

研究报告

1

D HQU

D HQU

31 28

3134

D HQU

D HQU

HQU4

D HQU

HQU4/6

P BJ

D HQU

D HQU

HQU4/6

H N

D HQU BQI

H

C

H !I

U !N

D !N

R !M

J !

T

BJ

BJ

D HQU

BJ

BJ

B

!H

!J

BHJ

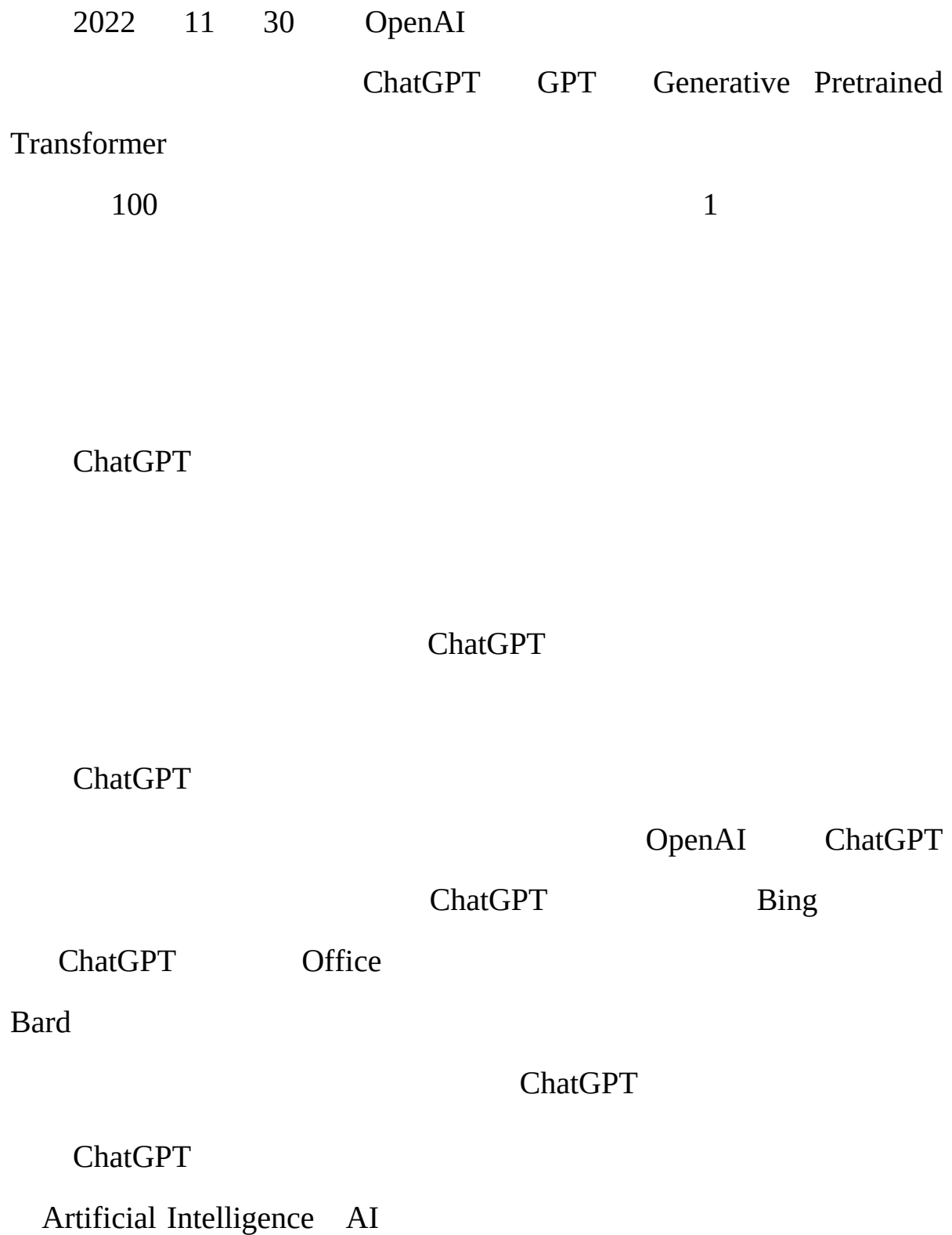
HQU5

D HQU



TSINGHUA PBCSF
清华五道口

1-1				7
2-1	AI	ChatGPT		9
2-2	ChatGPT			10
2-3		ChatGPT		11
3-1				17
3-2				19
3-3				27
3-4				28
4-1				29
4-2				30
4-3	ChatGPT			32
4-4		AI		33
4-5				34
4-6	ChatGPT	Plugins		35
1-1				5
3-1	ChatGPT			15
3-2				21
3-3				23
3-4				24



