

# IA em ambiente empresarial

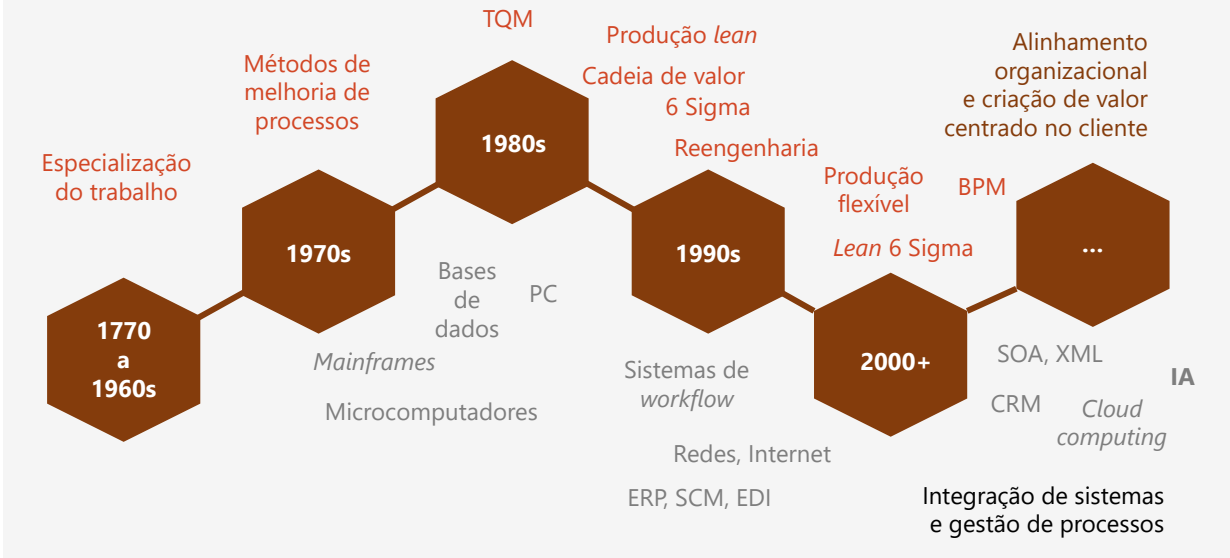
Gestão de Informação e dos Processos de Negócio



1

## Enquadramento

Aula 21  
**2**



2

### → **Diversidade de informação para melhores decisões**

A extração e interpretação de informação a partir de diferentes fontes de informação para **potenciar melhorias nos processos de decisão** é, ainda hoje, uma preocupação permanente na gestão



### → **Quantidade para melhores decisões**

Não só a diversidade da informação é importante, mas também a questão da **quantidade de informação** é também relevante.

A **IA** tem aqui um papel a desempenhar



## Enquadramento

Aula 21

5

### → Cada vez mais informação!



5

## Enquadramento

Aula 21

6

### → Tempo para melhores decisões

Necessidade de tomar decisões em tempo adequado/oportuno.



6

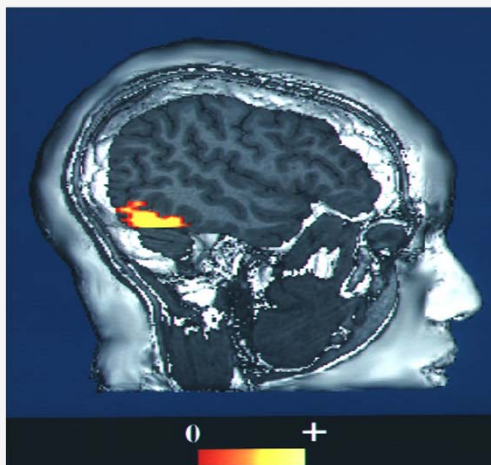
## → Decisores para tomar decisões

Um dos fatores mais importantes para um processo de tomada de decisão é o próprio **agente de decisão**.



7

## Áreas de reconhecimento facial



8

Procura por mais informação

Aula 21  
**9**

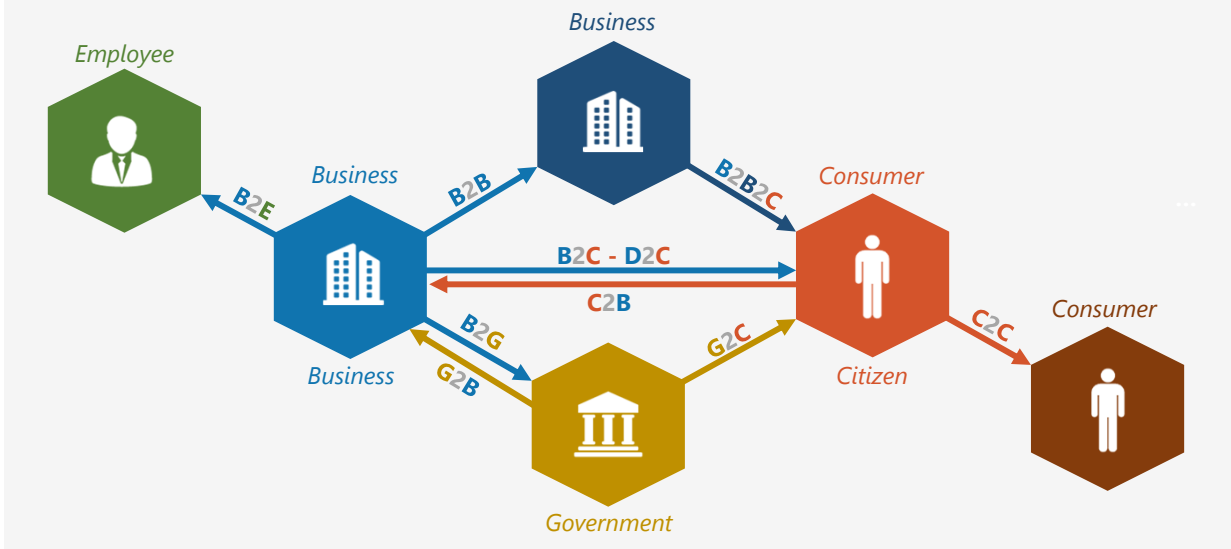
→ **Cada vez mais informação**



9

Modelos de negócio eletrónico

Aula 21  
**10**



10

Procura por mais informação

Aula 21  
**11**

É possível obter **mais informação** para tomada de decisão?



[https://www.nicepng.com/ourpic/u2t4y3y3t4t4r5t4\\_squeezed-orange/](https://www.nicepng.com/ourpic/u2t4y3y3t4t4r5t4_squeezed-orange/)

11

Procura por mais informação

Aula 21  
**12**

Maior capacidade para lidar com informação **não estruturada**.



<https://documentmedia.com/article-3204-Intelligent-Documents-Processing.html>

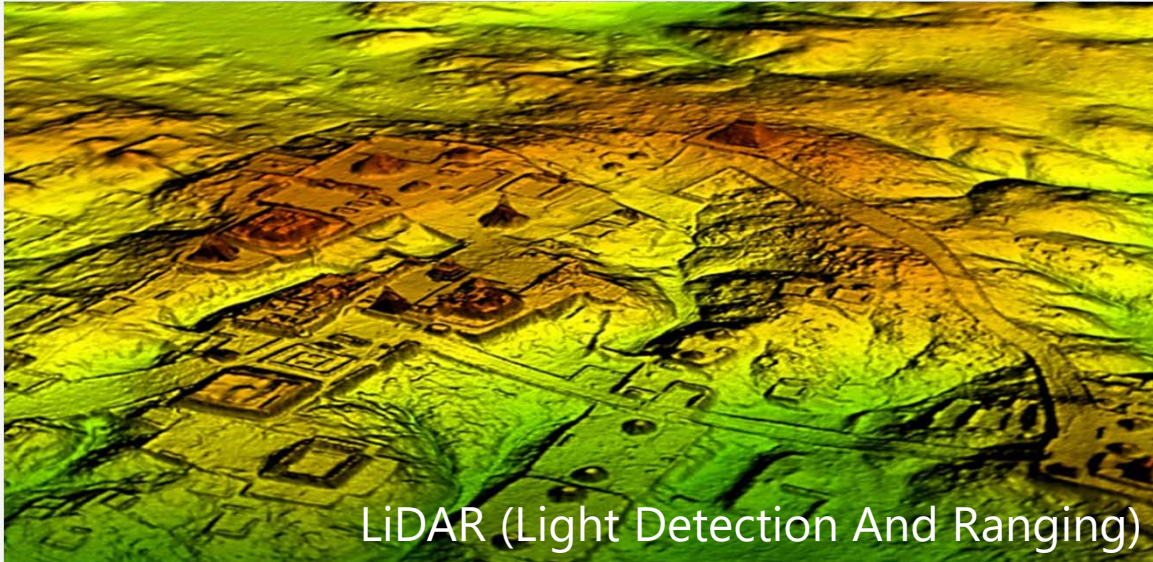
12



## Procura por mais informação

Aula 21

13



LiDAR (Light Detection And Ranging)

13

## Trabalhando com Informação não estruturada

Aula 21

14

### Pré-processamento do texto

Os dados obtidos a partir de *posts*, avaliações, comentários, observações e críticas em textos, nomeadamente RSO, são **altamente não estruturados**.

Assim o pré-processamento é uma etapa crítica, que impacta significativamente a qualidade/validade do trabalho seguinte.

14

### Extração de características (feature extraction)

Processo de **ML** no qual informações relevantes ou características são selecionadas e extraídas dos dados brutos.

Na análise de sentimentos/emoções, as características podem incluir frequências de palavras, sequências de palavras adjacentes ou outros atributos linguísticos.

