds noise and learns how to work backwards to the original image.

Modelos de difusão

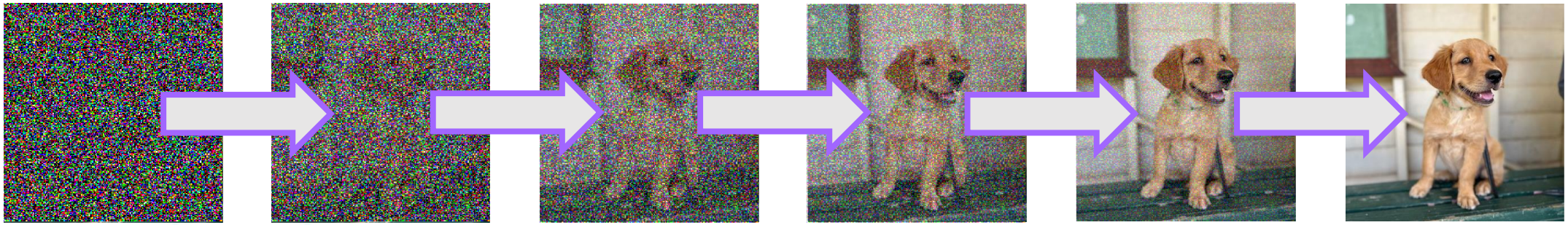
Diffusion model

Adds noise and learns how to work backwards to the original image.

Modelos de difusão

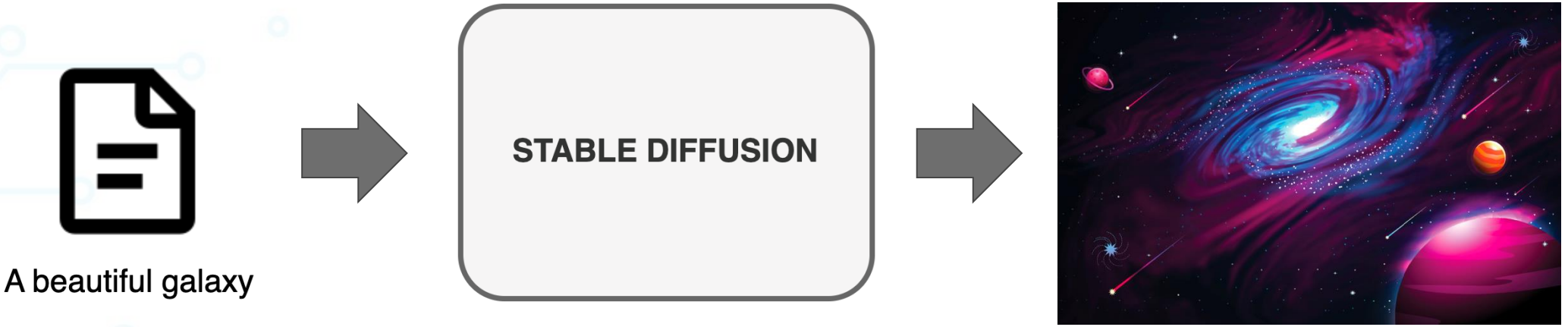
Diffusion model

Adds noise and learns how to work backwards to the original image.

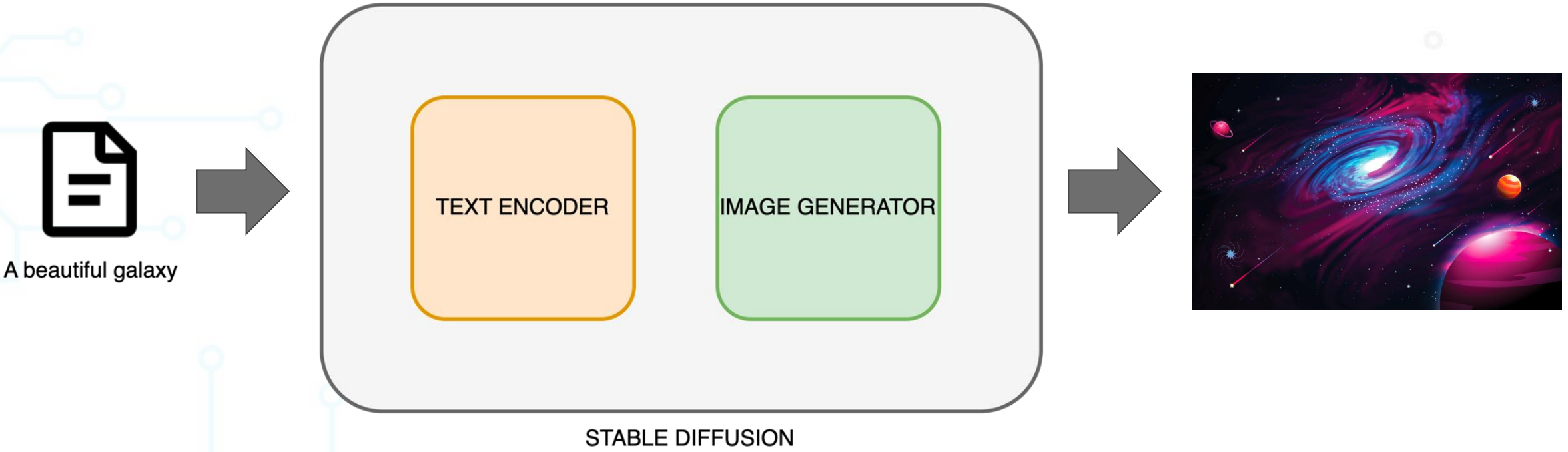


Trained model works from random noise to generated an image.

