

...

Desvendando Empresas:

Análise de Dados Textuais com NLP

...

Por: Andrey de Oliveira

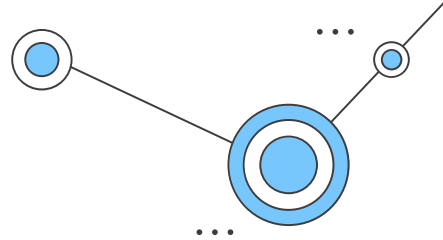
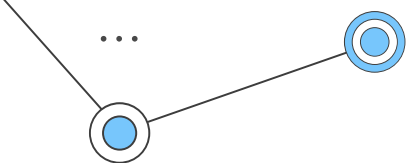


Análise Exploratória

	id	desc
911	1242114d-5fdd-4a83-8219-0af3ed39addf	Interhouse is an e-commerce platform that sells additives for fuel and lubricants.
1193	2a98f36a-db52-498d-8792-319b5e89e82d	Skai is an AI-powered platform that offers consumer and market insights, strategic planning, omnichannel media activation, testing, and measurement services to brands and retailers. It is the go-to-market engine that enables smarter decisions and better outcomes with actionable intelligence in making customer connections. The platform includes a suite of data-driven products for market intelligence, omnichannel media activation, testing, and measurement, enabling brands and agencies to make predictions, improve strategic planning and drive growth on the most important customer channels. Skai's machine-learning algorithms and proprietary NLP enable companies to listen, predict and keep pace with the consumer journey. The company is on a mission to empower businesses with data they can trust, insights they can use, and the impact they can measure. It was founded in Tel Aviv, Tel Aviv, Israel in 2006 by Yoav Izhar-Prato, Nir Cohen, Amos Talmor, and Alon Sheaffer. Operating in 15 international locations and backed by Sequoia Capital, Arts Alliance, Tenaya Capital, and Bain Capital Ventures, Skai™ generates over \$350 billion in annualized revenue for the world's top brands. Read More
1530	4721021d-4efb-4d77-aacd-d30eab334a5f	ThreatQuotient is a threat intelligence platform (TIP) that centrally manages and correlates unlimited external sources with all internal security and analytics solutions for contextual, operationalized intelligence in a single pane of glass. It provides Indicator Nurturing, which goes beyond enrichment to help customers tailor indicators of compromise (IOCs) more specifically to their infrastructure. It improves the efficiency and effectiveness of security operations with a platform that accelerates and simplifies investigations and collaboration within and across teams and tools. Integrating an organization's existing processes and technologies into a unified workspace, ThreatQuotient's solutions reduce noise, highlight top priority threats, and automate processes to provide greater focus and decision support while maximizing limited resources. Its threat-centric approach supports multiple use cases including incident response, threat hunting, spear phishing, alert triage, and vulnerability management, and also serves as a threat intelligence platform. ThreatQuotient has been recognized for product excellence and industry leadership by several award programs in 2021. It has won four 2021 Globec Cybersecurity Excellence Awards, including Gold for Security Investigation, Silver for both Threat Intelligence Technology and Hot Security Company of the Year, and Bronze for Threat Hunting, Detection, Intelligence, and Response. ThreatQuotient also won three 2021 Cybersecurity Excellence Awards: Gold for Security Investigation, and Silver for both Threat Detection, Intelligence and Response, and Best Cybersecurity Company. Additionally, ThreatQuotient was recognized as McAfee Partner of the Year for the second consecutive year and the company was listed in the 2021 CRN Partner Program Guide. Headquartered in Northern Virginia, ThreatQuotient was founded in 2013 by Ryan Trost and Wayne Chiang. Read More