15

### ➔ **Machine learning – Vector databases**

Uma base de dados vetorial é um tipo de base de dados que armazena e gere dados não estruturados, como texto, imagens ou áudio, em **embeddings** vetoriais (vetores de alta dimensionalidade) para facilitar a busca e recuperação rápida de objetos semelhantes.

16

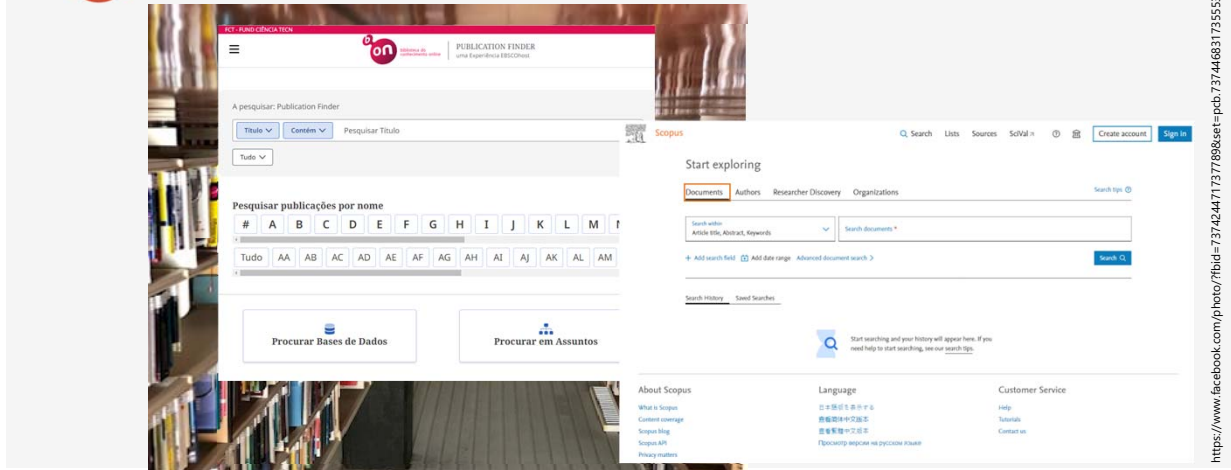


## Conceitos adicionais

Aula 21

17

### → *BD relacionais vs vetorais*



https://www.facebook.com/photo/?fbid=737424471737789&set=pcb-737446831735553

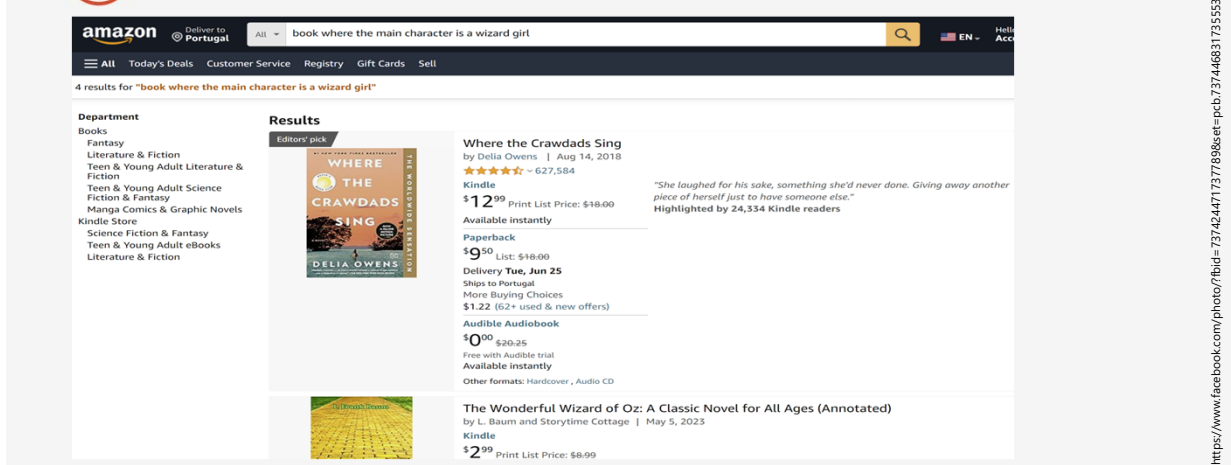
17

## Conceitos adicionais

Aula 21

18

### → *BD relacionais vs vetorais*

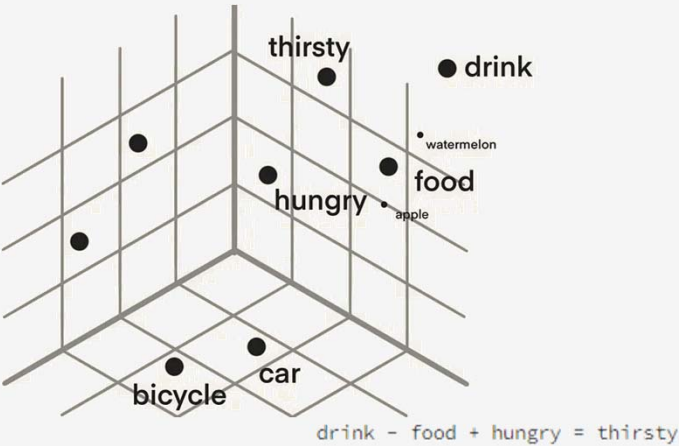


https://www.facebook.com/photo/?fbid=737424471737789&set=pcb-737446831735553

18

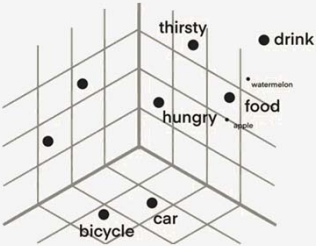
➔ **Machine learning - Vector embedding (VE)**

**Grape** ➡ [0.2, 1.1, 1.2, 3.1...]  
**Castle** ➡ [0.1, 2.8, 3.9, 4.2...]  
**Orange** ➡ [2.3, 0.9, 3.1, 5.3...]  
**Person** ➡ [5.0, 0.8, 0.1, 6.4...]



19

As pesquisas em VD pretendem encontrar documentos, imagens ou outros dados que sejam semelhantes ou “**aproximados**” de alguma maneira (por exemplo características ou contexto).

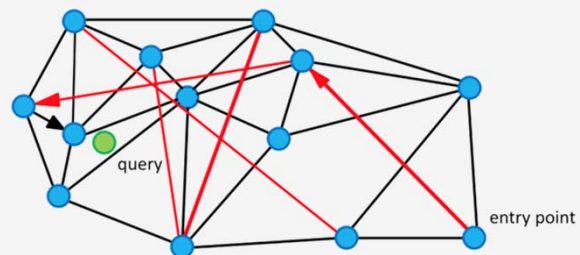


20

## Conceitos adicionais

**Exact Nearest Neighbor (kNN)**

O resultado de uma pesquisa é calculado comparando o **vetor da query** com **todos** os outros vetores da VD.

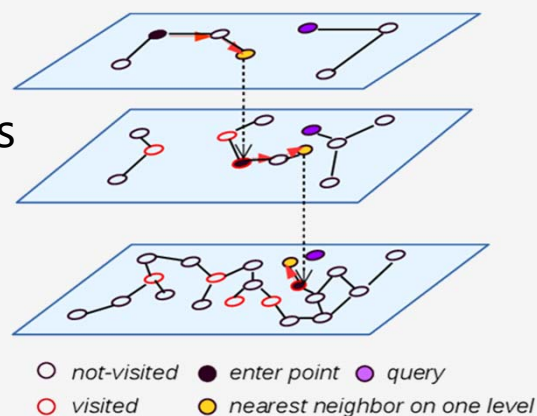


21

## Conceitos adicionais

**Approximate Nearest Neighbor (ANN)**

Com o **ANN**, troca-se alguma precisão por velocidade para recuperar os objetos aproximadamente mais semelhantes a uma consulta.