National Science and Technology Council NSTC 2016

2019

2020

2017

2020

National Science Foundation NSF

1.4

AI

2023 2 24

ChatGPT

ChatGPT

ChatGPT

ChatGPT

ChatGPT

GPT

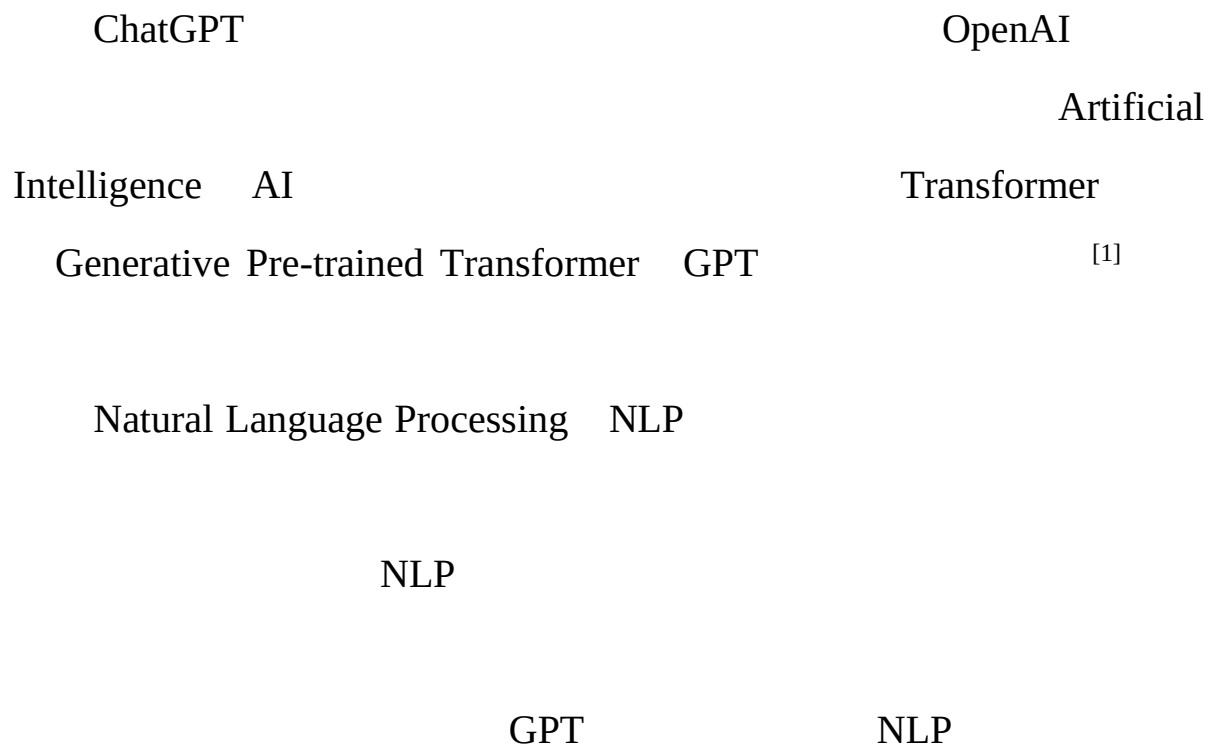