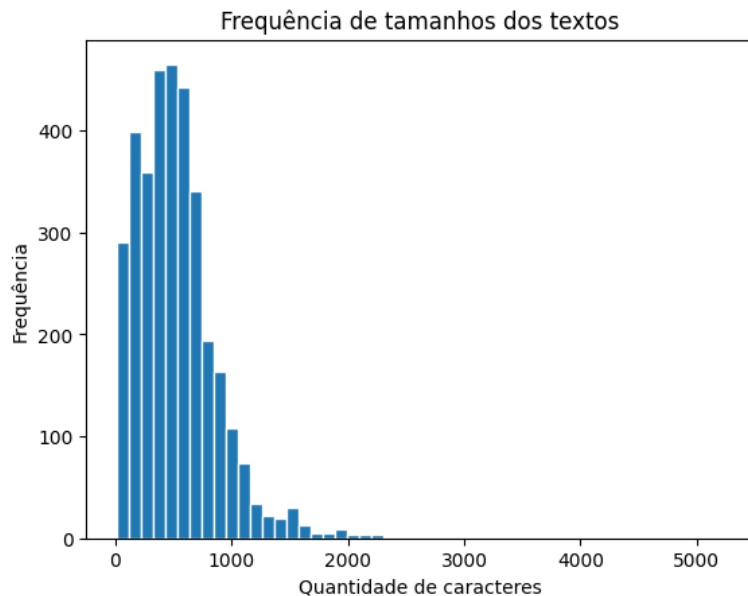
- Dimensão: 3617 linhas , 2 colunas
- Removidos:
 - 64 amostras sem descrição
 - 101 amostras com descrição duplicada
- Nova dimensão: 3452 amostras

...

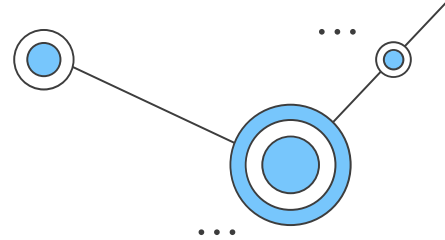
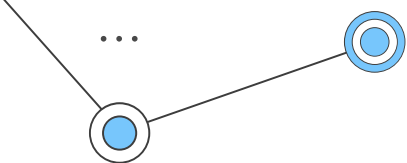


Análise Exploratória

Tamanho das Frases

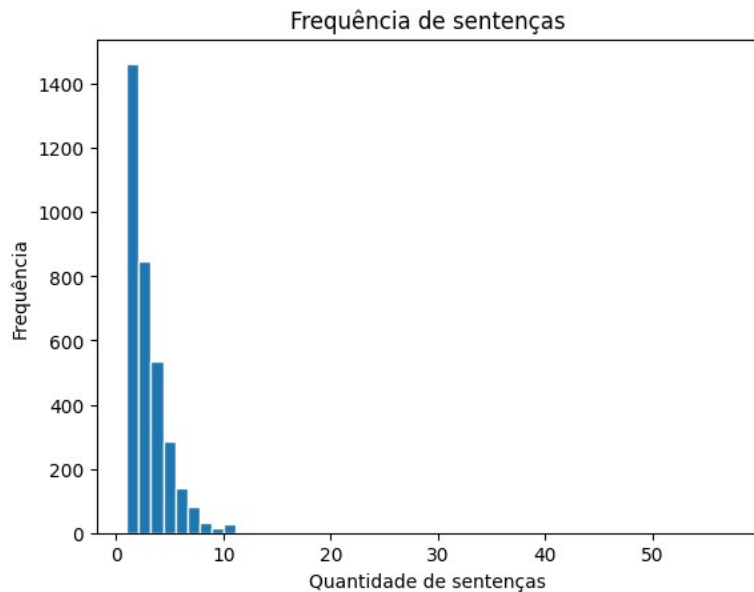


count	3452.000000
mean	527.343859
std	379.409009
min	11.000000
25%	276.750000
50%	475.000000
75%	688.000000
max	5228.000000

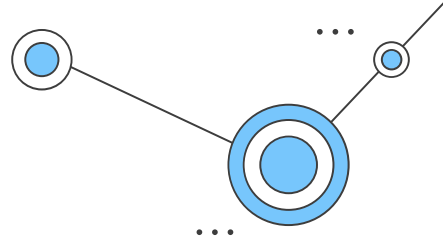
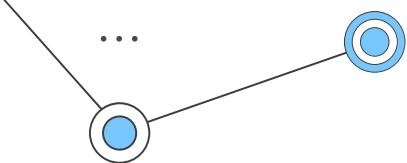