Stable Diffusion

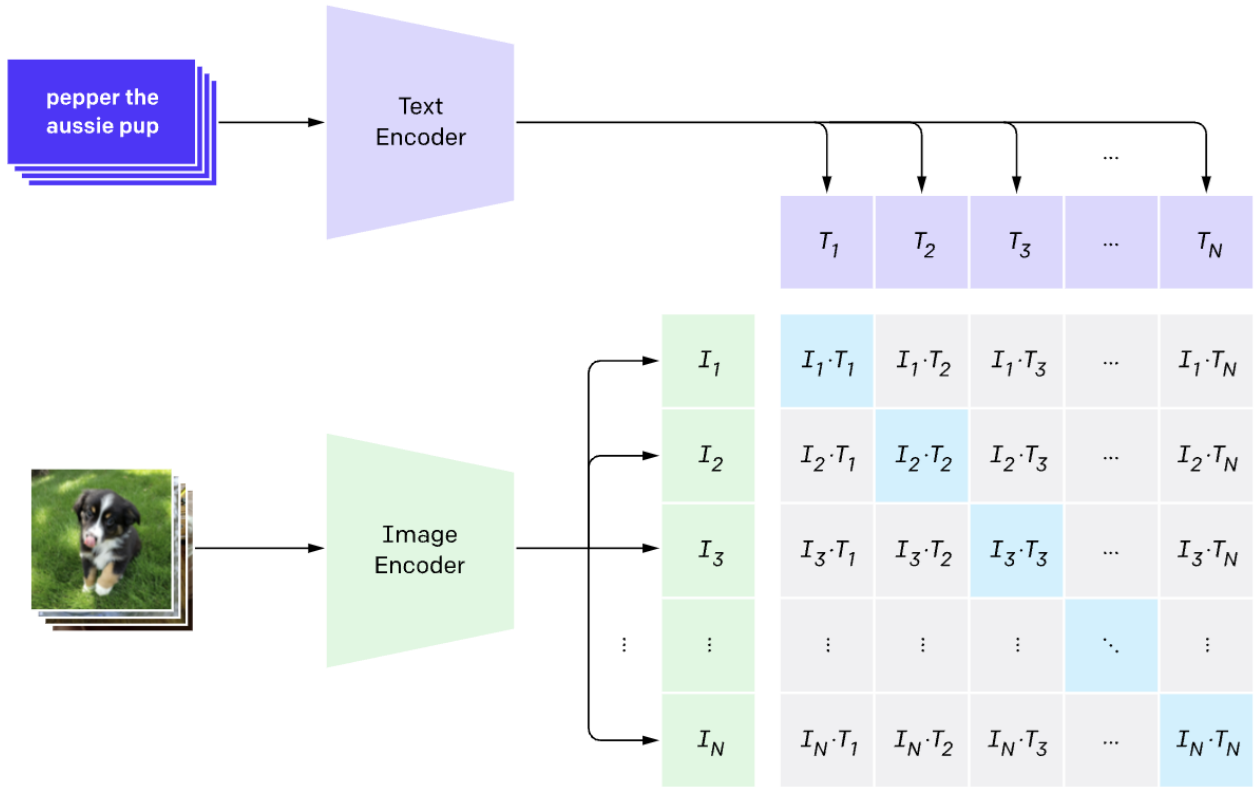


Stable Diffusion



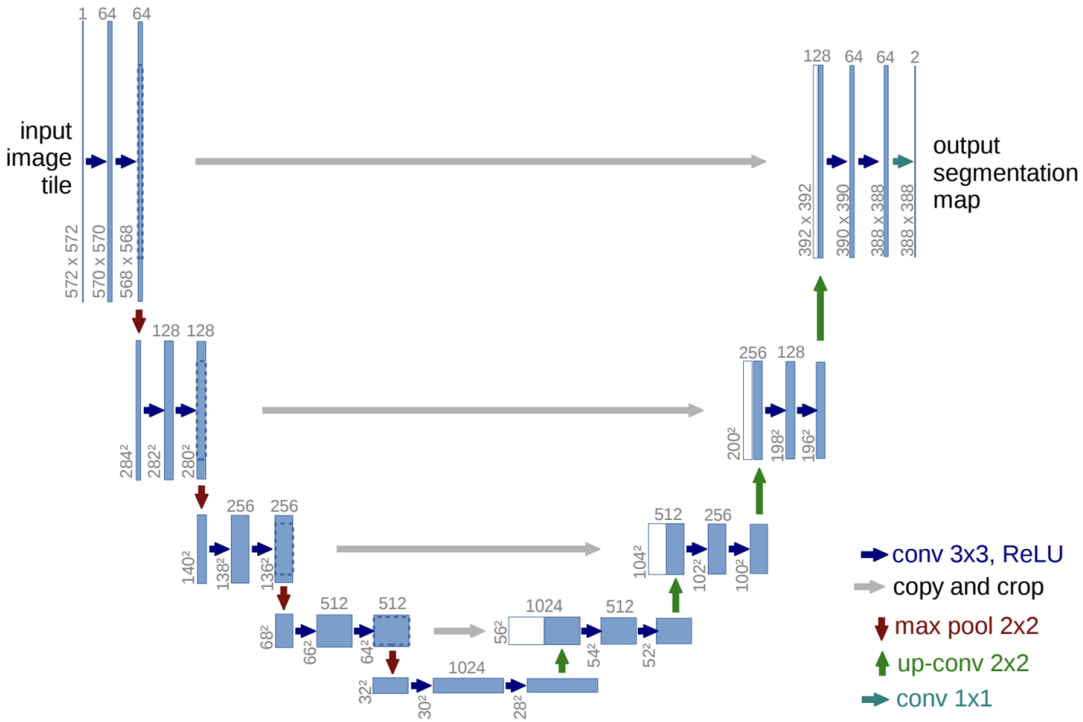
Stable Diffusion

Text Encoder - CLIP



Stable Diffusion

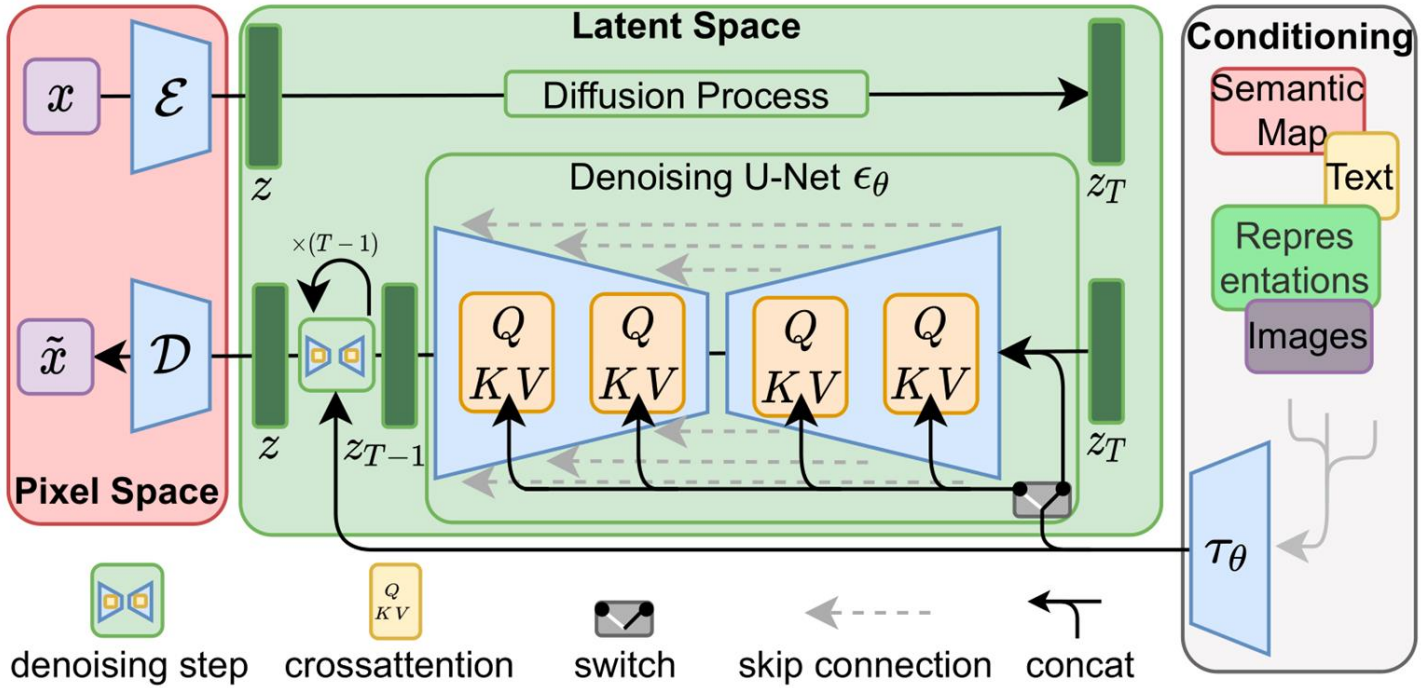
Preditor de ruído - UNet Condicionada



Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation