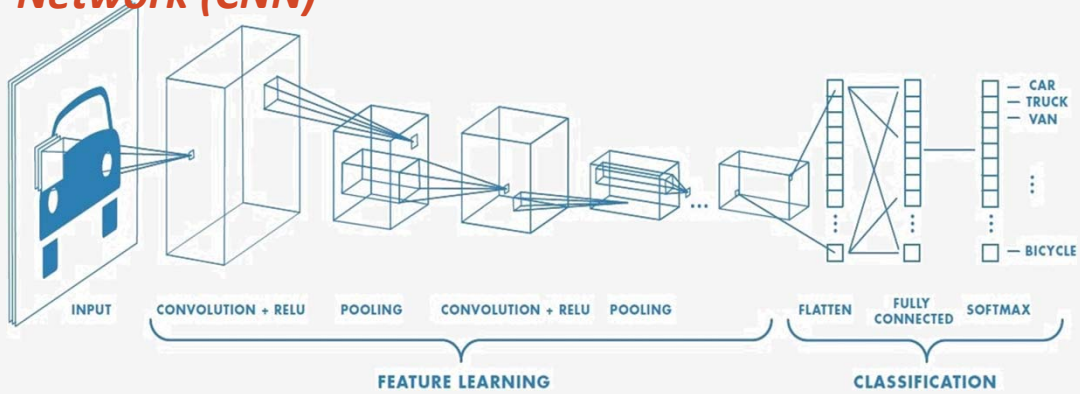
22

Conceitos adicionais

Aula 21

23

→ **Deep Learning - Convolution Neural Network (CNN)**



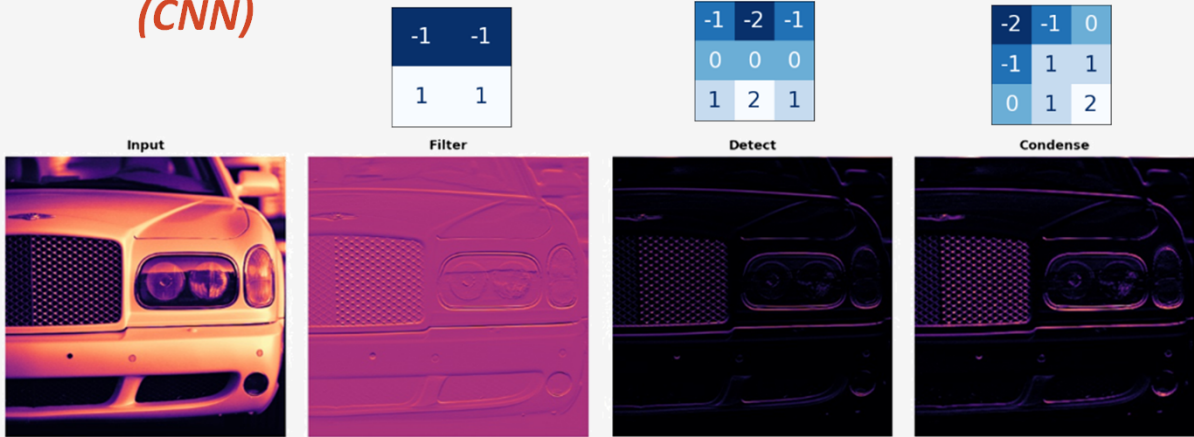
23

Conceitos adicionais

Aula 21

24

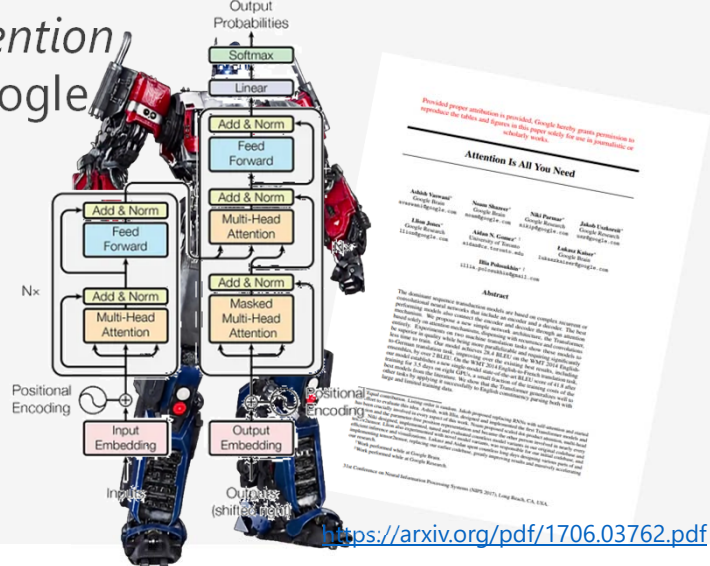
→ **Deep Learning - Convolution Neural Network (CNN)**



24

2017: Ano da mudança

O Research Paper "Attention Is All You Need" da Google apresentou uma **nova** arquitetura de **deep learning** conhecida como o: **Transformer**.



Arquitetura "Transformer"

## ➔ Deep Learning – Transfer Learning

Bidirectional Encoder Representations from Transformers (BERT)



→ **Deep Learning – Long-short term memory (LSTM), Bidirectional Long-short term memory (BI-LSTM)**

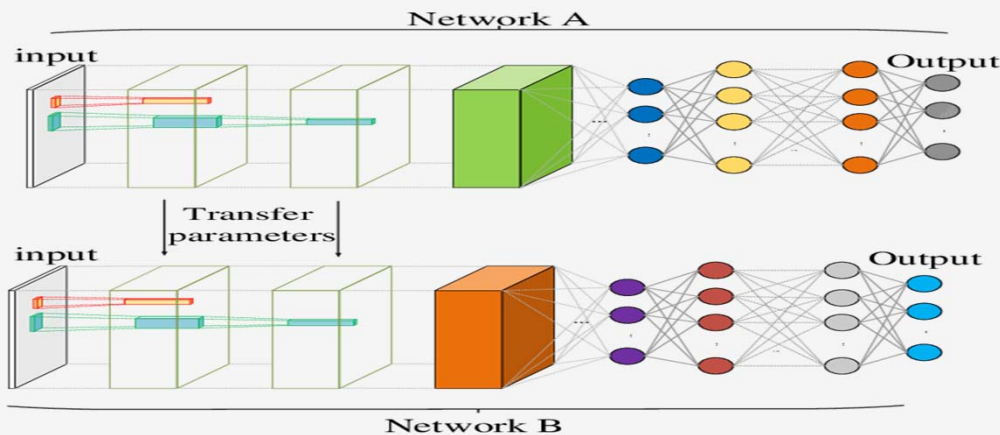
Apple is something that competitors simply cannot reproduce.

and

Apple is something that I like to eat.

27

→ **Deep Learning – Transfer Learning**



28



Desafios do processamento de linguagem natural

"Está um dia maravilhoso!"

"Adoro o meu iPhone, mas a bateria podia ser melhor."

"Tenho saudades de ti!"

"Este meu carro dá-me saudades do anterior e desejo pelo próximo."

29

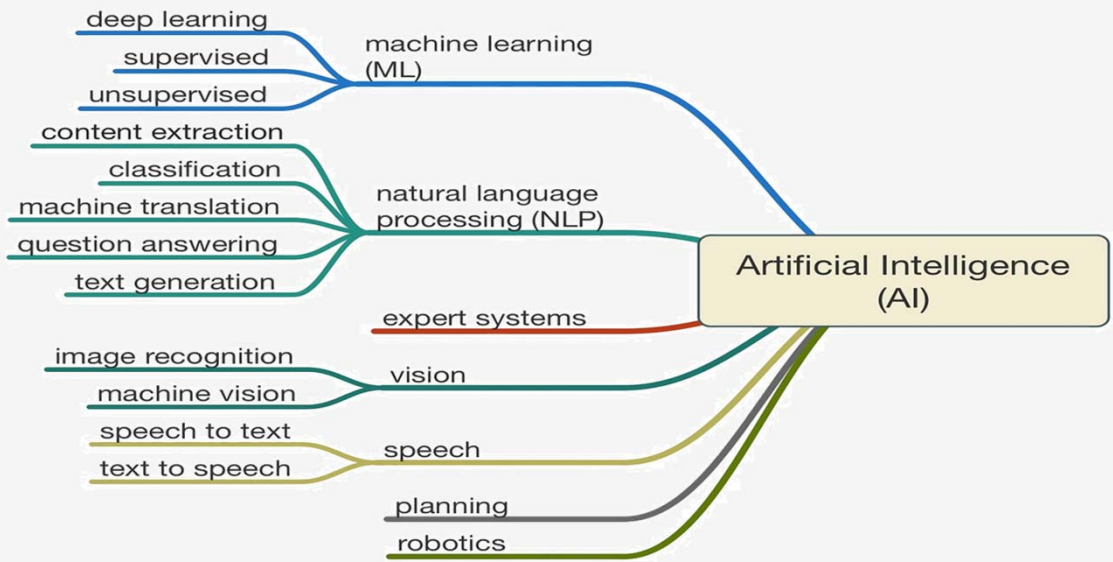
Desafios do processamento de linguagem natural



30

Tipos de inteligência artificial

Aula 21  
**31**

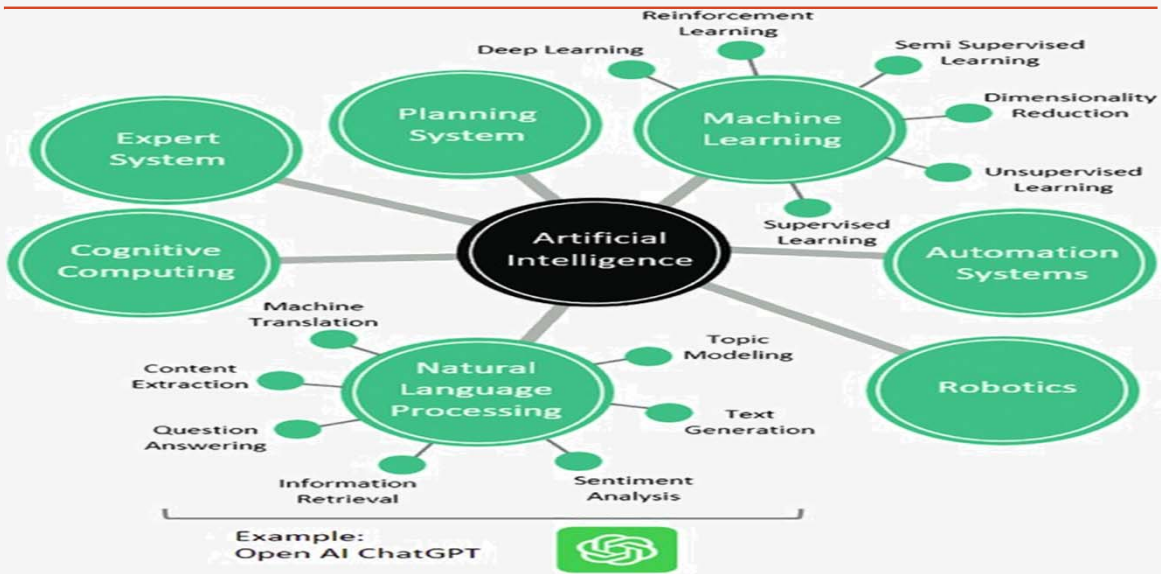


<https://chetan.kumar.medium.com/artificial-intelligence-definition-types-examples-technologies-96e875c7d9b6>

31

Tipos de inteligência artificial

Aula 21  
**32**



[https://www.researchgate.net/publication/373938258\\_To\\_what\\_extent\\_does\\_ChatGPT\\_understand\\_genet/figures/101-1](https://www.researchgate.net/publication/373938258_To_what_extent_does_ChatGPT_understand_genet/figures/101-1)

32

## Tipos de inteligência artificial

Aula 21

33

### TYPES OF AI



#### Reactive AI

Can only respond to the current state of the world. They do not have any memory of past events, and they cannot plan for the future.