Análise Exploratória

Quantidade de frases

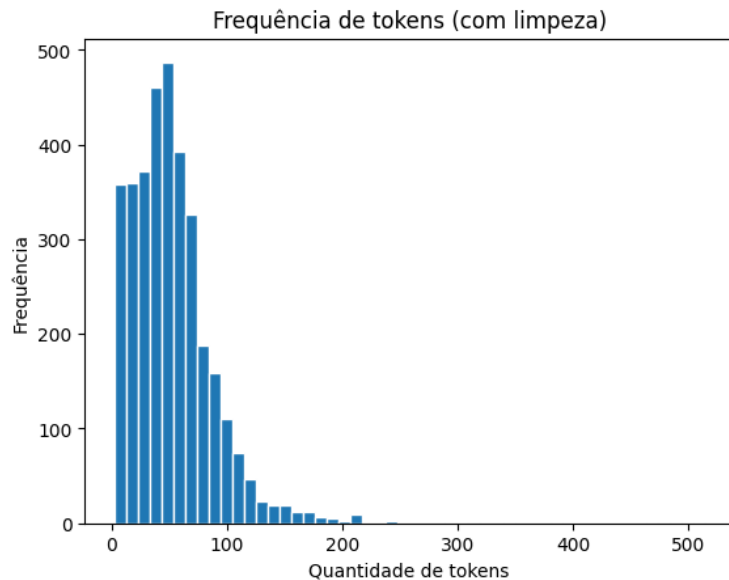


count	3452.000000
mean	3.152086
std	2.265304
min	1.000000
25%	2.000000
50%	3.000000
75%	4.000000
max	57.000000



Análise Exploratória

Quantidade de tokens



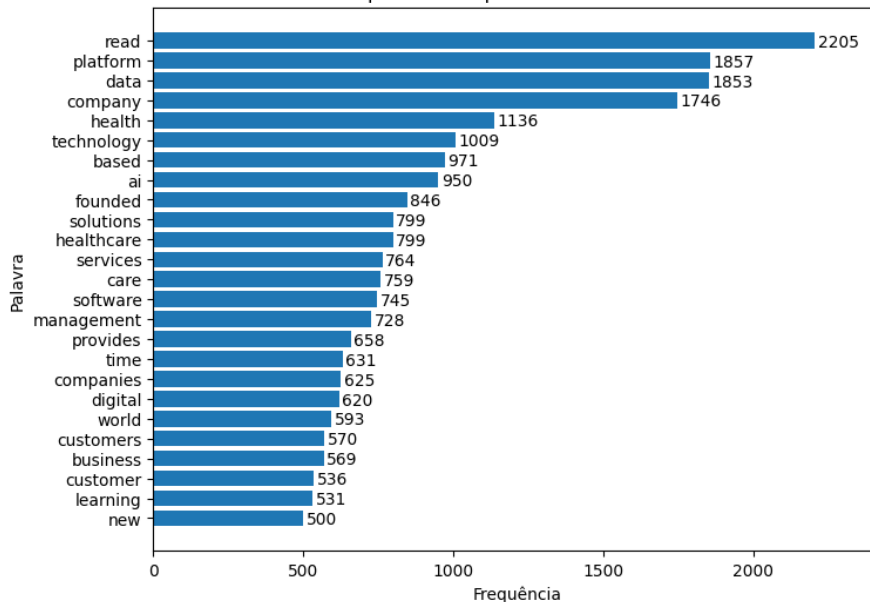
count	3452.000000
mean	51.980881
std	37.616130
min	2.000000
25%	27.000000
50%	46.000000
75%	68.000000
max	514.000000