Wikipedia Reddit

ChatGPT

ChatGPT

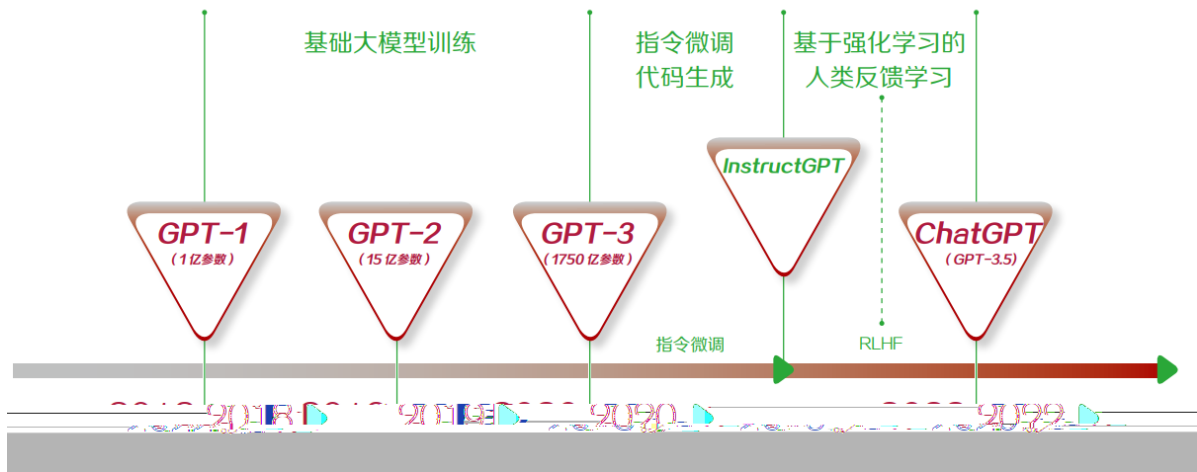
D HQU

D HQU



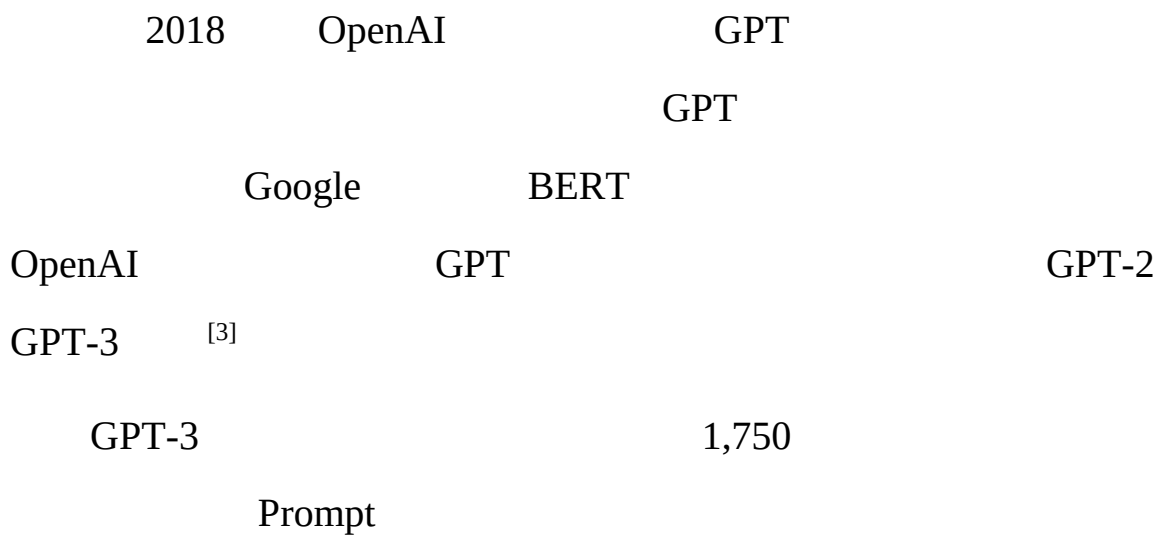
						Oracle
						Google
						OpenAI Microsoft
			20	50	1-1	