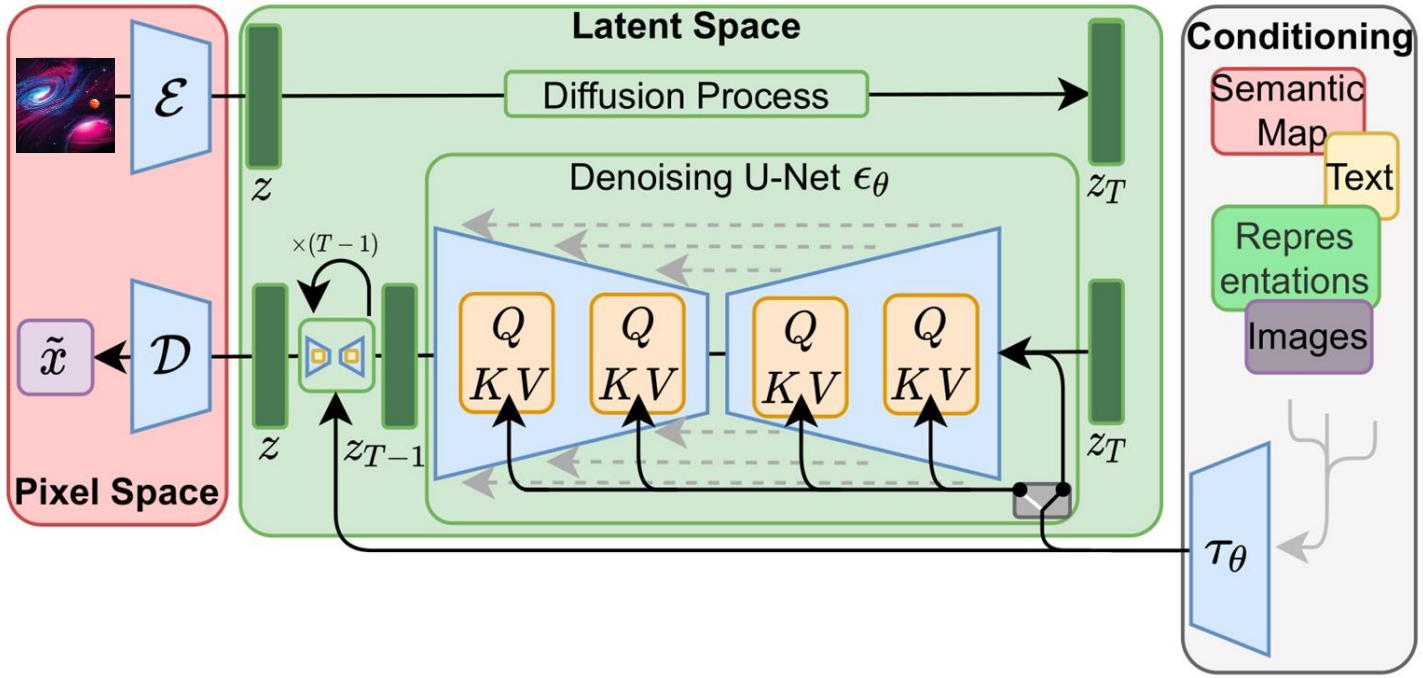
Stable Diffusion

Treinamento



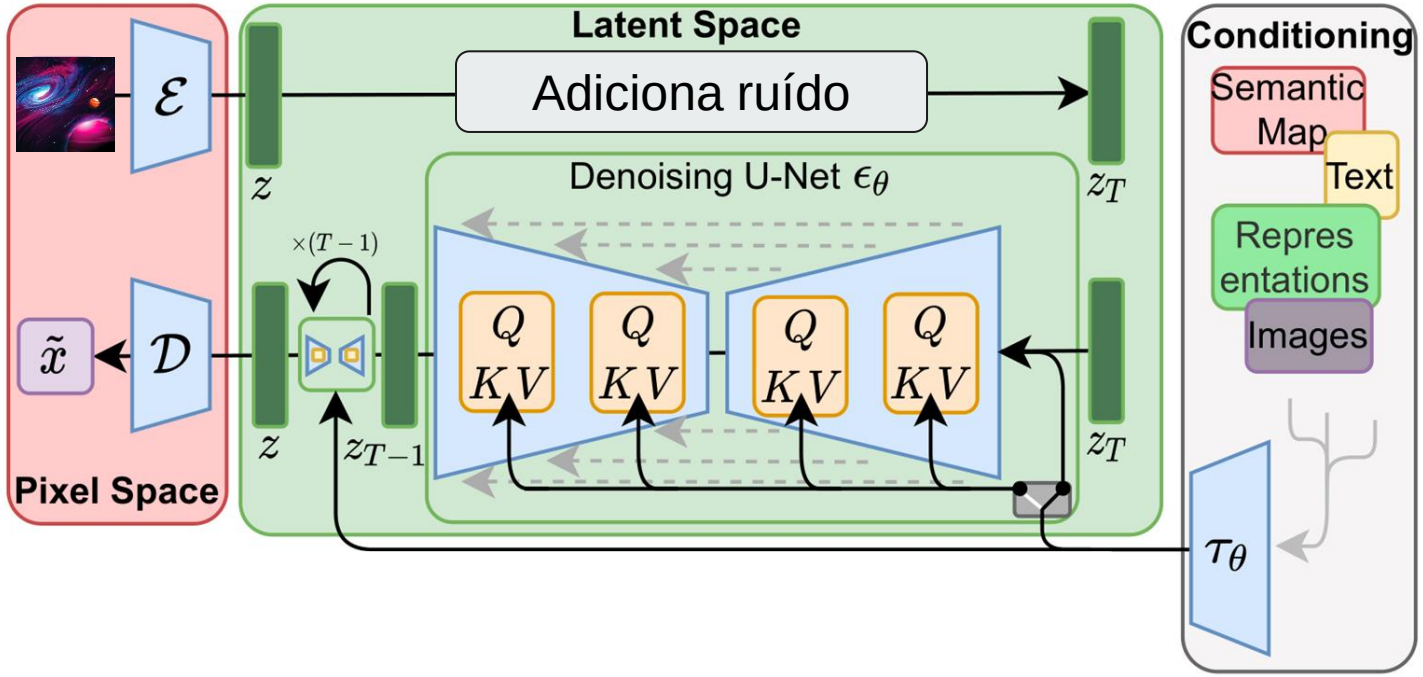
Stable Diffusion

Treinamento



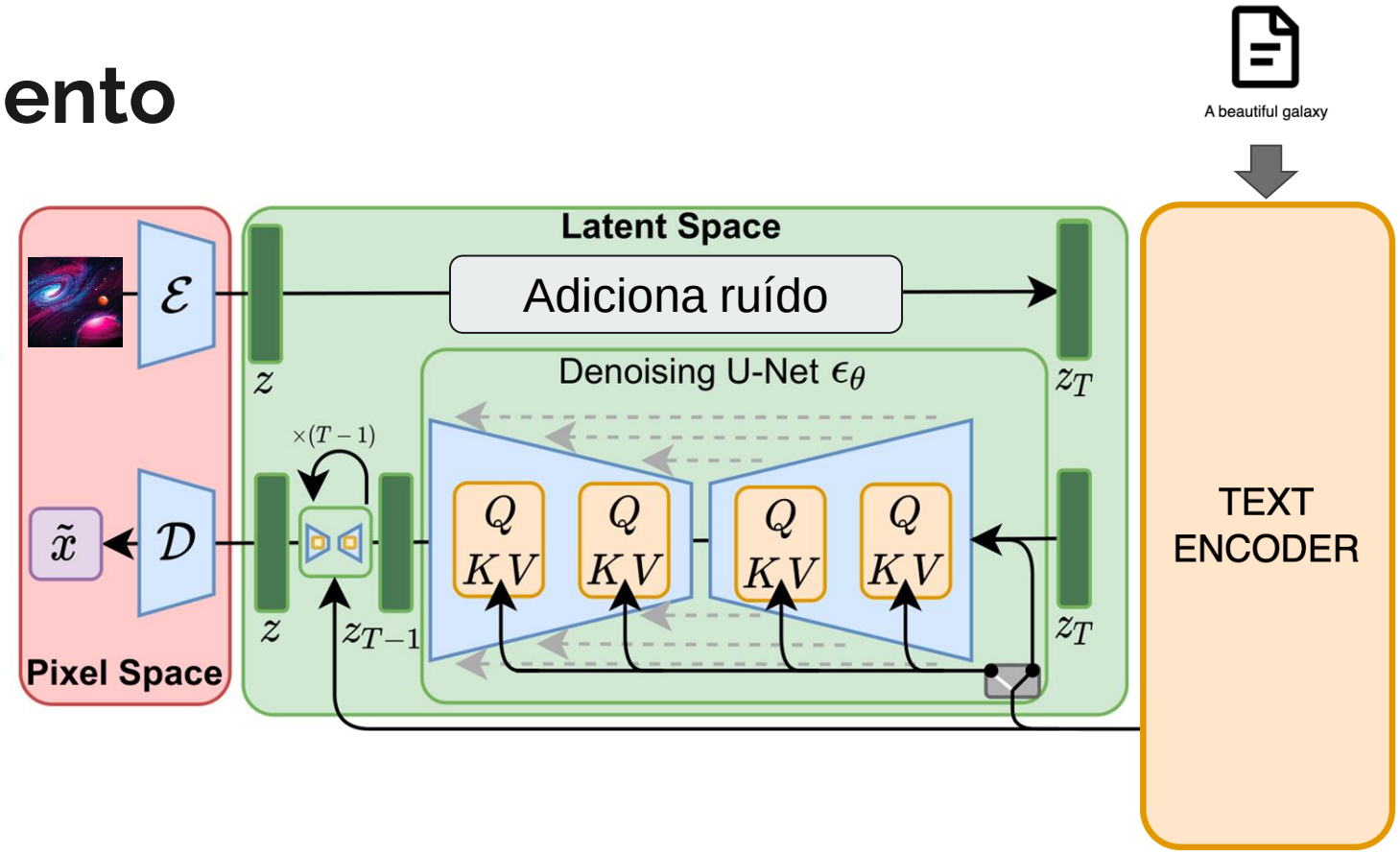
Stable Diffusion

Treinamento



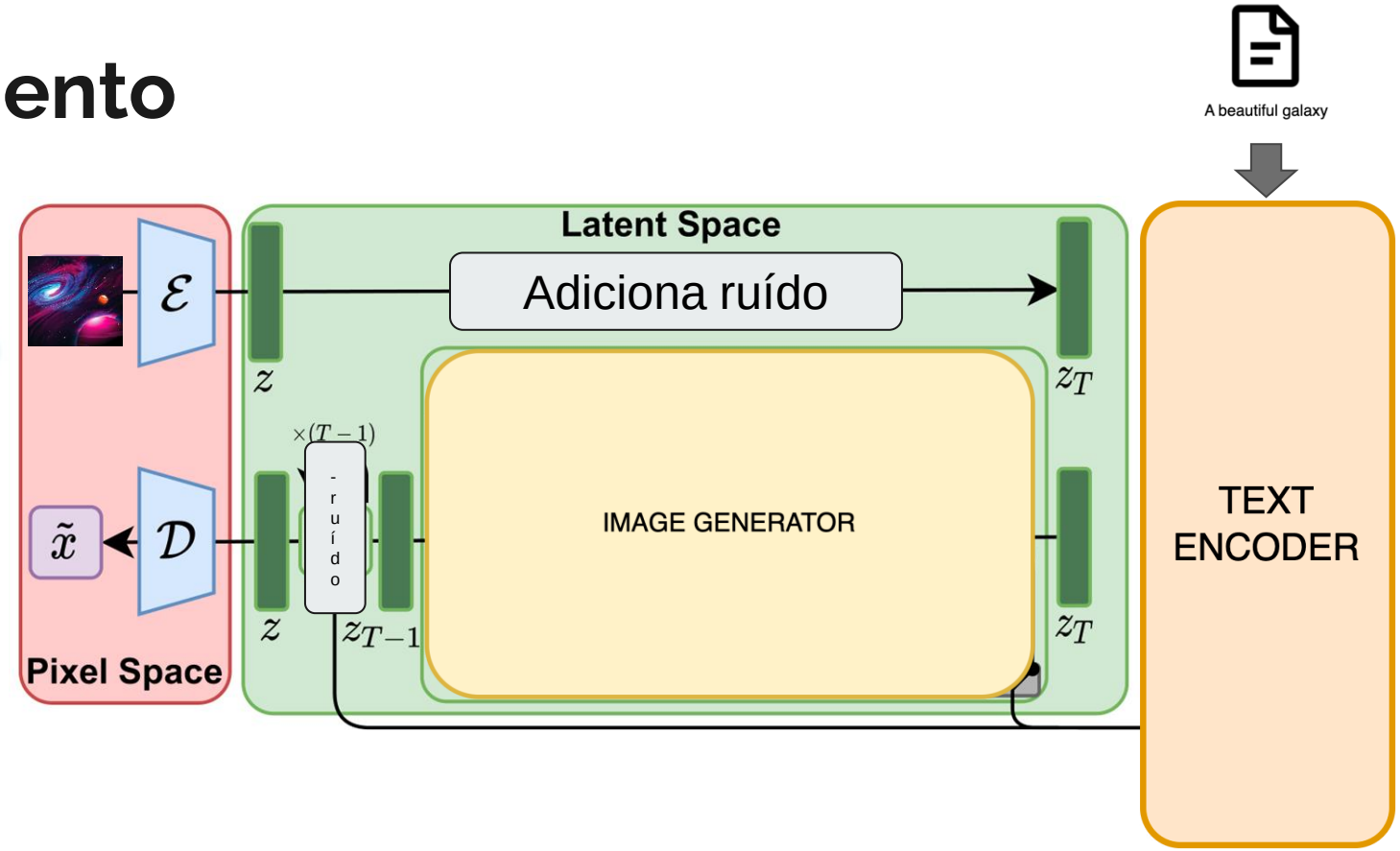
Stable Diffusion

Treinamento