Análise Exploratória

Principais Tokens

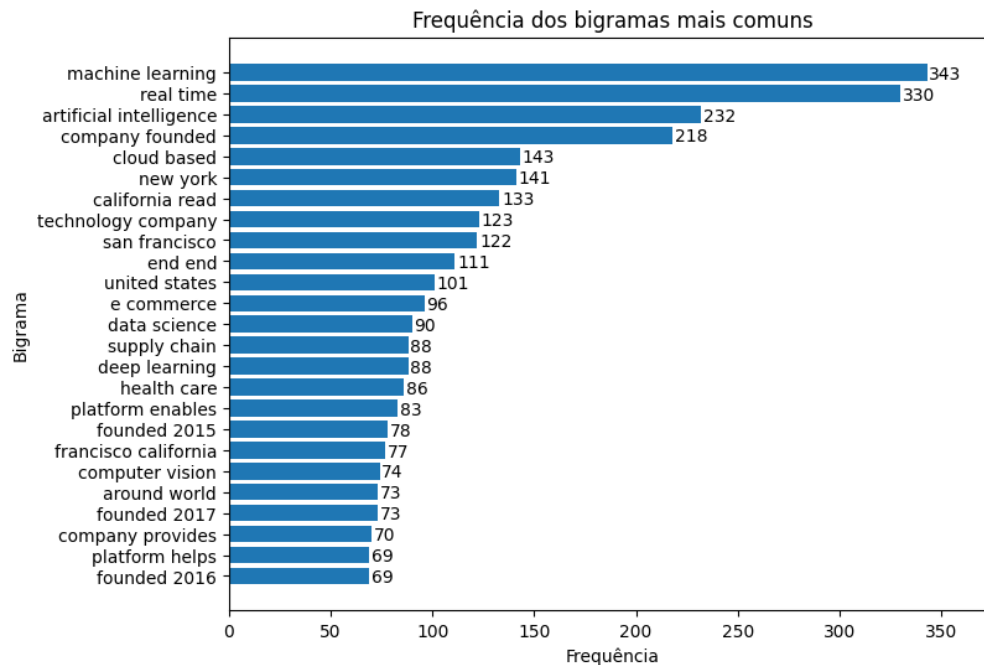


Frequência das palavras mais comuns



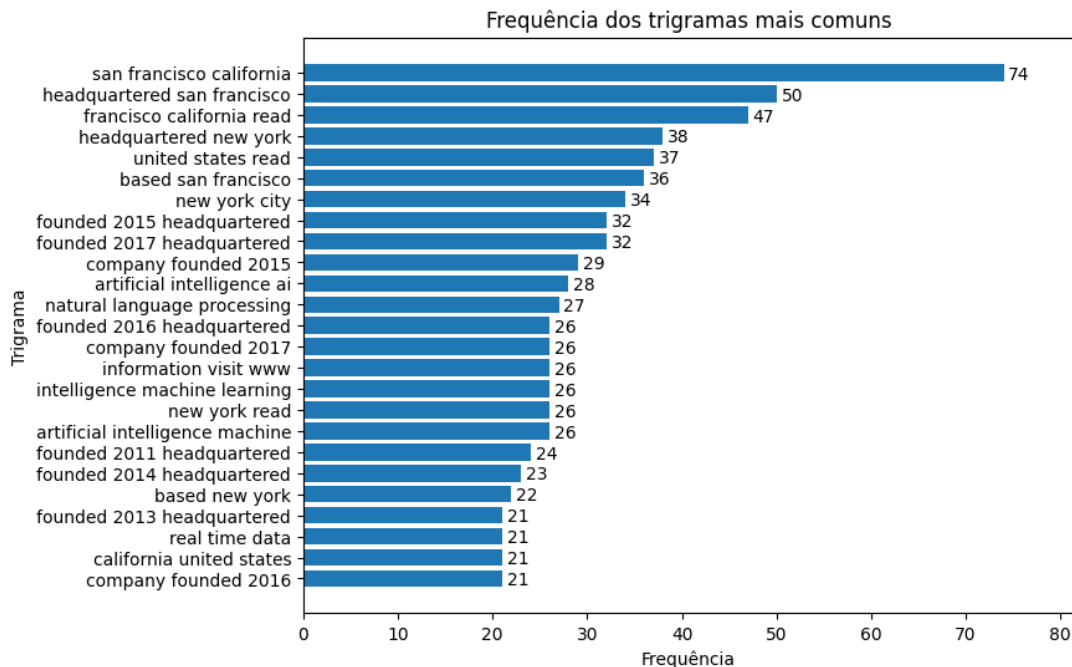
Análise Exploratória

Principais Bigramas



Análise Exploratória

Principais Trigramas



Análise de Tópicos

Latent Dirichlet Allocation (LDA)



platform
solution
patient
health
company
care
healthcare
read
service
data



company
platform
read
data
brand
customer
business
digital
ai



world
technology
data
business
company
customer
platform
read
ai



brand
read
customer
online
platform
logistics
return
retailer
product
delivery



company
platform
business
read
security
service
cloud
management
payment
solution



therapeutic
drug
company
disease
read
treatment
product
development
developing
discovery

- Número de tópicos testados: 2 a 10
- Número de tópicos com maior coerência: 6

Análise de Tópicos

BERTopic

Topic	Count	Name
0	-1	867 -1_data_read_platform_company
1	0	1264 0_health_care_healthcare_medical
2	1	276 1_security_data_cloud_risk
3	2	126 2_payments_payment_financial_businesses
4	3	87 3_insurance_claims_underwriting_insurers
5	4	64 4_ai_learning_intelligence_business
6	5	64 5_logistics_freight_fulfillment_shipping
7	6	61 6_retailers_retail_commerce_store
8	7	56 7_training_learning_workforce_workers
9	8	41 8_marketing_customer_data_marketers
10	9	41 9_video_tv_streaming_radio
11	10	39 10_autonomous_driving_vehicles_ai
12	11	37 11_food_waste_plants_conserwater
13	12	36 12_conversational_ai_voice_assistants
14	13	33 13_advertising_media_ad_publishers
15	14	33 14_climate_weather_carbon_emissions