Representations from Transformers BERT Bidirectional Encoder
GPT



ChatGPT

1-1



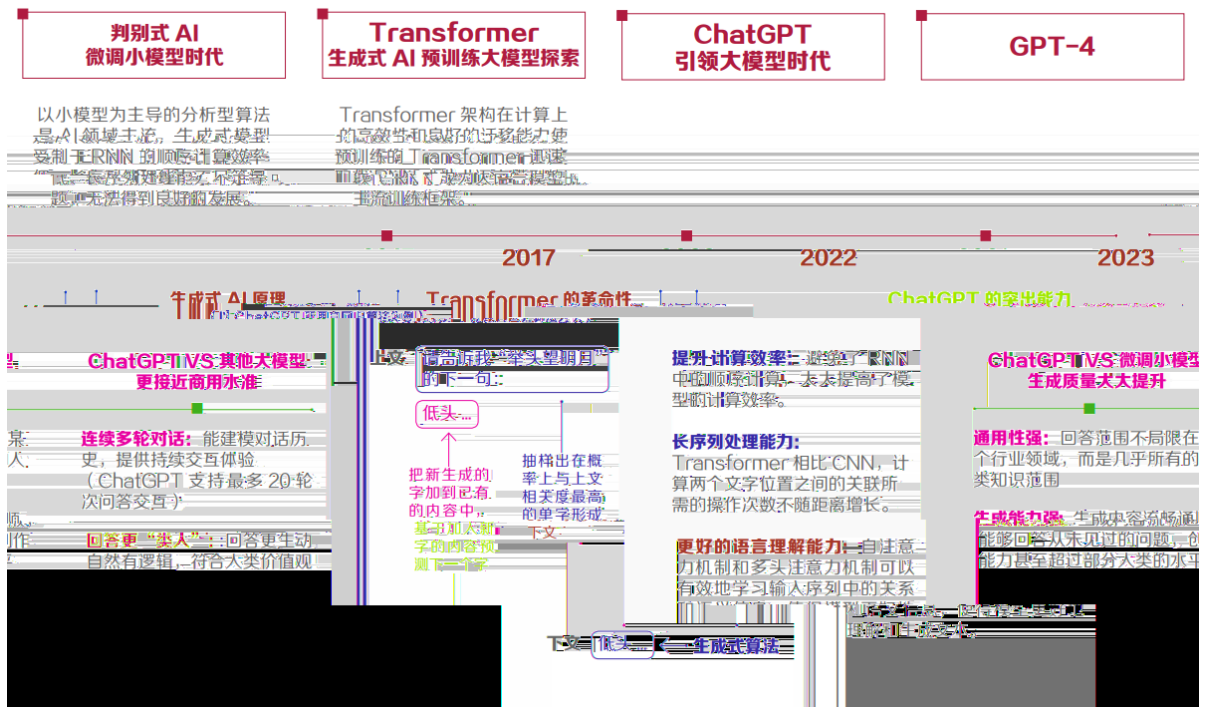
ChatGPT

—

GPT-3

In-

[5]



GPT-3.5

ChatGPT

GPT-3.5

Reinforcement learning with human feedback RHLF

[6]