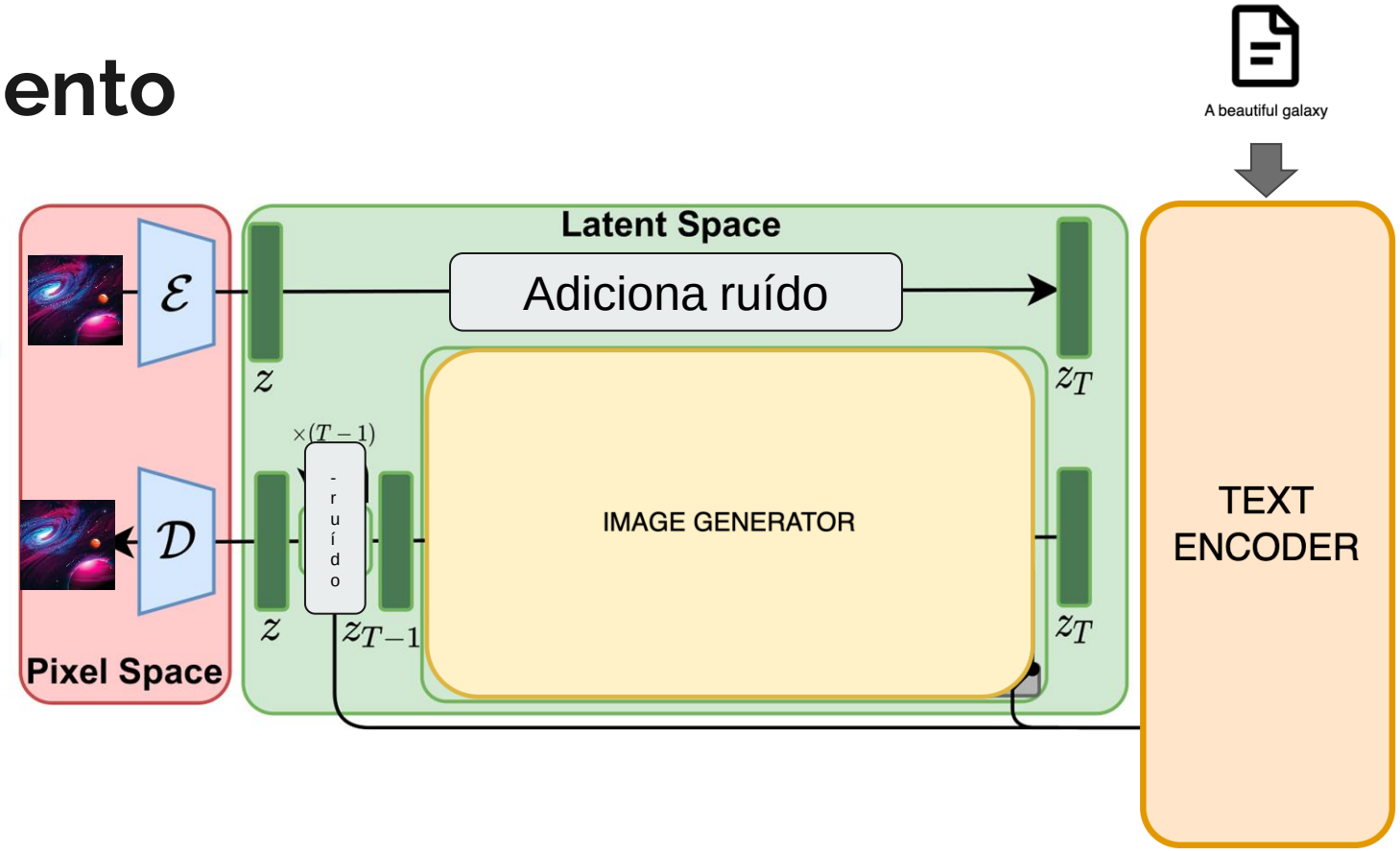
Stable Diffusion

Treinamento



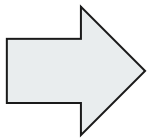
Stable Diffusion

Treinamento

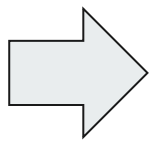


Textual Inversion

A beautiful galaxy

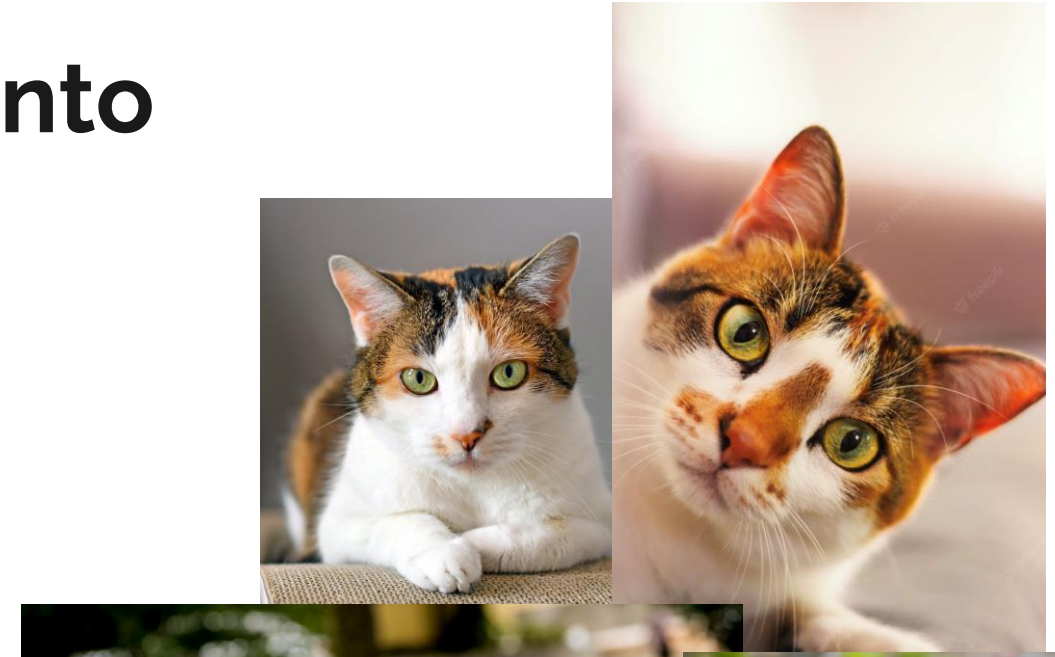


STABLE DIFFUSION



Textual Inversion