#### Limited-memory

Can remember past events and use this information to make decisions. Cannot reason about the future or understand the intentions of other agents.



#### Theory of Mind

Can understand the thoughts and intentions of other agents. Cooperate with other agents and to achieve goals that would be impossible for a single agent.



#### Self Aware

Can understand the current state of the world and its own state. Can reason about the future and understand the intentions of other agents.

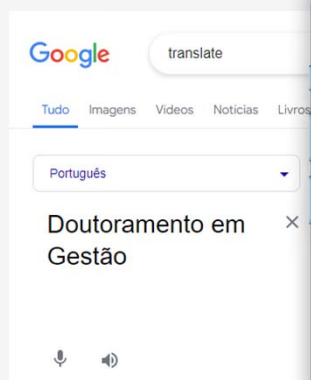
33

## Curiosidade #02

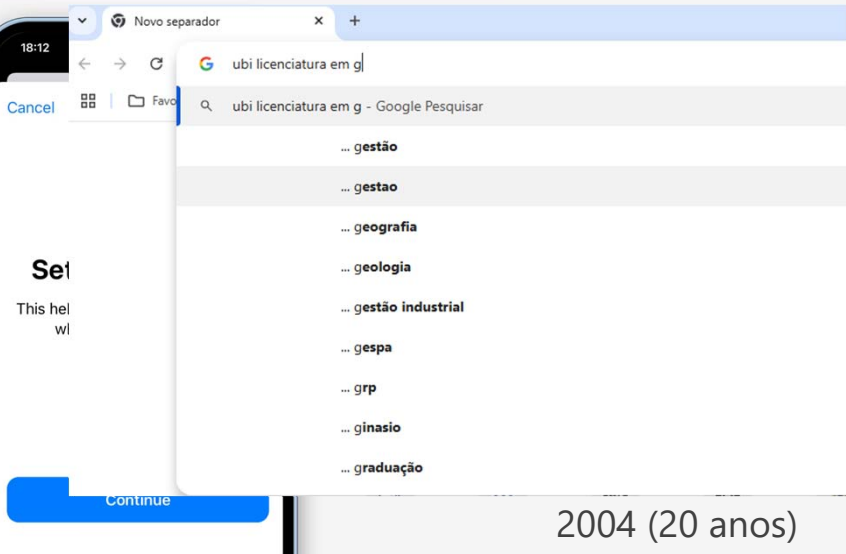
Aula 21

34

É anterior a 201



2006 (18 anos)



2004 (20 anos)

34

## IA Generativa

Aula 21

35

É um tipo de inteligência artificial (IA) que usa técnicas de aprendizagem profunda e conjuntos de dados massivamente grandes para entender, resumir, gerar e prever **novo conteúdo**.

### Artificial Intelligence

Is the field of study

### Machine Learning

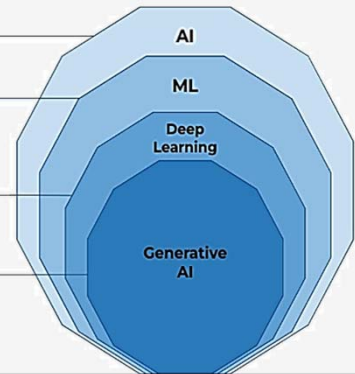
Is a branch of AI that focus on the creation of intelligent machines that learn from data. Another very well know branch inside AI is **Optimization**.

### Deep Learning

Is a subset of Machine Learning methods, based on **Artificial Neural Networks**. Examples: CNNs, RNNs

### Generative AI

A type of ANNs that generate data that is similar to the data it was trained on. Examples: GANs, LLMs



<https://gcloud.devoteam.com/pt-pt/blog/generative-ai/>

35

IA: Geração | **Previsão** | C

Aula 21

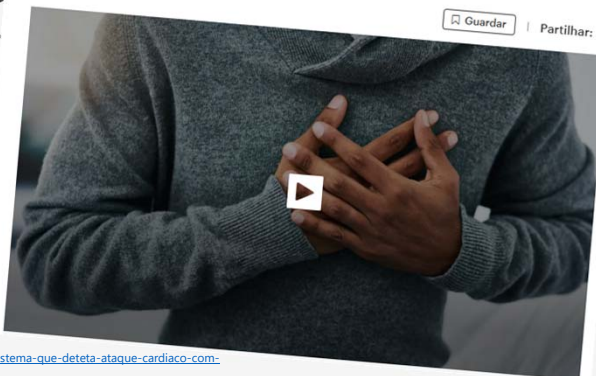
36

A IA **preditiva**, em inteligência artificial **generativa**, usa dados históricos para prever eventos, ou mesmo para tomar decisões. Organizações utilizam a IA para a tomada de decisões orientadas por dados.

### PAÍS Investigadora portuguesa cria sistema que deteta ataque cardíaco com recurso à Inteligência Artificial

A doença coronária é considerada a principal causa de morte no mundo, sendo a dor torácica um dos motivos mais comuns de admissão nos serviços de urgência.

Rui Carlos Teixeira, Ana Veloso e Miguel Castro  
17:25, 30 jan.2024



<https://sicnoticias.pt/pais/2024-01-30-Investigadora-portuguesa-cria-sistema-que-deteta-ataque-cardiaco-com-recurso-a-Inteligencia-Artificial-ac7cf8e8>

36

Através de técnicas de processamento de linguagem natural (NLP) e Machine Learning, a IA pode ser usada para conversação

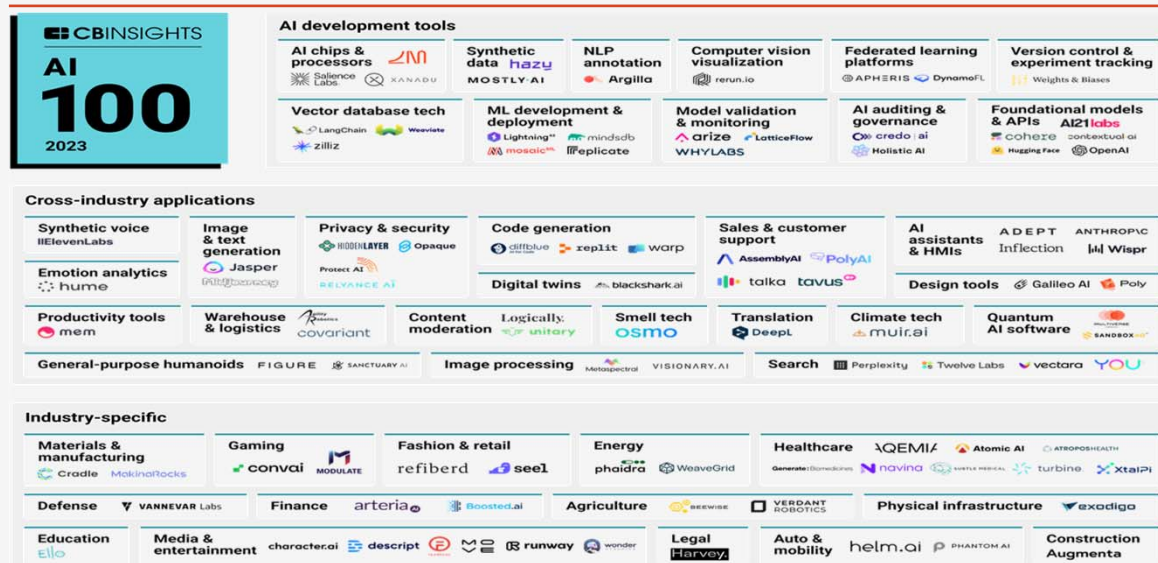
assistentes  
atendimento  
linguagem  
de forma