2-2

ChatGPT



ChatGPT

OpenAI

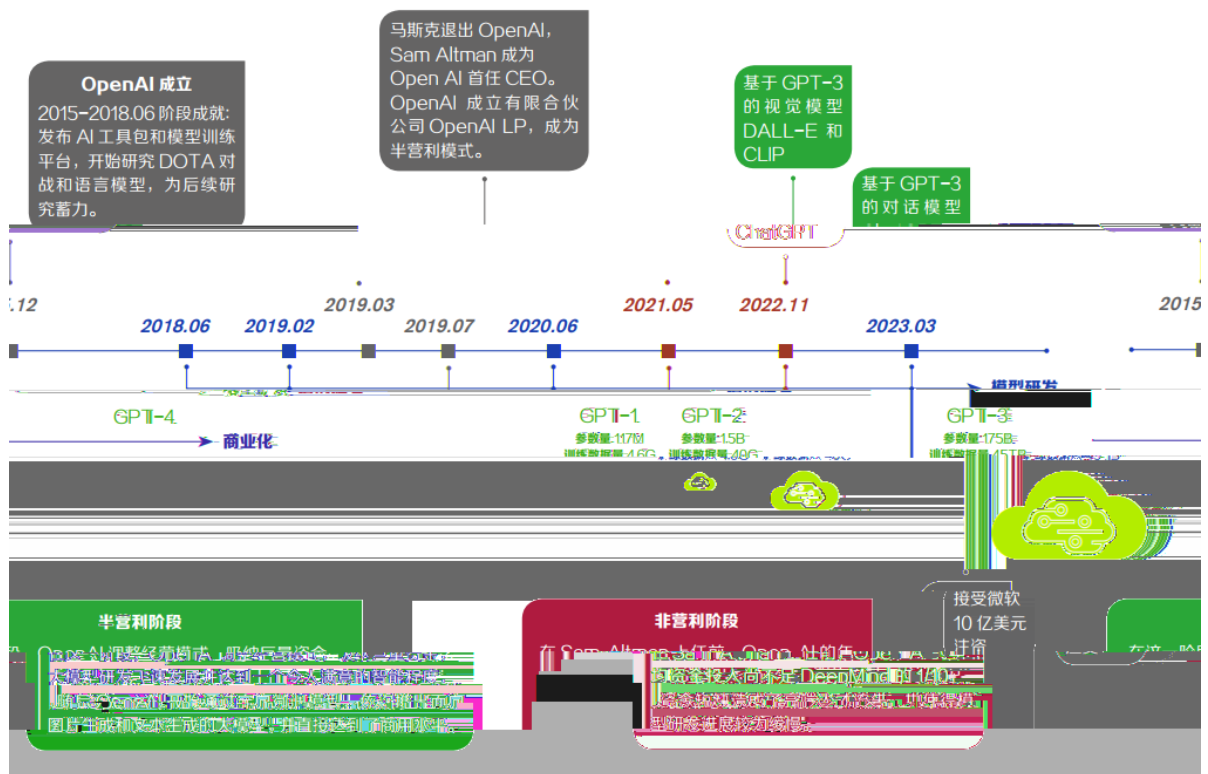
ChatGPT

OpenAI

Artificial General Intelligence AGI

AI

2-3

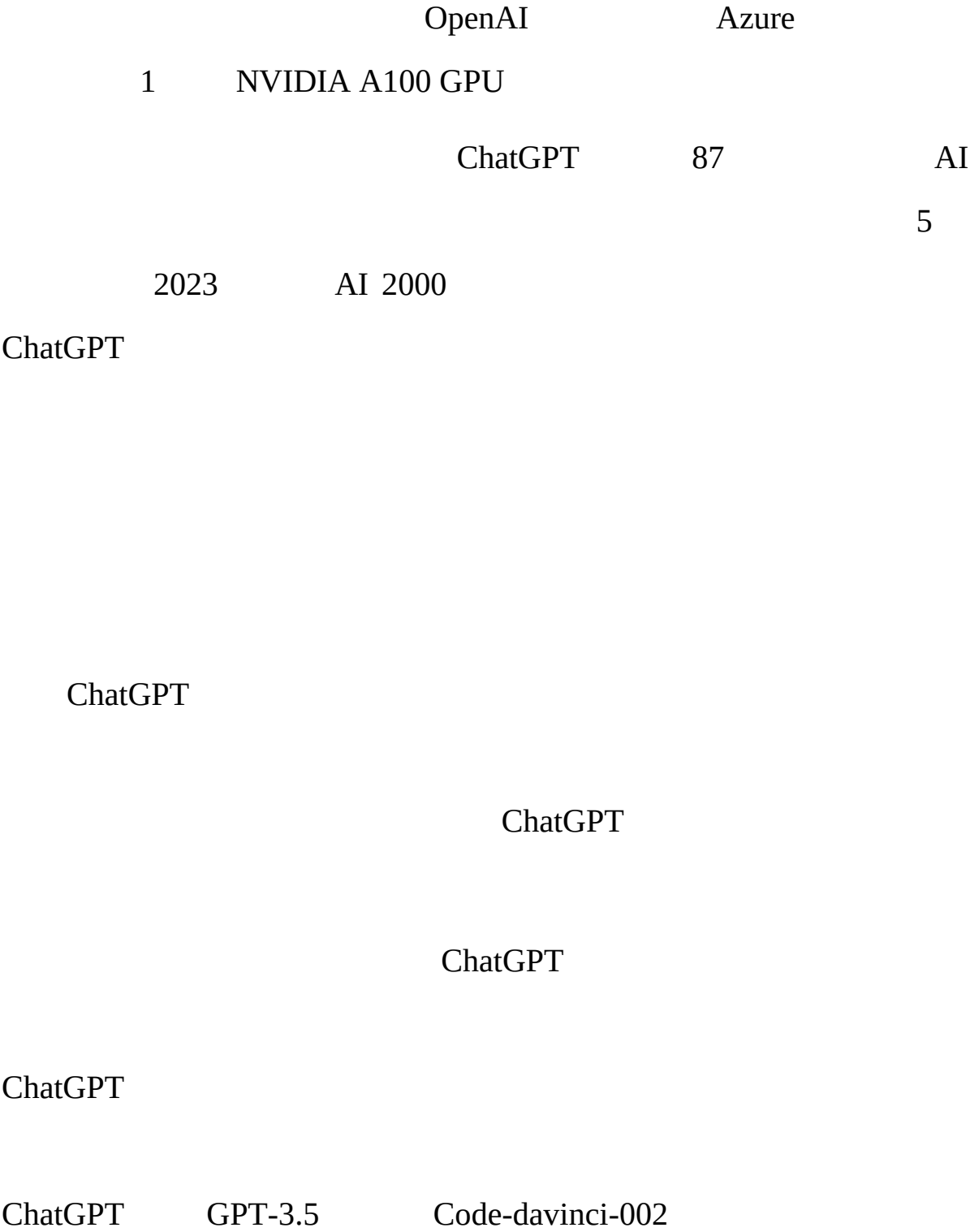


GPT-3

45TB

GPT-1

ChatGPT



scaling law

ChatGPT

[7]