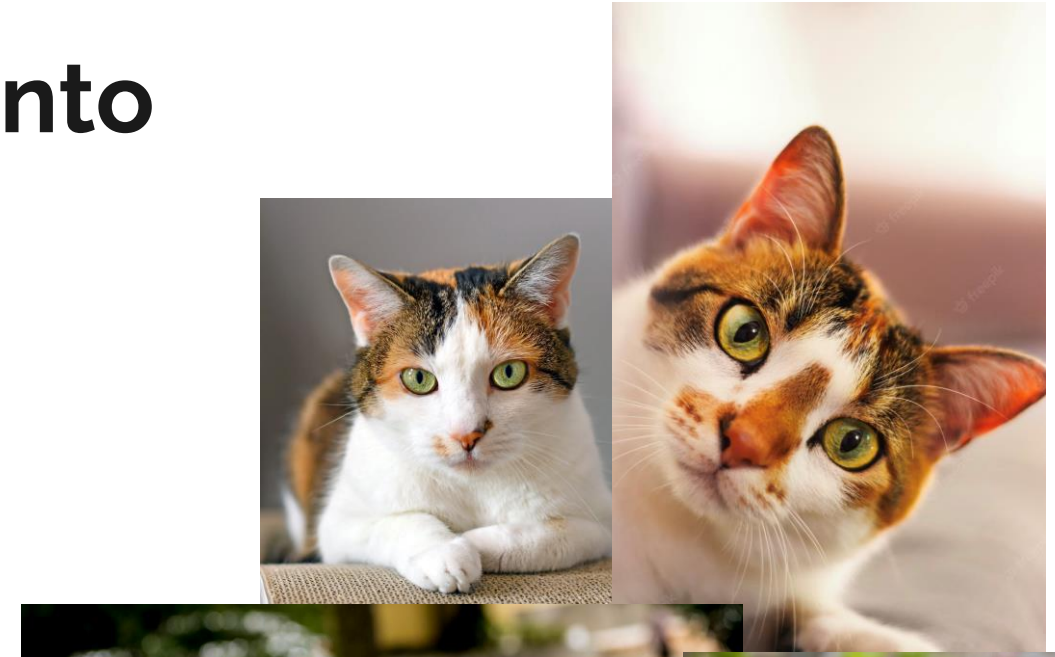
Dados fora do treinamento



Textual Inversion

Dados fora do treinamento

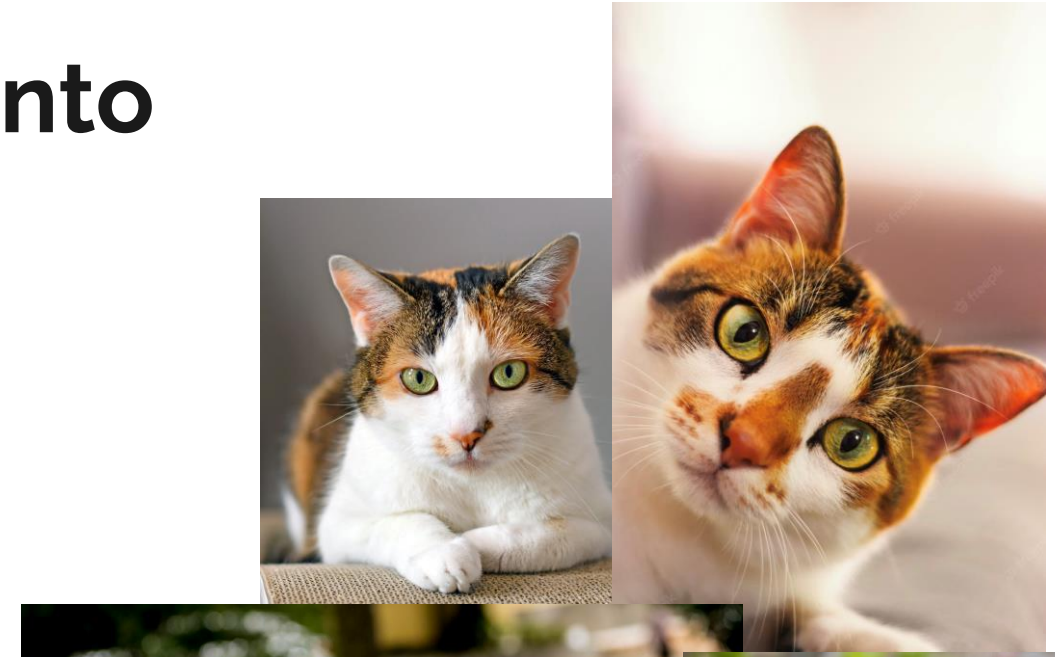
- Um gato em uma praça ✅



Textual Inversion

Dados fora do treinamento

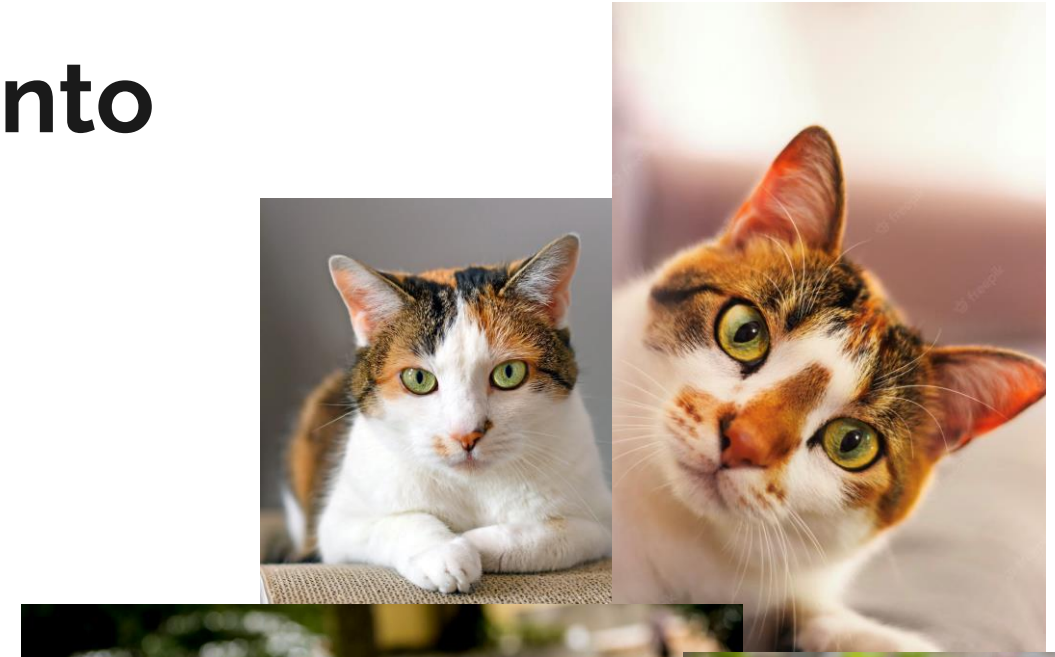
- Um gato em uma praça ✅
- Um cachorro na grama ✅