texto ou



37



<https://www.cbinsights.com/research/artificial-intelligence-top-startups-2023/>

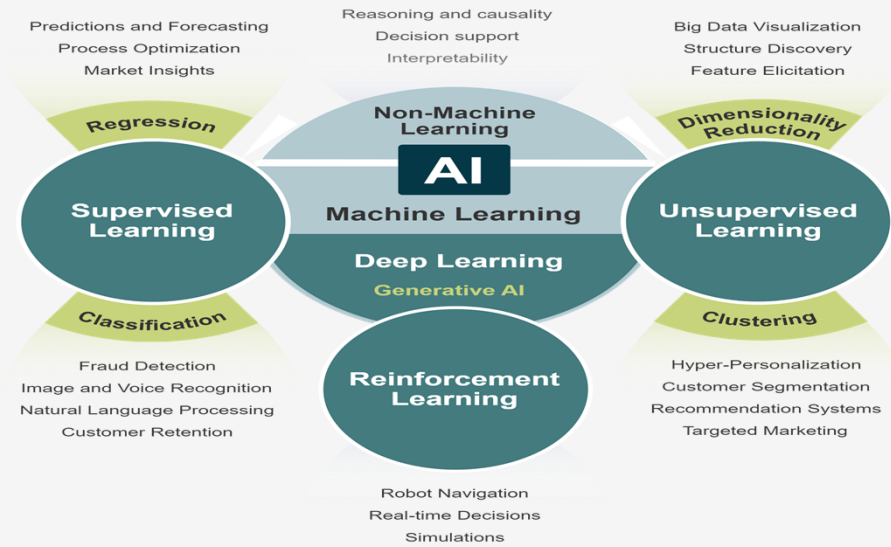
38



## Análise de sentimentos e detecção de emoções

Aula 21

39

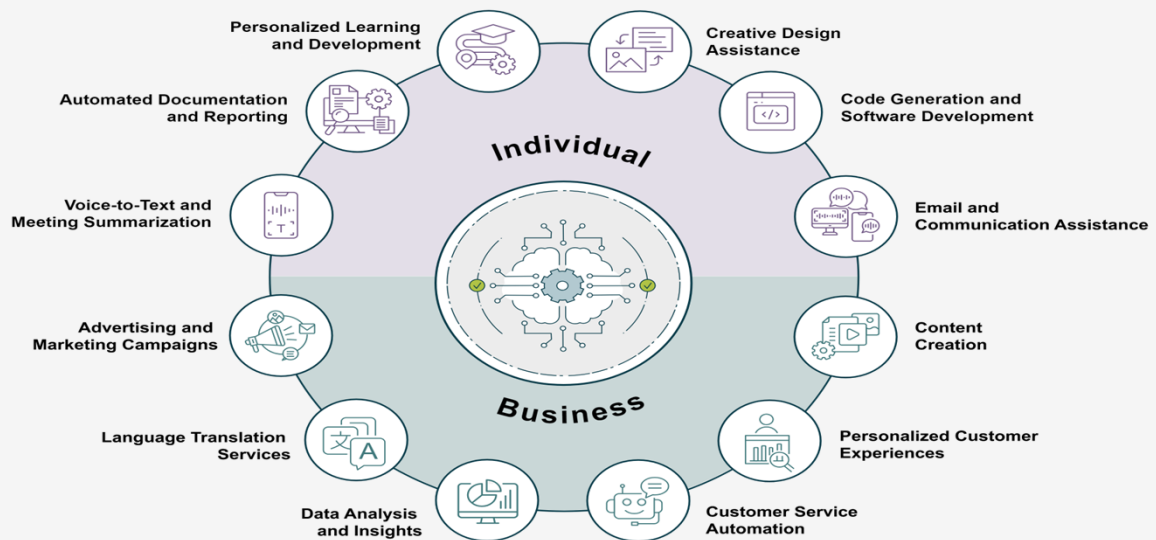


39

## Aplicações em ambien

Aula 21

40



40



## → **Feedback e avaliações de clientes**

Avaliação de produtos e serviços



41

## → **Interações nas redes sociais**

Monitorização de marcas

Desempenho de campanhas/iniciativas

Teste de Textos Publicitários

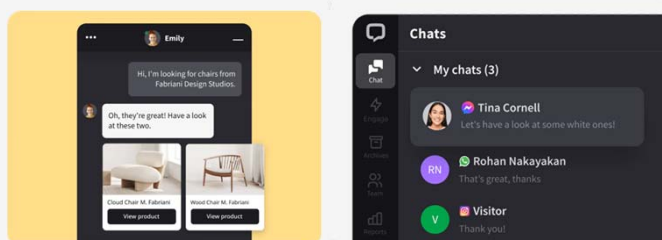


42

## → Atendimento ao cliente

*Chats ao vivo*

Centros de atendimento (*call centers*)



[https://www.livechat.com/?a=wbV/EeeyG8uim\\_campaign=pp\\_livechat-default&utm\\_source=PP](https://www.livechat.com/?a=wbV/EeeyG8uim_campaign=pp_livechat-default&utm_source=PP)

43

## → Estudos de mercado e percepção do consumidor

Transcrições de grupos de foco

Respostas a inquéritos

Detetar tendências



44

## → **Análise da concorrência**

Avaliar a concorrência

Detetar tendências



<https://www.knowledgeagent.de/en/products/competitor-analysis/pp>

45

## → **Feedback sobre produtos e experiência dos utilizadores**

Feedback de testes "beta"

Inquéritos de experiência do utilizador



46

## → Assistência à comunicação

Apoio à elaboração de *Emails*

Apoio ao contato com clientes



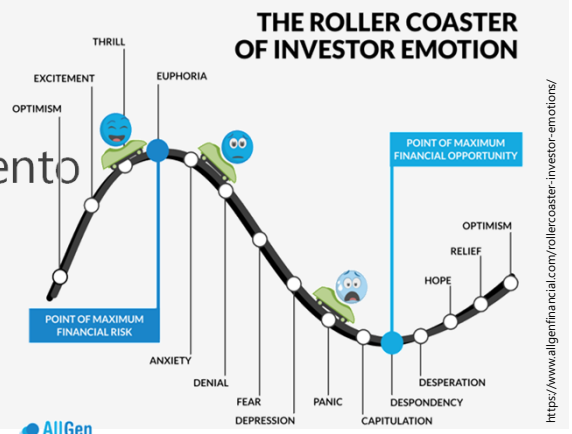
## → Relatórios financeiros

Relatórios de contas

Impacto em cotações

Sentimento de investimento

Deteção de fraudes



## → Automação de documentos

Criação de documentos automáticos

Geração de relatórios



[https://www.flaticon.com/free-icon/legal-document\\_5576437](https://www.flaticon.com/free-icon/legal-document_5576437)

49

Trata-se de um tipo de IA generativa especificamente arquitetada para **gerar conteúdo baseado em texto**.



50

show only:

100 MMLU

89.8 = human expert

70 = IDEAL

pre-2022 2022 2023 2024

Legend: anthropic (red), chinese (orange), google (yellow), meta (green), microsoft (teal), mistral (blue), openAI (purple), other (pink)

Key models labeled: GPT-3, GPT-4 Classic, Claude 3 Opus, Claude 3.5 Sonnet\*, GPT-4o, Gemini 1.5 Pro, Gemini 1.5 Flash, Gemini 1.5 Nano, Gemini 1.5 Ultra, Gemini 1.5 Pro-9B, Gemini 1.5 Pro-72B, Gemini 1.5 Pro-120B, Gemini 1.5 Pro-280B, Gemini 1.5 Pro-512B, Gemini 1.5 Pro-1024B, Gemini 1.5 Pro-2048B, Gemini 1.5 Pro-4096B, Gemini 1.5 Pro-8192B, Gemini 1.5 Pro-16384B, Gemini 1.5 Pro-32768B, Gemini 1.5 Pro-65536B, Gemini 1.5 Pro-131072B, Gemini 1.5 Pro-262144B, Gemini 1.5 Pro-524288B, Gemini 1.5 Pro-1048576B, Gemini 1.5 Pro-2097152B, Gemini 1.5 Pro-4194304B, Gemini 1.5 Pro-8388608B, Gemini 1.5 Pro-16777216B, Gemini 1.5 Pro-33554432B, Gemini 1.5 Pro-67108864B, Gemini 1.5 Pro-134217728B, Gemini 1.5 Pro-268435456B, Gemini 1.5 Pro-536870912B, Gemini 1.5 Pro-1073741824B, Gemini 1.5 Pro-2147483648B, Gemini 1.5 Pro-4294967296B, Gemini 1.5 Pro-8589934592B, Gemini 1.5 Pro-17179869184B, Gemini 1.5 Pro-34359738368B, Gemini 1.5 Pro-68719476736B, Gemini 1.5 Pro-137438953472B, Gemini 1.5 Pro-274877906944B, Gemini 1.5 Pro-549755813888B, Gemini 1.5 Pro-1099511627776B, Gemini 1.5 Pro-2199023255552B, Gemini 1.5 Pro-4398046511104B, Gemini 1.5 Pro-8796093022208B, Gemini 1.5 Pro-17592186044416B, Gemini 1.5 Pro-35184372088832B, Gemini 1.5 Pro-70368744177664B, Gemini 1.5 Pro-140737488355328B, Gemini 1.5 Pro-281474976710656B, Gemini 1.5 Pro-562949953421312B, Gemini 1.5 Pro-1125899906842624B, Gemini 1.5 Pro-2251799813685248B, Gemini 1.5 Pro-4503599627370496B, Gemini 1.5 Pro-9007199254740992B, Gemini 1.5 Pro-18014398509481984B, Gemini 1.5 Pro-36028797018963968B, Gemini 1.5 Pro-72057594037927936B, Gemini 1.5 Pro-144115188075855872B, Gemini 1.5 Pro-288230376151711744B, Gemini 1.5 Pro-576460752303423488B, Gemini 1.5 Pro-1152921504606846976B, Gemini 1.5 Pro-2305843009213693952B, Gemini 1.5 Pro-4611686018427387904B, Gemini 1.5 Pro-9223372036854775808B, Gemini 1.5 Pro-18446744073709551616B, Gemini 1.5 Pro-36893488147419103232B, Gemini 1.5 Pro-73786976294838206464B, Gemini 1.5 Pro-147573952589676412928B, Gemini 1.5 Pro-295147905179352825856B, Gemini 1.5 Pro-590295810358705651712B, Gemini 1.5 Pro-1180591620717411303424B, Gemini 1.5 Pro-2361183241434822606848B, Gemini 1.5 Pro-4722366482869645213696B, Gemini 1.5 Pro-9444732965739290427392B, Gemini 1.5 Pro-18889465931478580854784B, Gemini 1.5 Pro-37778931862957161709568B, Gemini 1.5 Pro-75557863725914323419136B, Gemini 1.5 Pro-151115727451828646838272B, Gemini 1.5 Pro-302231454903657293676544B, Gemini 1.5 Pro-604462909807314587353088B, Gemini 1.5 Pro-1208925819614629174706176B, Gemini 1.5 Pro-2417851639229258349412352B, Gemini 1.5 Pro-4835703278458516698824704B, Gemini 1.5 Pro-9671406556917033397649408B, Gemini 1.5 Pro-19342813113834066795298816B, Gemini 1.5 Pro-38685626227668133590597632B, Gemini 1.5 Pro-77371252455336267181195264B, Gemini 1.5 Pro-154742504910672534362390528B, Gemini 1.5 Pro-309485009821345068724781056B, Gemini 1.5 Pro-618970019642690137449562112B, Gemini 1.5 Pro-1237940039285380274899124224B, Gemini 1.5 Pro-2475880078570760549798248448B, Gemini 1.5 Pro-4951760157141521099596496896B, Gemini 1.5 Pro-9903520314283042199192993792B, Gemini 1.5 Pro-19807040628566084398385987584B, Gemini 1.5 Pro-39614081257132168796771975168B, Gemini 1.5 Pro-79228162514264337593543950336B, Gemini 1.5 Pro-158456325028528675187087900672B, Gemini 1.5 Pro-316912650057057350374175801344B, Gemini 1.5 Pro-633825300114114700748351602688B, Gemini 1.5 Pro-1267650600228229401496703205376B, Gemini 1.5 Pro-2535301200456458802993406410752B, Gemini 1.5 Pro-5070602400912917605986812821504B, Gemini 1.5 Pro-10141204801825835211973625643008B, Gemini 1.5 Pro-20282409603651670423947251286016B, Gemini 1.5 Pro-40564819207303340847894502572032B, Gemini 1.5 Pro-81129638414606681695789005144064B, Gemini 1.5 Pro-162259276829213363391578010288128B, Gemini 1.5 Pro-324518553658426726783156020576256B, Gemini 1.5 Pro-649037107316853453566312041152512B, Gemini 1.5 Pro-1298074214633706907132624082305024B, Gemini 1.5 Pro-2596148429267413814265248164610048B, Gemini 1.5 Pro-5192296858534827628530496329220096B, Gemini 1.5 Pro-10384593717069655257060992658440192B, Gemini 1.5 Pro-20769187434139310514121985316880384B, Gemini 1.5 Pro-41538374868278621028243970633760768B, Gemini 1.5 Pro-83076749736557242056487941267521536B, Gemini 1.5 Pro-166153499473114484112975882535043072B, Gemini 1.5 Pro-332306998946228968225951765070086144B, Gemini 1.5 Pro-664613997892457936451903530140172288B, Gemini 1.5 Pro-1329227995784915872903807060280344576B, Gemini 1.5 Pro-2658455991569831745807614120560689152B, Gemini 1.5 Pro-5316911983139663491615228241121378304B, Gemini 1.5 Pro-10633823966279326983230456482242756608B, Gemini 1.5 Pro-21267647932558653966460912964485513216B, Gemini 1.5 Pro-42535295865117307932921825928971026432B, Gemini 1.5 Pro-85070591730234615865843651857942052864B, Gemini 1.5 Pro-170141183460469231731687303715884105728B, Gemini 1.5 Pro-340282366920938463463374607431768211456B, Gemini 1.5 Pro-680564733841