ChatGPT

ChatGPT

ChatGPT

AI

ChatGPT

ChatGPT

ChatGPT

InstructGPT

96%

ChatGPT

ChatGPT

ChatGPT

GPT-3.5

ChatGPT

[8]

ChatGPT

ChatGPT

Beam Search

3-1

ChatGPT

ChatGPT

RLHF

ChatGPT OpenAI

OpenAI

ChatGPT

ChatGPT

Natural Language

Understanding NLU

3-1

ChatGPT

2023 2 24

ChatGPT

NLP

	ChatGPT	
	2022 11 21 12 18	
5.3 1	3.1415926...5.3 5.3 3.1415926... 1A = r ^2=3.1415926... x 1 ^ 2 3.1415926...	
	AI	

2023	3	OpenAI	GPT-4	ChatGPT
------	---	--------	-------	---------

Application Programming Interface API

B 360

Large Language Model LLM

LLM

[9]



3-1

AI

know-how

ChatGPT

AI

AIGC

AIGC

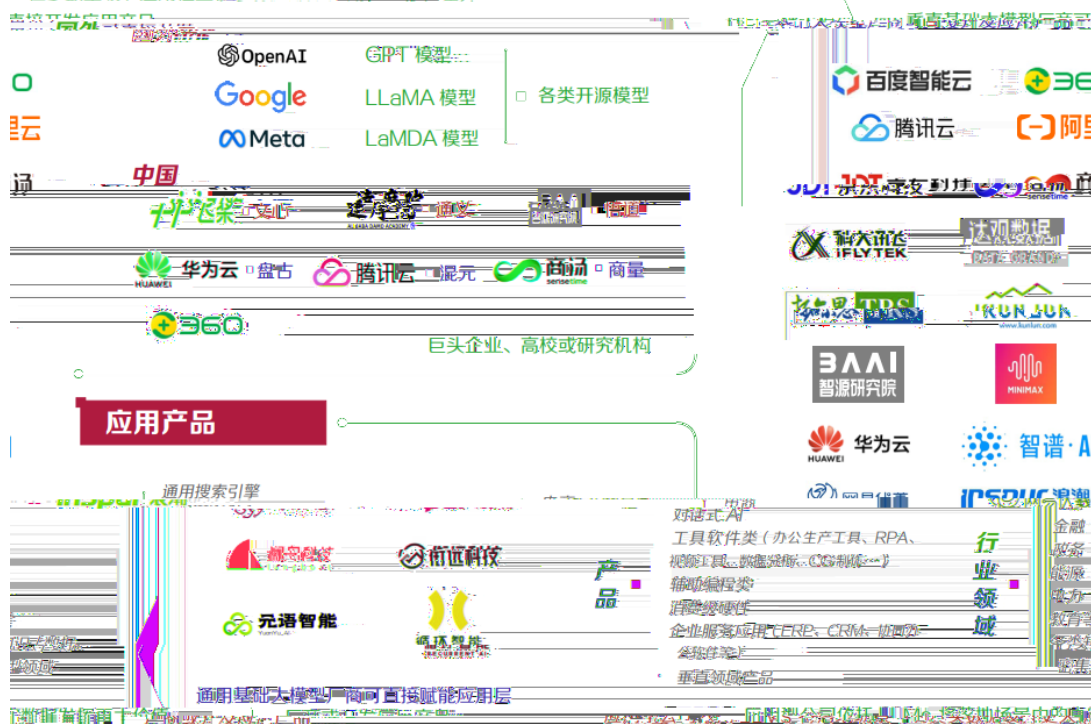
3-2

LLMs



垂直基础大模型、工具平台

中国地质大学(北京)