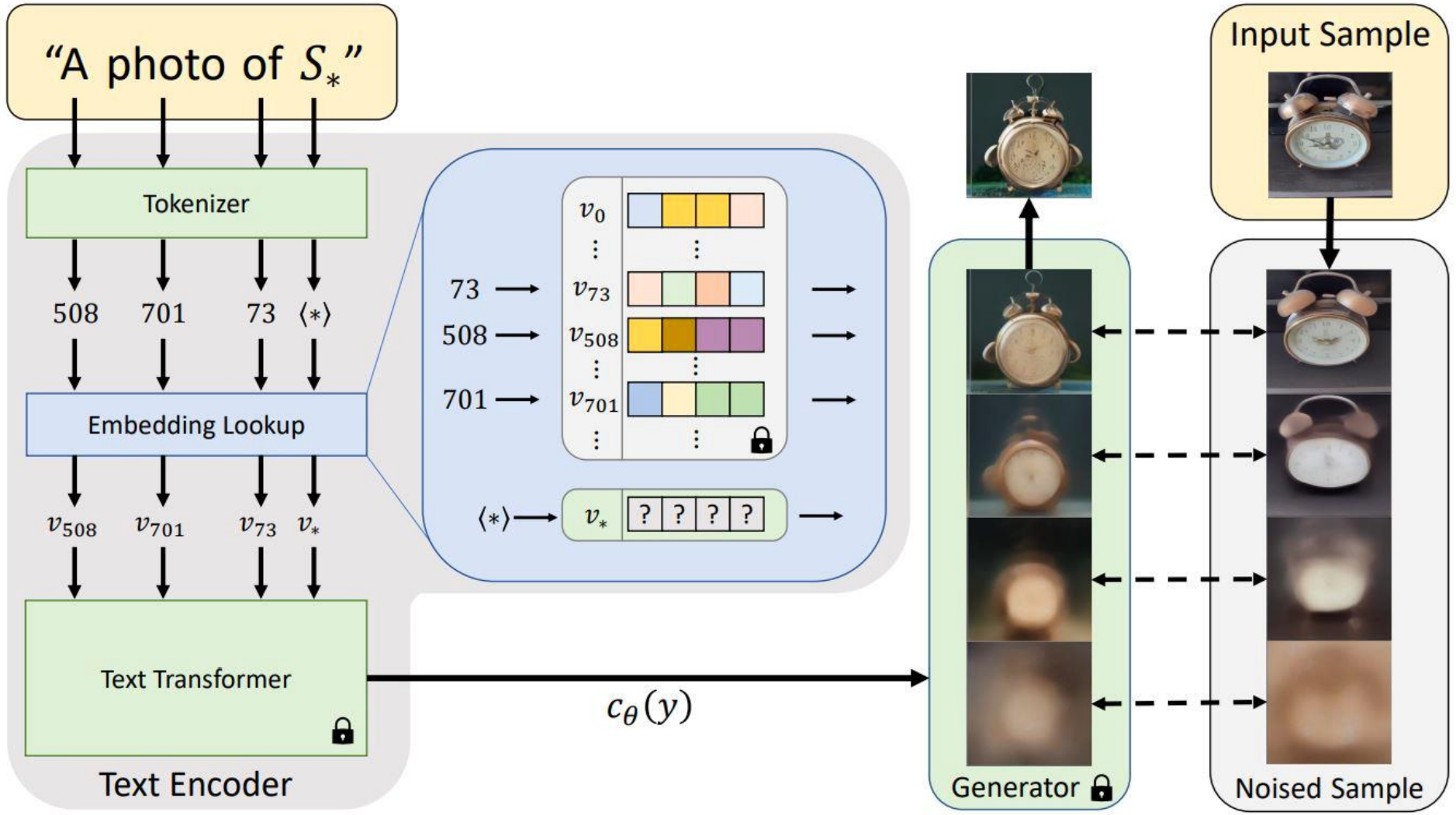
Textual Inversion

Dados fora do treinamento

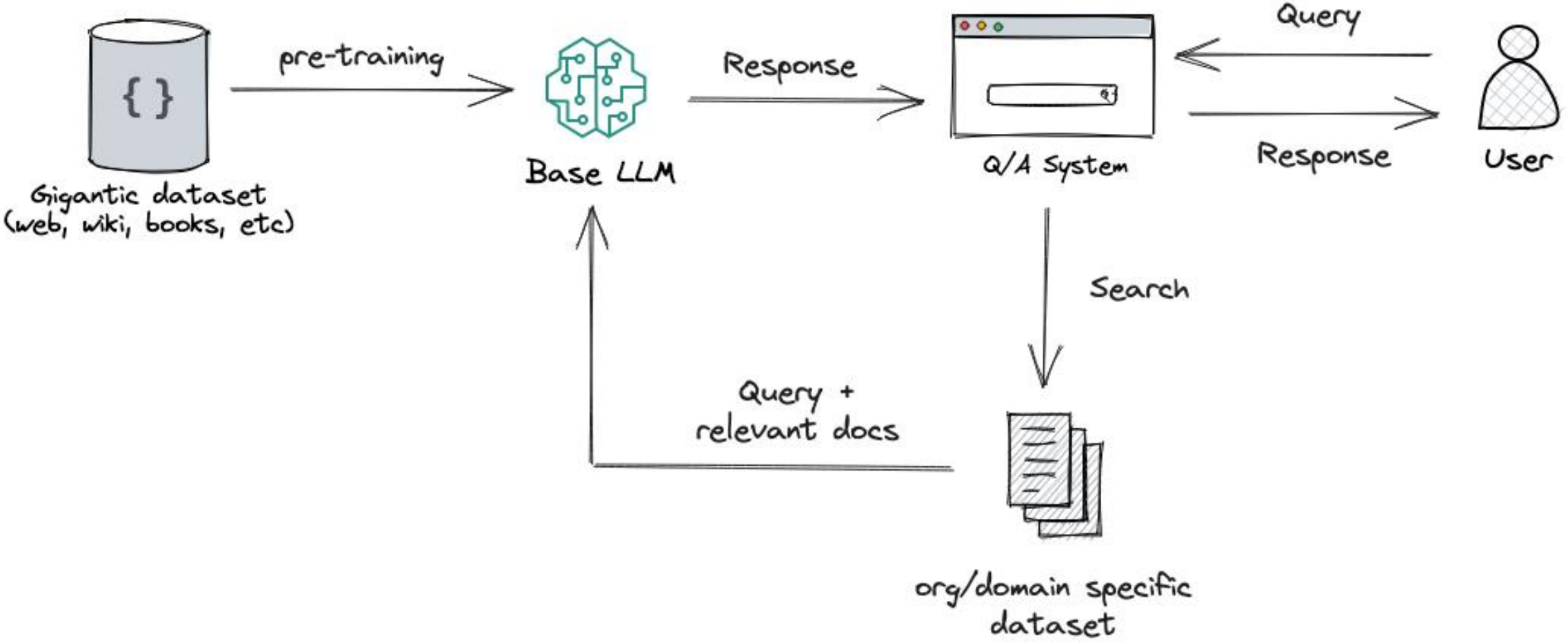
- Um gato em uma praça ✅
- Um cachorro na grama ✅
- Um coelho no mato ✖