26

Diagrama de uma rede neural profunda com 5 camadas: 5 nós de entrada (azuis), 5 nós ocultos (vermelhos), 5 nós ocultos (amarelos) e 3 nós de saída (verdes). As conexões são representadas por linhas coloridas correspondentes aos nós.

$$5 \times 8 + 8 + 8 \times 4 + 4 + 4 \times 3 + 3 = 99 \text{ parâmetros}$$

52



A importância da *Prompt*

53

Uma **Prompt** é um modo de interação entre um humano e um *LLM* que permite ao modelo gerar a saída pretendida.



53

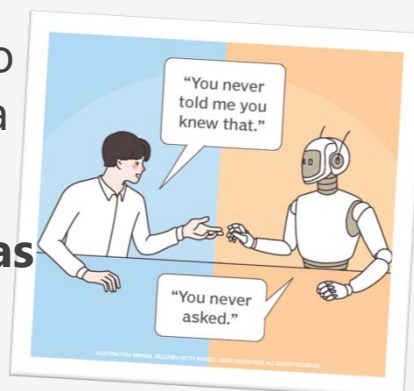
A importância da *Prompt*

54

### → Criação de *Prompts*

Requer uma consideração cuidadosa para garantir que o resultado obtido corresponda ao desejado.

**Formular as perguntas certas** é essencial para alcançar os resultados pretendidos.



54

## A importância da *Prompt*

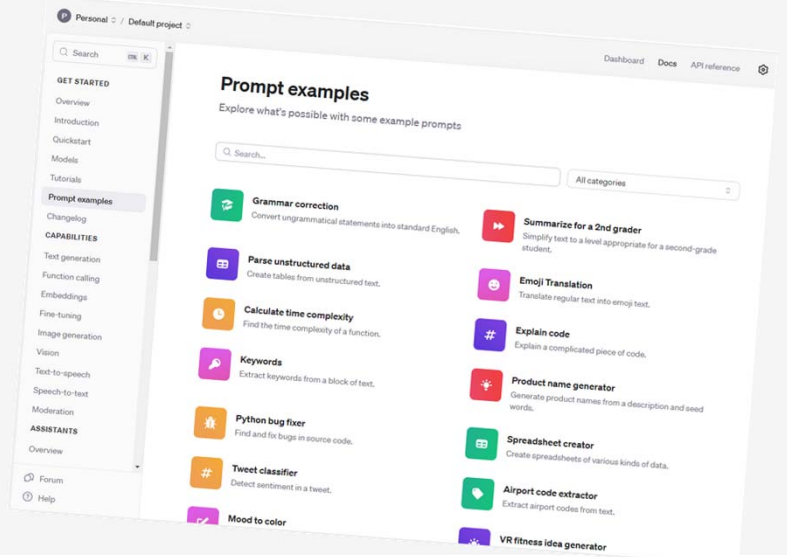
Aula 21

55



**Frameworks**

Exemplos na  
API OpenAI

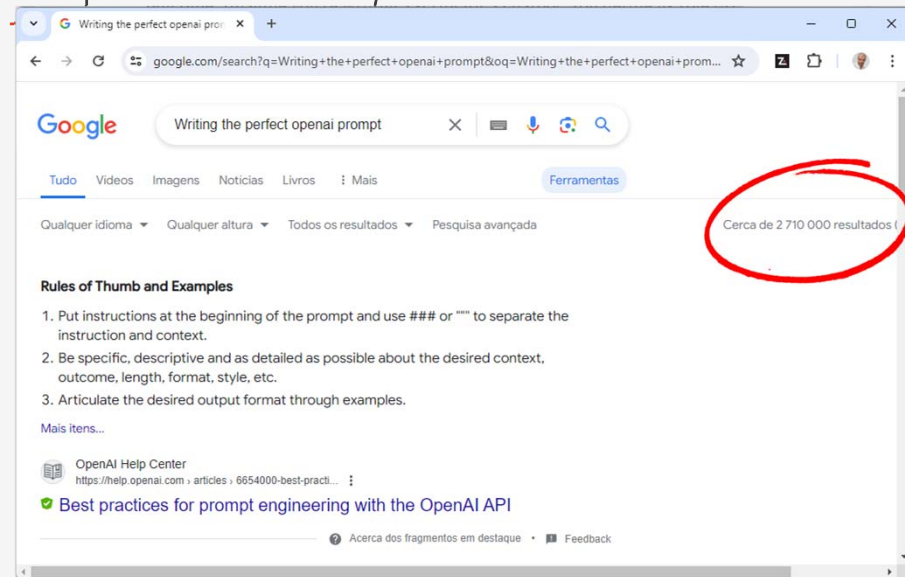


55

## A importância da *Prompt*

Aula 21

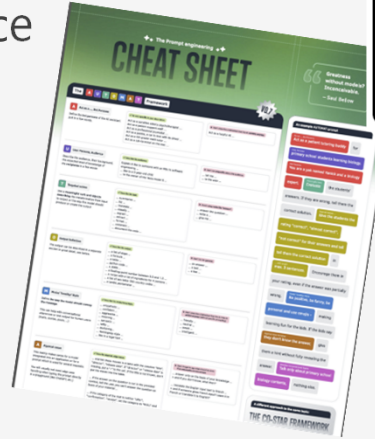
56



56

## A importância da *Prompt*

- A** Act as a ...
- U** User Persona & Audience
- T** Targeted Action
- O** Output Definition
- M** Mode / Tonality / Style
- A** Atypical Cases
- T** Topic Whitelisting



Maximilian Vogel

57

## A importância da *Prompt*

### THE CO-STAR FRAMEWORK

- C** Context
- O** Objective
- S** Style
- T** Tone
- A** Audience
- R** Response