16	15	31	15_reality_ar_augmented_blippar
17	16	29	16_science_data_machine_learning
18	17	24	17_content_cms_quintype_headless
19	18	23	18_delivery_grocery_order_restaurant
20	19	22	19_printing_3d_manufacturing_additive
21	20	21	20_blockchain_development_aigatha_computing
22	21	21	21_energy_grid_buildings_building
23	22	18	22_cleaning_laundry_service_dry
24	23	17	23_safety_concrete_sensor_sensors
25	24	17	24_lidar_radar_autonomous_sensing
26	25	16	25_workers_productivity_parsable_realwear
27	26	16	26_air_light_molekule_lighting
28	27	15	27_video_content_valossa_ai
29	28	13	28_retailers_vida_size_online
30	29	12	29_aifi_profitability_retailers_intellimize
31	30	11	30_loyalty_programs_giift_redefining
32	31	11	31_aerial_drone_swift_propeller
33	32	10	32_orion_voice_customer_plivo



Análise de Tópicos

BERTopic



patients
healthcare
platform
read
patient
clinical
health
care
company

privacy
secure
security
cyber
cloud
risk
cybersecurity
identity
data
fraud

businesses
payment
banking
financial
cash
fintech
payments
mobile
card
credit

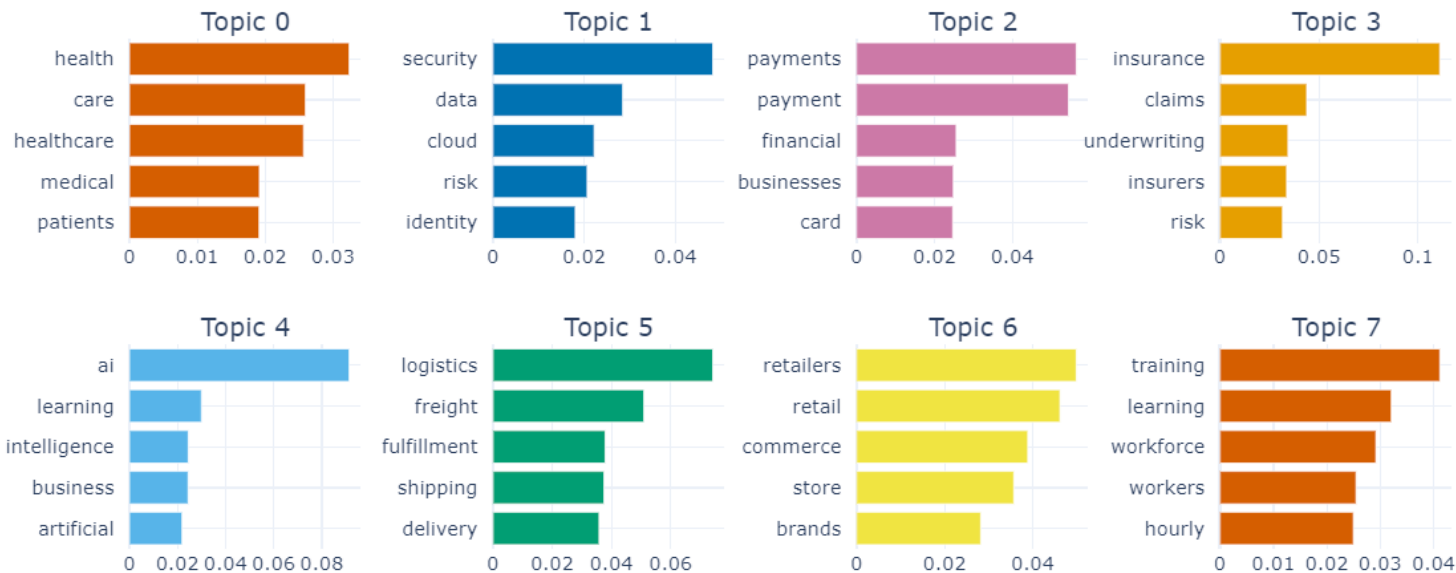
underwriting
read
insurers
platform
insurance
claims
risk
policy
carriers
claim

business
learning
intelligence
ai
artificial
machine
deep
ml
models
data

Análise de Tópicos

BERTopic

Topic Word Scores



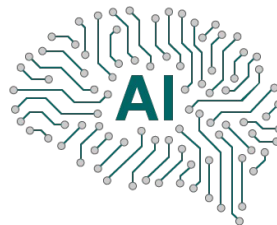
Extração de Informação

Question Answering

Principais tópicos:

Qual o nome das empresas que trabalham com IA?

0_health_care_healthcare_medical	1264
-1_data_read_platform_company	867
1_security_data_cloud_risk	276
2_payments_payment_financial_businesses	126
3_insurance_claims_underwriting_insurers	87
4_ai_learning_intelligence_business	64
5_logistics_freight_fulfillment_shipping	64
6_retailers_retail_commerce_store	61
7_training_learning_workforce_workers	56
9_video_tv_streaming_radio	41





Extração de Informação

Question Answering



Empresas que trabalham com IA

	desc	Probability
	Skan.ai uses cognitive technologies to map, model, and manage business processes by creating digital twins of human/digital interactions. Skan.ai ...	0.814415
	DeepSig is an early-stage startup that develops deep learning software to reinvent wireless communications. By replacing core wireless technology ...	0.793050
	Graphcore is a startup that develops a microprocessor designed for AI and machine learning applications. It has created a new processor, the Intel...	1.000000
	Syntiant is moving artificial intelligence and machine learning from the cloud to edge devices. Syntiant's advanced chip solutions merge deep lear...	0.896949
	AI21 Labs specializes in developing AI systems with an unprecedented capacity to understand and generate natural language. AI21 Labs was co-founde...	0.822778

Probability

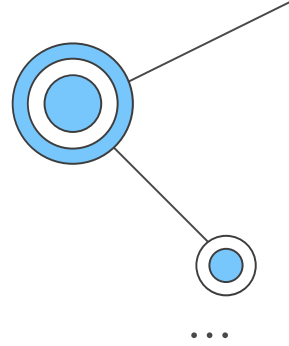
count	64.000000
mean	0.902791
std	0.072346
min	0.793050
25%	0.837425
50%	0.894372
75%	0.971484
max	1.000000



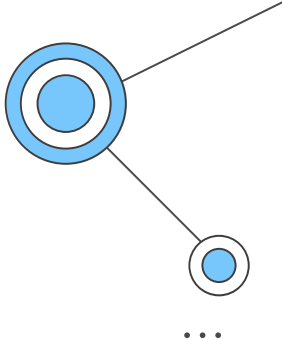
Extração de Informação

Question Answering

Nome das empresas que trabalham com IA



desc	company_name
Skan.ai uses cognitive technologies to map, model, and manage business processes by creating digital twins of human/digital interactions. Skan.ai ...	Skan.ai
DeepSig is an early-stage startup that develops deep learning software to reinvent wireless communications. By replacing core wireless technology ...	DeepSig
Graphcore is a startup that develops a microprocessor designed for AI and machine learning applications. It has created a new processor, the Intel...	Graphcore
Syntiant is moving artificial intelligence and machine learning from the cloud to edge devices. Syntiant's advanced chip solutions merge deep lear...	Syntiant
AI21 Labs specializes in developing AI systems with an unprecedented capacity to understand and generate natural language. AI21 Labs was co-founde...	AI21 Labs
Beyond Limits is an Industrial and Enterprise-grade AI technology company that covers the full range of Artificial Intelligence, purpose-built for...	Beyond Limits
Noodle.ai offers pioneering business solutions in Enterprise Artificial Intelligence, a unique collaboration among business executives, process ex...	Noodle.ai
Explosion is a software company specializing in developer tools for Artificial Intelligence and Natural Language Processing.	Explosion
Tenstorrent is enabling a new era in AI and deep learning with its breakthrough processor architecture and software. The company's mission is to d...	Tenstorrent
Pixis is a California-based technology company that provides codeless AI infrastructure to enable customers to scale accurate data-driven marketin...	Pixis



Empresas que trabalham com IA





Desvendando Empresas:

Análise de Dados Textuais com NLP ...

Obrigado

Perguntas?