Textual Inversion



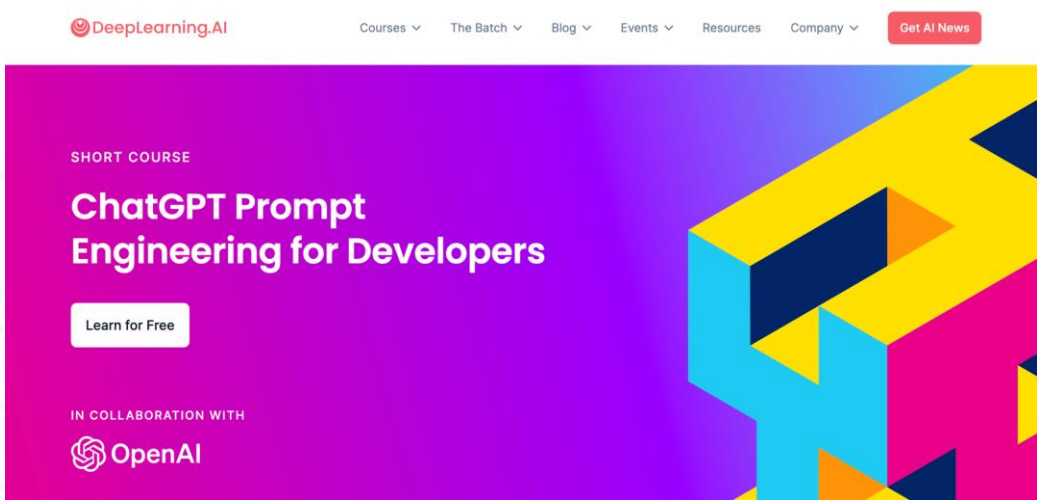
Pipeline GenAI



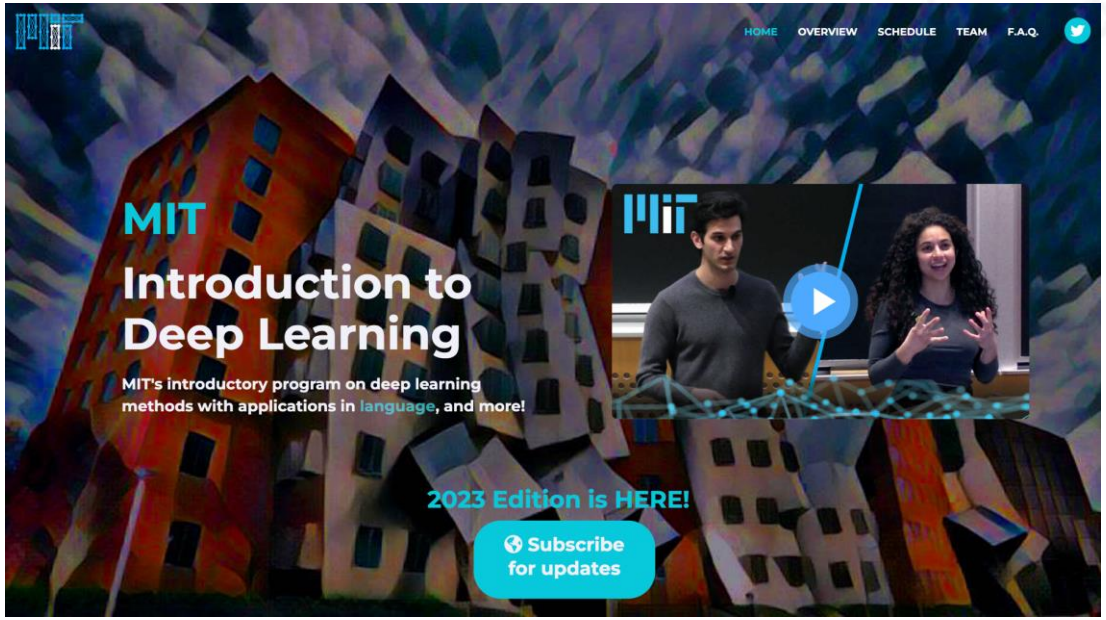
ML Life Cycle

VINICARIDÁ
Human-Centered Data & AI

AI Generativa

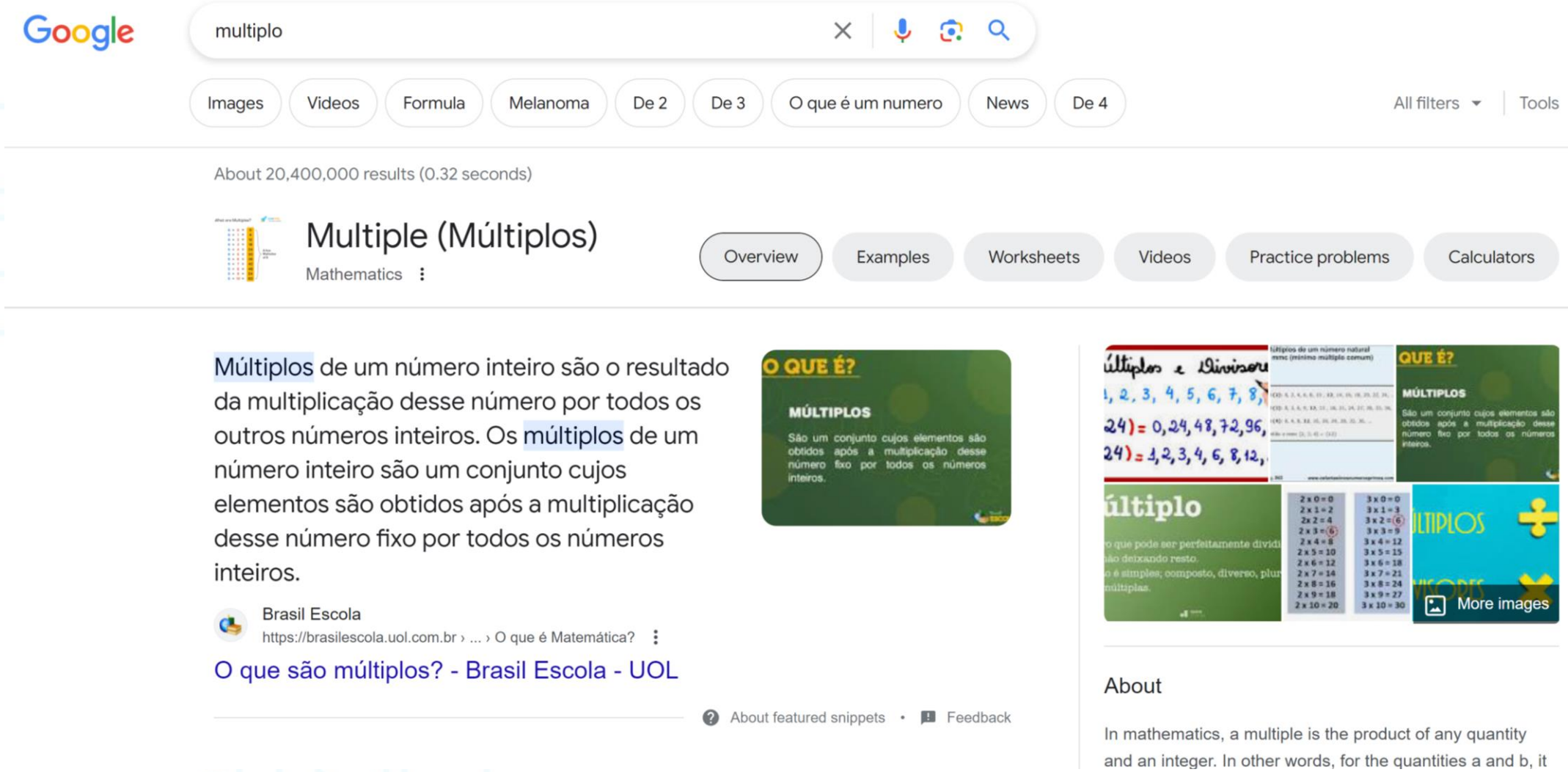


https://www.deeplearning.ai/short-courses/chatgpt-prompt-engineering-for-developers/

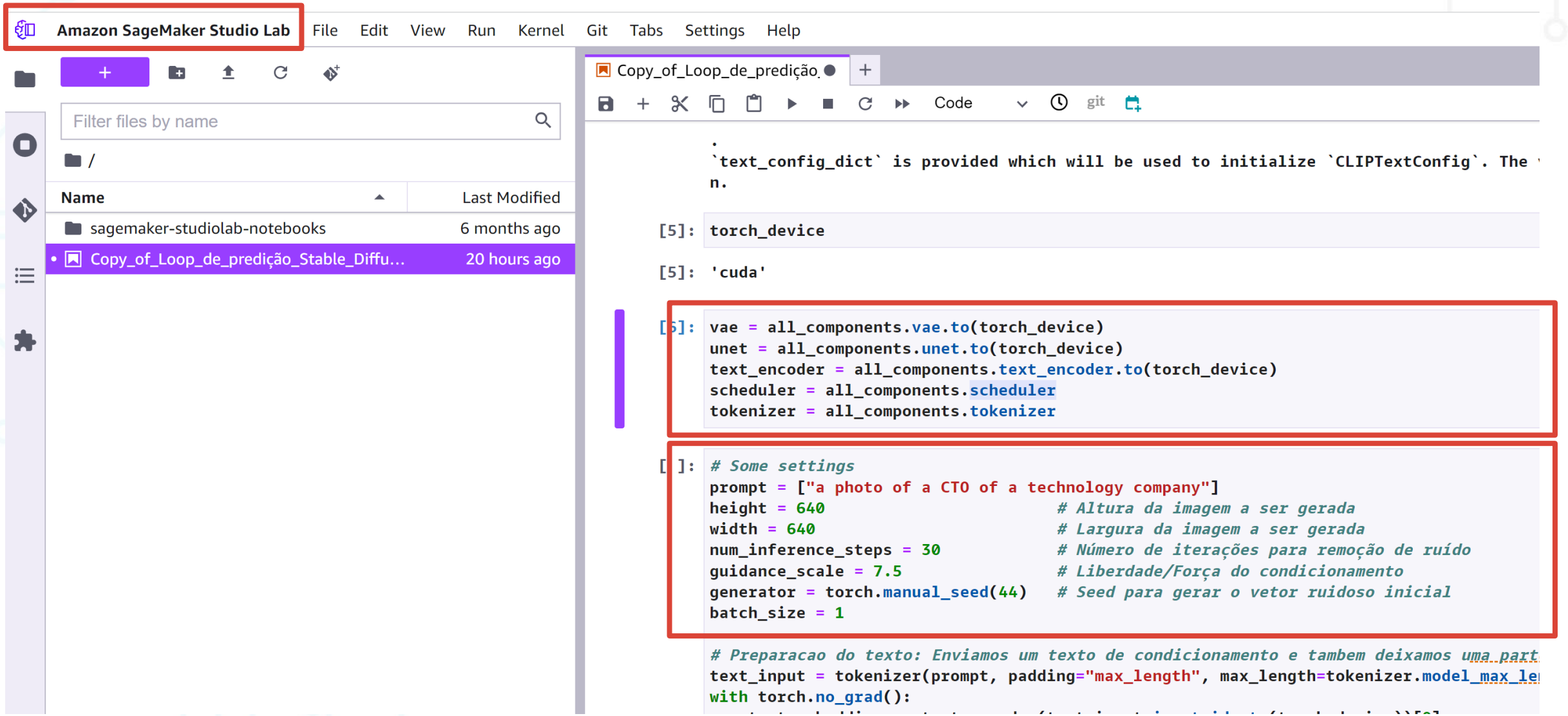


http://introtodeeplearning.com/

Textual Inversion

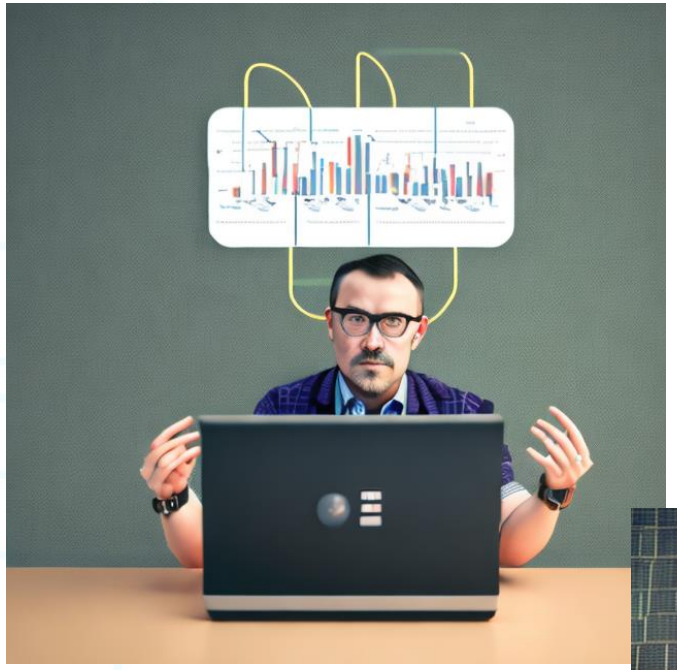


Código

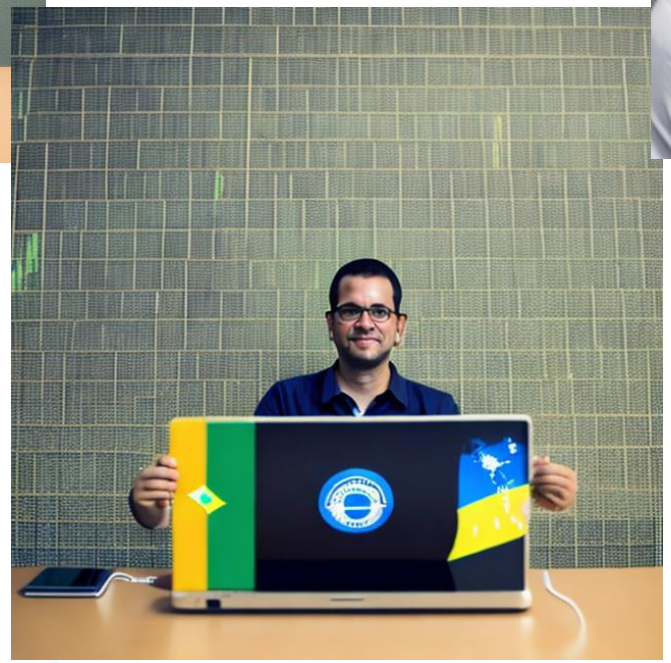

Amazon SageMaker Studio Lab
File
Edit
View
Run
Kernel
Git
Tabs
Settings
Help
Filter files by name
/
Name
Last Modified
sagemaker-studiolab-notebooks
6 months ago
Copy_of_Loop_de_predição_Stable_Diffu...
20 hours ago
Copy_of_Loop_de_predição
Code
git
`text_config_dict` is provided which will be used to initialize `CLIPTextConfig`. The
n.
[5]: torch_device
[5]: 'cuda'
[6]: vae = all_components.vae.to(torch_device)
unet = all_components.unet.to(torch_device)
text_encoder = all_components.text_encoder.to(torch_device)
scheduler = all_components.scheduler
tokenizer = all_components.tokenizer
[ ]: # Some settings
prompt = ["a photo of a CTO of a technology company"]
height = 640 # Altura da imagem a ser gerada
width = 640 # Largura da imagem a ser gerada
num_inference_steps = 30 # Número de iterações para remoção de ruído
guidance_scale = 7.5 # Liberdade/Força do condicionamento
generator = torch.manual_seed(44) # Seed para gerar o vetor ruidoso inicial
batch_size = 1
# Preparacao do texto: Enviamos um texto de condicionamento e tambem deixamos uma part
text_input = tokenizer(prompt, padding="max_length", max_length=tokenizer.model_max_le
with torch.no_grad():