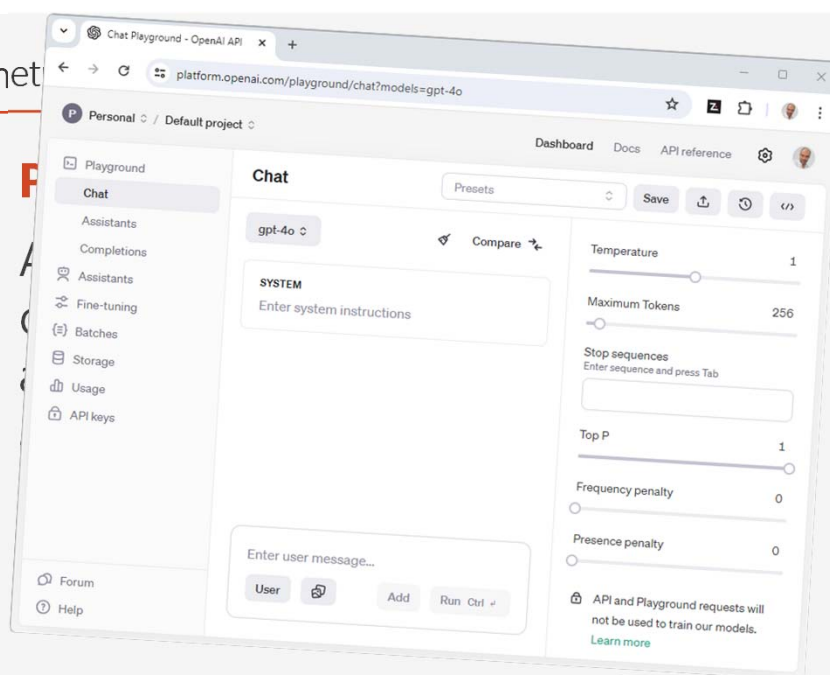
Maximilian Vogel

58

## Parametrização

Aula 21

59



59

## Parametrização avançada

Aula 21

60

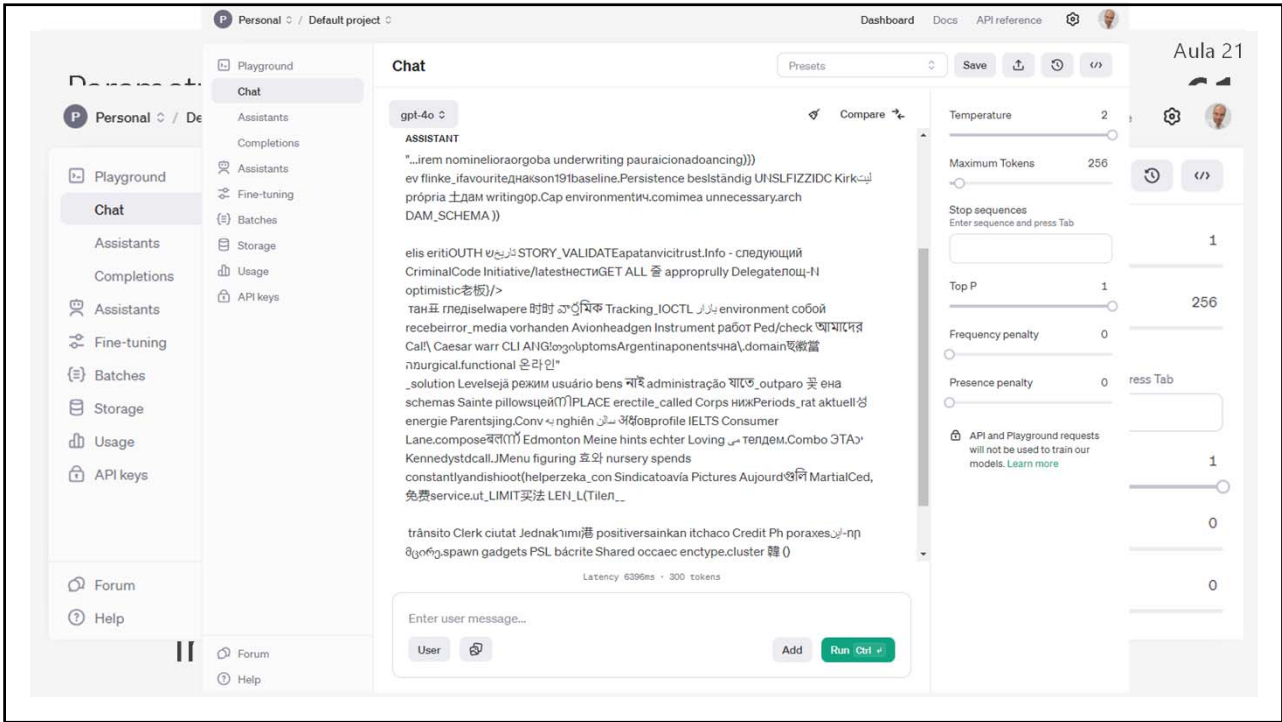
### → Temperatura

Controla o grau de aleatoriedade das respostas e o seu valor varia entre 0 e 2.



James Yang: <https://www.wsj.com/articles/hallucination-when-chatbots-and-people-see-what-isnt-there-91c6c885>

60



61

Parametrização avançada

Aula 21

62

➔ **Maximum length (em tokens)**

relacionada com o número máximo de tokens que serão gerados.

Para texto em **inglês**, 1 **token** equivale aproximadamente a **0,75 palavras** (ou 4 caracteres). Em **português**, 1 **token** equivale aproximadamente a **1 palavra** (ou 5 caracteres).

62

➔ **Maximum length (em tokens)**

Este número **inclui** os *tokens* da **prompt** de entrada.

**Output** = **MaxL** - **Input**

GPT-4o

GPT-4o ("o" for "omni") is our most advanced model. It is multimodal (accepting text or image inputs and outputting text), and it has the same high intelligence as GPT-4 Turbo but is much more efficient—it generates text 2x faster and is 50% cheaper. Additionally, GPT-4o has the best vision and performance across non-English languages of any of our models. GPT-4o is available in the OpenAI API to paying customers. Learn how to use GPT-4o in our [text generation guide](#).

| MODEL             | DESCRIPTION                                                                                                                                          | CONTEXT WINDOW | TRAINING DATA  |
|-------------------|------------------------------------------------------------------------------------------------------------------------------------------------------|----------------|----------------|
| gpt-4o            | <b>New GPT-4o</b><br>Our most advanced, multimodal flagship model that's cheaper and faster than GPT-4 Turbo. Currently points to gpt-4o-2024-05-13. | 128,000 tokens | Up to Oct 2023 |
| gpt-4o-2024-05-13 | gpt-4o currently points to this version.                                                                                                             | 128,000 tokens | Up to Oct 2023 |

GPT-4 Turbo and GPT-4

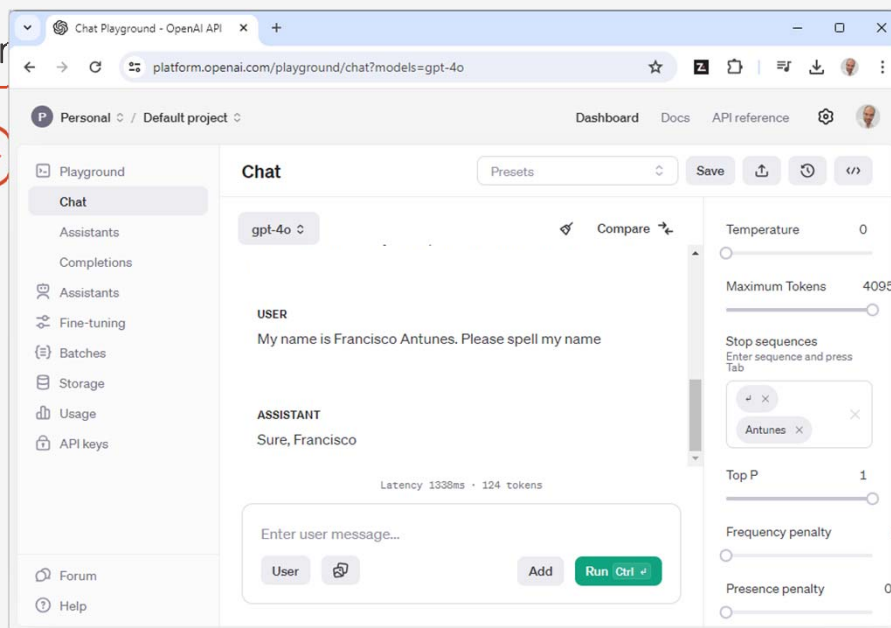
GPT-4 is a large multimodal model (accepting text or image inputs and outputting text) that can solve difficult problems with greater accuracy than any of our previous models, thanks to its broader general knowledge and advanced reasoning capabilities. GPT-4 is available in the OpenAI API to [paying customers](#). Like gpt-3.5-turbo, GPT-4 is optimized for chat but works well for traditional completions tasks using the [Chat Completions API](#). Learn how to use GPT-4 in our [text generation guide](#).

| MODEL                  | DESCRIPTION                                                                                                                                                    | CONTEXT WINDOW | TRAINING DATA  |
|------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------|----------------|
| gpt-4-turbo            | The latest GPT-4 Turbo model with vision capabilities. Vision requests can now use JSON mode and function calling. Currently points to gpt-4-turbo-2024-04-09. | 128,000 tokens | Up to Dec 2023 |
| gpt-4-turbo-2024-04-09 | GPT-4 Turbo with Vision model. Vision requests can now use JSON mode and function calling. gpt-4-turbo currently points to this version.                       | 128,000 tokens | Up to Dec 2023 |
| gpt-4-turbo-preview    | GPT-4 Turbo preview model. Currently points to gpt-4-                                                                                                          | 128,000 tokens | Up to Dec 2023 |

Parar

Aula 21

65


<https://blog.clutch.ca/posts/how-long-do-car-brakes-last>

65

Parametrização avançada

Aula 21

66

## → Top P

Com a amostragem **Top p**, em vez de considerar todo o vocabulário, a próxima palavra será escolhida a partir de um conjunto menor de palavras que, coletivamente, têm uma **probabilidade cumulativa** acima do valor **Top p**.

66



ASSISTANT

Em "Ecos do Futuro", um grupo de intrépidos alunos de doutoramento em gestão descobre a poderosa ferramenta da análise de sentimentos e vislumbra um mundo de possibilidades para a gestão empresarial. Sob a orientação do enigmático professor Munhoz, eles embarcam numa jornada que revela não apenas os padrões emocionais por trás de decisões de negócios, mas também os profundos impactos que palavras e sentimentos podem ter nas estratégias corporativas.

Unidos por um desejo fervoroso de inovar, esses jovens acadêmicos mergulham em dados que vão desde avaliações de clientes até interações nas redes sociais, desvendando insights que prometem revolucionar a relação entre empresas e consumidores. Contudo, seu caminho é repleto de desafios. Entre dilemas éticos sobre privacidade, a pressão das grandes corporações e o próprio ceticismo acadêmico, o grupo enfrenta uma montanha-russa emocional enquanto luta para provar que as emoções podem, de fato, governar o mundo dos negócios.