Demo

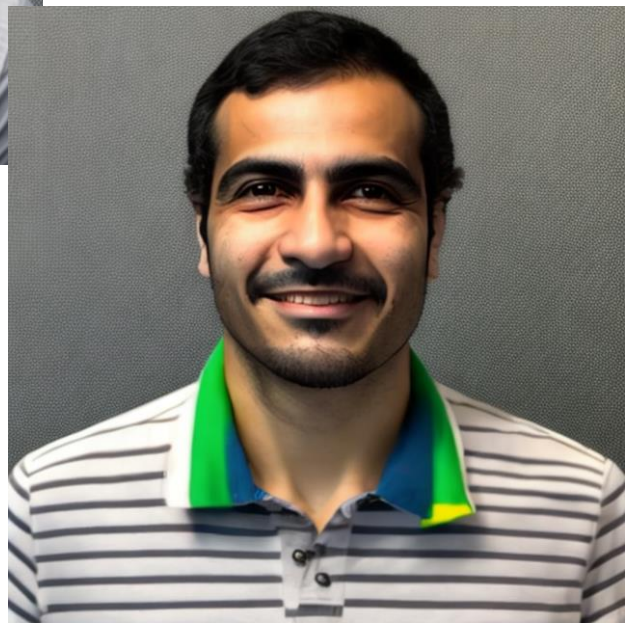
a photo of a CTO of a technology company



a photo of a CTO of a technology company in Brazil who is male, white and blonde



a photo of a CTO of a technology company in Brazil



a photo of Evandro Pires, who is the CTO of a technology company in Brazil

Demo

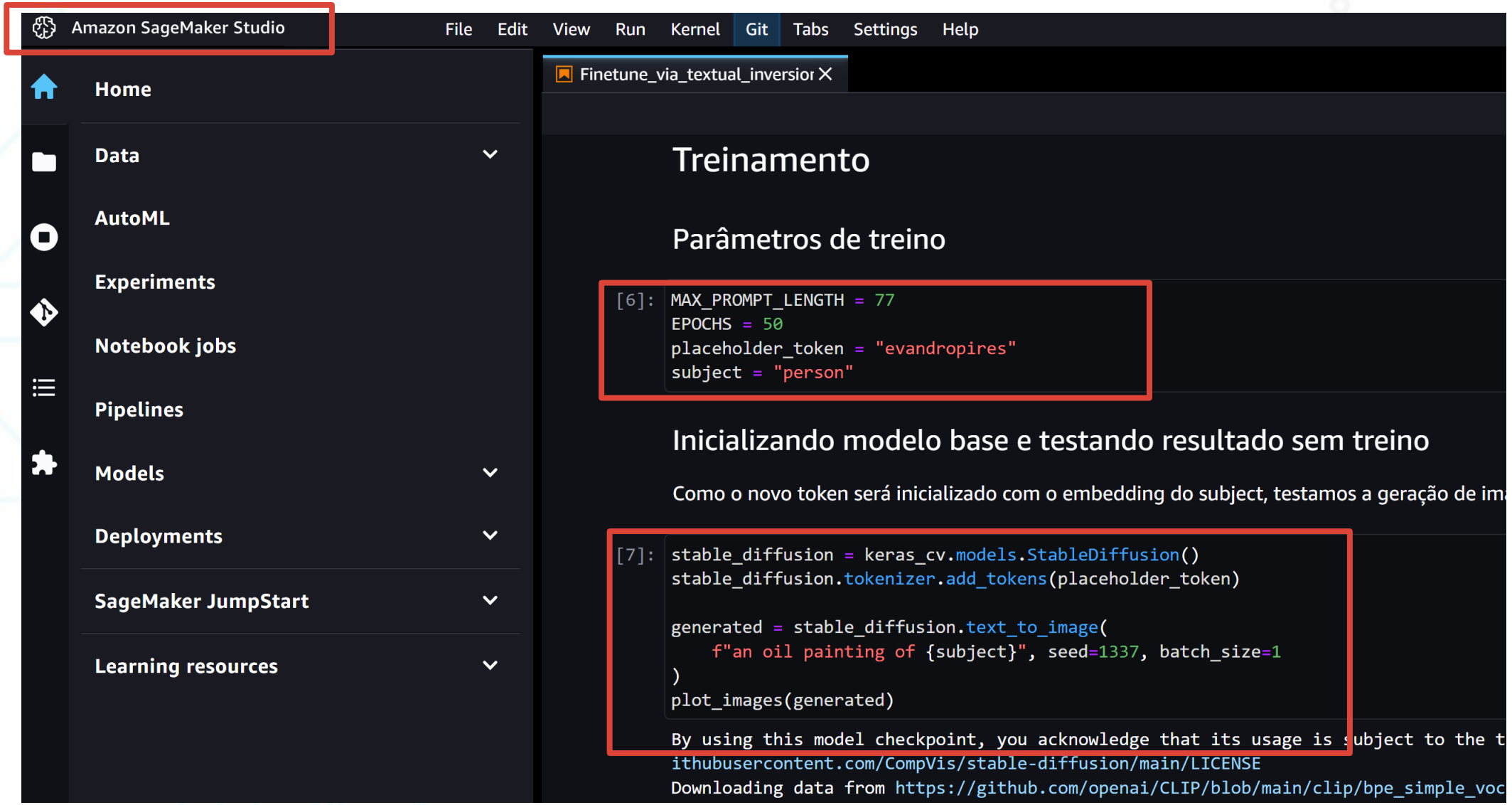








Código


Amazon SageMaker Studio
File
Edit
View
Run
Kernel
Git
Tabs
Settings
Help
Home
Data
AutoML
Experiments
Notebook jobs
Pipelines
Models
Deployments
SageMaker JumpStart
Learning resources
Finetune_via_textual_inversior
Treinamento
Parâmetros de treino
[6]:
MAX_PROMPT_LENGTH = 77
EPOCHS = 50
placeholder_token = "evandropires"
subject = "person"
Inicializando modelo base e testando resultado sem treino
Como o novo token será inicializado com o embedding do subject, testamos a geração de im
[7]:
stable_diffusion = keras_cv.models.StableDiffusion()
stable_diffusion.tokenizer.add_tokens(placeholder_token)
generated = stable_diffusion.text_to_image(
f"an oil painting of {subject}", seed=1337, batch_size=1
)
plot_images(generated)
By using this model checkpoint, you acknowledge that its usage is subject to the t
ithubusercontent.com/CompVis/stable-diffusion/main/LICENSE
Downloading data from https://github.com/openai/CLIP/blob/main/clip/bpe_simple_voc

Código

VINI CARIDÁ
Human-Centered Data & AI

```
[8]: train_ds = assemble_dataset(
        file_paths=["evandro1.jpg",
                    "evandro2.jpg",
                    "evandro3.jpg",
                    "evandro4.jpg",
                    "evandro5.png",
                    "evandro6.png",
                    "evandro7.png",
                    "evandro8.png",
                    "evandro9.png"],
        placeholder_token=placeholder_token,
        prompts=[
            "a photo of a {}",
            "a rendering of a {}",
            "a cropped photo of the {}",
            "the photo of a {}",
            "a photo of a clean {}",
            "a photo of my {}",
            "a photo of the cool {}",
            "a close-up photo of a {}",
            "a bright photo of the {}",
            "a cropped photo of a {}",
            "a photo of the {}",
            "a good photo of the {}",
            "a photo of one {}",
            "a close-up photo of the {}",
            "a rendition of the {}",
            "a photo of the clean {}",
            "a rendition of a {}",
            "a photo of a nice {}",
            "a good photo of a {}",
            "a photo of the nice {}",
            "a photo of the small {}",
```

```
[18]: generated = trainer.stable_diffusion.text_to_image(
          f" a photo of {placeholder_token} who is the CTO of a technology company in Brazil",
          batch_size=3,
          num_steps=50
      )
      plot_images(generated)
```

Demo








Demo






VINICARIDÁ

Human-Centered Data & AI

@vinicius caridá

@vfcarida

vfcarida@gmail.com

https://linktr.ee/vfcarida