"\*\*Ecos do Futuro\*\*" é uma narrativa vibrante que entrelaça ciência, ética e humanidade, questionando até onde a análise de sentimentos pode ir e o que acontece quando deciframos o código das emoções que movem o mercado global

Aula 21

67

io

67

Parametrização avançada

ASSISTANT

Em "Sentimentos em Gestão", um grupo de alunos de doutoramento em gestão de uma prestigiada universidade decide explorar um campo inovador: a análise de sentimentos aplicada à gestão empresarial. Liderados pela carismática e visionária professora Dr.ª Clara Mendes, os estudantes mergulham em um projeto ambicioso que visa transformar a maneira como as empresas entendem e respondem às emoções de seus clientes e funcionários.

Cada membro do grupo traz uma habilidade única para a mesa: João, o gênio da programação, desenvolve algoritmos sofisticados; Ana, a especialista em psicologia organizacional, interpreta os dados com uma sensibilidade aguçada; Pedro, o comunicador nato, traduz as descobertas em estratégias práticas; e Luísa, a analista de mercado, identifica oportunidades de aplicação no mundo real.

À medida que avançam, eles enfrentam desafios técnicos, dilemas éticos e a resistência de executivos tradicionais. No entanto, a paixão pelo potencial transformador da análise de sentimentos os mantém unidos. Em uma reviravolta surpreendente, uma grande corporação se interessa pelo projeto, oferecendo uma oportunidade que pode catapultar suas carreiras e revolucionar a gestão empresarial.

"Sentimentos em Gestão" é uma história de inovação,

Aula 21

68

Maximum Tokens

256

Stop sequences

Enter sequence and press Tab

Top P

0.01

Frequency penalty

0

Presence penalty

0

API and Playground requests will not be used to train our models.

[Learn more](#)

68

34

## → *Frequency penalty*

Um dos problemas na geração de textos é a **repetição**.

Ao aplicar penalizações a palavras frequentes, este parâmetro "incentiva" o modelo a diversificar o uso da linguagem.



69

### ASSISTANT

Zebra ziguezagueia,  
Zefiro zela,  
Zona zero,  
Zumbido zela.

Zebra zoraide,  
Zela zigoto,  
Zumbindo, zige-zagueando,  
Zoo zensacional!

Ziguezague zênite,  
Zonza, zigente,  
Zelosa, zumba,  
Zona zêbrula.

Zebra zangada,  
Zé zebra ziguezagueia,  
Zona zebra! zigzagueando  
Zoando, zumbizando.

Zero zanaga zebreira,  
Zeros zebritos zanzam.  
Zum! Zim! Zum!  
Zelosos zemrosos zaróicos!

Zenit zalenteio,  
Za-pimba! Za-zoura!  
Zeferino zoa.

Zirigaita zapumbaqueira,  
Zu-zu Zu...  
Zzz... Zebra zeja zen.

Maximum Tokens 256

Stop sequences

Enter sequence and press Tab

Top P 1

Frequency penalty 2

Presence penalty 0

API and Playground requests will not be used to train our models.



70

→ **Presence penalty**

Tem o mesmo objetivo do parâmetro anterior.

Esta penalidade é única (aditiva) e aplicada a um *token* que apareceu pelo menos uma vez, como um marcador Booleano (1 ou 0).

Média de 98 tokens de input (1 token até 1923 tokens) + Prompt com 26 tokens  
Máximo de 10 tokens de output  
833 execuções (arredondar para 900)

**OpenAI GPT API Pricing Calculator**

Wondering about how the OpenAI GPT and other AI model pricing works? Here's a pricing calculator for OpenAI GPT API, Anthropic Claude API, Mistral API and Google Gemini API.

OpenAI gpt-4o OpenAI gpt-3.5-turbo OpenAI gpt-4o OpenAI gpt-4 Anthropic claude-3-haiku Anthropic claude-3-sonnet Anthropic claude-3-opus Mistral small Mistral large Google gemini-1.5-flash Google gemini-1.5-pro

An execution includes both the prompt sent and the response. Language: English

| Tokens per execution | Words per execution | Price for 1 execution | Price for 900 executions |
|----------------------|---------------------|-----------------------|--------------------------|
| 100                  | 75                  | ~\$0.00075            | ~\$0.68                  |
| 1000                 | 750                 | ~\$0.00750            | ~\$6.75                  |
| 10000                | 7500                | ~\$0.07500            | ~\$67.50                 |
| 100000               | 75000               | ~\$0.75000            | ~\$675.00                |
| 1000000              | 750000              | ~\$7.50000            | ~\$6750.00               |

<https://gptforwork.com/tools/openai-chatgpt-api-pricing-calculator>

## LLM – Python - LangChain

```

1  # Import the library
2  import openai
3  import tiktoken
4
5  # assign the API key to the environment by replacing
6  # API_KEY with your generated key
7  # Importa a chave secreta do projeto da API da OpenAI
8  from SecretAPIKeys import OpenAI_LangChain_KPSI_key
9  openai.api_key = OpenAI_LangChain_KPSI_key
10
11
12
13  from pathlib import Path
14  from ConnectToDatabase import *
15
16  # Função para contar tokens com o engine="text-davinci-003"
17  def count_tokens(text):
18      encoding = tiktoken.get_encoding("p50k_base")
19      num_tokens = len(encoding.encode(text))
20      return num_tokens
21
22
23
24  # Função para obter o sentimento de um PROMPT definido
25  def GetSentiment(PROMPT, PromptTokenSize, outputs):
26      # using OpenAI's Completion module that helps perform
27      # text manipulations
28      response = openai.Completion.create(
29          # model name used here is text-davinci-003
30          # there are many other models available under the
31          # umbrella of GPT-3
32          model="text-davinci-003",
33          # passing the user input
34          prompt=PROMPT,

```

73

## LLM – Python - LangChain

Para poder automatizar pedidos de classificação de *Posts* em RSO de forma **automática**, é necessário prever uma infraestrutura técnica que permita não só enviar e receber a informação a tratar por LLM, mas também que permita usar **diferentes** LLM.



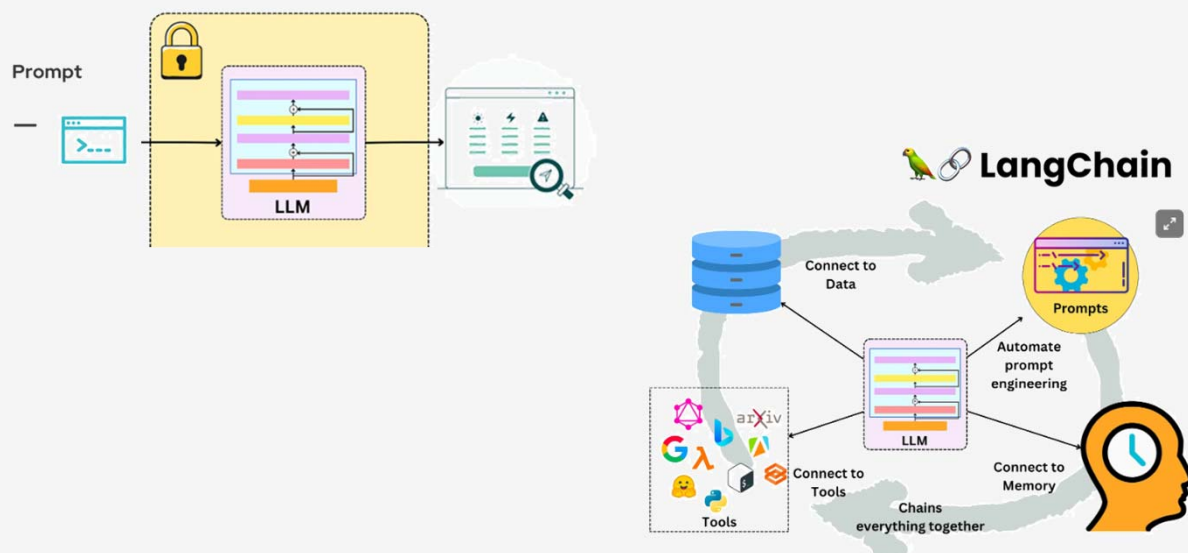
# LangChain

74

## LLM – Python - LangChain

Aula 21

75



75

## Recursos de programação

Aula 21

76

O **GitHub** é uma interface web que utiliza o *Git*, que é um software de controlo de versão de código aberto que permite que várias pessoas façam alterações separadas em páginas web ao mesmo tempo.



76

O **Hugging Face** é uma plataforma e comunidade que ajuda os utilizadores a criar, implementar e treinar modelos de *ML*, *publicando código aberto* (contrariamente ao efetuado pela *Google* ou pela *OpenAI*).



## Hugging Face

77

Além das abordagens mais tradicionais ou formais à aprendizagem de ferramentas computacionais, os recursos em vídeo podem diminuir o tempo de aprendizagem.



## YouTube

78

## Ficheiro JSON (JavaScript Object Notation)

**79**

Um ficheiro **JSON** é um formato ("leve") de intercâmbio de dados que é fácil de ler e escrever por seres humanos, mas que também é fácil de analisar e gerar por computadores.

É um formato de texto que é **completamente independente** da linguagem de programação.

{ j s o n }

79

## Ficheiro JSON (JavaScript Object Notation)

**80**

Um só ficheiro pode armazenar diversos de dados e objetos, **não requerendo uma estrutura fixa de dados** como, por exemplo, uma base de dados relacional.



80



## Resumo



A IA tornou-se cada vez mais presente na gestão, particularmente naquelas em que o excesso de informação e a necessidade de automatizar operações pode impossibilitar outra abordagem.



A IA tem várias modalidades, tendo respostas diversificadas consoante o problema a resolver, pelo que a IA generativa é apenas uma das suas facetas.



O potencial da utilização das ferramentas de IA não se encontra em ferramentas como o ChatGPT, ou outras de utilização livre de custos, mas sim através de soluções programáticas que permitem parametrizar soluções face às necessidades das empresas.