算力基础设施



Programming Interface API Application
3-2

3-2

API
OpenAI Google

3-3

ChatGPT

ChatGPT /

ChatGPT

https://www.thepaper.cn/newsDetail_forward_23592202

OpenAI	davinci	175B	2048	API	None
	text-davinci-001	175B	2048	API	
	code-davinci-002	175B	8000	API	
	text-davinci-002	175B	4000	API	+
	text-davinci-003	175B	4000	API	+ + RLHF
	text-curie-001	6.7B	2048	API	
	text-babbage-001	1.3B	2048	API	
	text-ada-001	350M	2048	API	
Google	LaMDA	137B			
	Flan-LaMDA	137B			
	T5	11B	512		
	UL2	20B	512		
	Flan-T5	11B	512		
	PaLM	540B			
	U-PaLM	540B			
	Flan-PaLM	540B			
	Flan-U-PaLM	540B			

DeepMind	Sparrow				+
	Chinchilla	70B			+
	Gopher	280B	2048		RLHF
BigScience	BLOOM	175B	2048		
	BLOOMZ	175B	2048		
	T0pp	11B	1024		
	mT0	11B	1024		
Meta	OPT	175B	2048		None
	OPT-IML	175B	2048		
	Galactica	120B	2048		
	LLaMA	65B	8192		
/	TNLG v2	530B	2048		
Eleuther	GPT-J	6B	2048		
	GPT-NeoX	20B	2048		
Cohere	xlarge	52.4B	2048	API	
	large v20220720	13.1B	2048	API	
	medium v20220720	6.1B	2048	API	
	small v20220720	410M	2048	API	
AnthropicAI	Claude	52B	8192		
		26B			

>ž .7N°Aã4ù W Á /" >ž

96;	N°Aã4ù W Á	ø H	EÉ >Kμ ÜL† I	AõL\$ Ĩ E	N°Aã4ù ¢H
Google	CuBERT	-	1024	6\$/Æ	6.6M
äE¥	CodeBERT	125M	512	6\$/Æ	3.5G
äE¥	GraphCodeBERT	125M	512	6\$/Æ	3.5G
äE¥	CodeGPT	124M	1024	6\$/Æ	Python 1.1M Java 1.6M
Case Western Reserve University	CoText	-	1024	6\$/Æ	-
Salesforce	CodeT5	60M/223M/770M	512	6\$/Æ	8.35G
University of California Los Angeles	PLBART	140M	512	6\$/Æ	655G
Columbia University					
Salesforce	CodeGen	350M/2.7B/6. 1B/16.1B	2048	6\$/Æ	825G
Meta	InCoder	1. 3B/6.7B 300M/2.5B	2048	6\$/Æ	159G
OpenAI	Codex	300M/2.5B 12B	1024	API	159G
DeepMind	AlphaCode	300M/1B 3B/9B/41B	1536	` ¢ 6	715.1G
„ p	PanGuCoder	317M/2.6B	1024	` ¢ 6	147G
\$; „] œ	CodeGeeX	13B	2048	` ¢ 6	-
aiXcoder	aiXcoder L	1.3B	-	` ¢ 6	-
aiXcoder	aiXcoderXL	13B	-	` ¢ 6	-

´ Ê μB#@6] W Á.J ¶ e 3 » L q?..J0¬ p ° >ž] F Ò

B#@6] W Á.J ¶ e 3 » L û ë \$ 8E¹],>ž 3-4 M 0 ¼B#@6]

W Á.J ¶ e,° 3 » L q?..J0¬ p ° ¬ >ž] F Ò > 5Ä 3 p Á ,

/" c5ÄB#@6] W Á.J ¶ e)æ(î > ^1% f1%B#@6] W Á.J ¶ e

/		
	Geoffrey Hinton	GPT BERT
	Tomas Mikolov	Word2Vec
	Chris Manning	
	Quoc Le	Transformer BERT GPT
	Alec Radford	Transformer (GPT)
	Ilya Sutskever	OpenAI
		2.0 1.75 2021 M6 GLM-130B
	CoAI	SentiLARE StoryGPT 12 AI AI

		AI	+
		NLP	NLP
	NLP	CCF-A/B	NLP
		70	NLP
		FudanNLP	FastNLP
	NLP		
		AIR	
		16	1999
			2014 9

OpenAI

AI

800

NVIDIA A100

1500

GPT

Google

TPU v4

1000

1300

Meta

NVIDIA V100

900

1400

AMD Instinct MI100

700

1200

AI

600

NVIDIA

A30

1000

AI

LLMs

ChatGPT

Google

Gopher

LaMDA

Meta

Llama

360

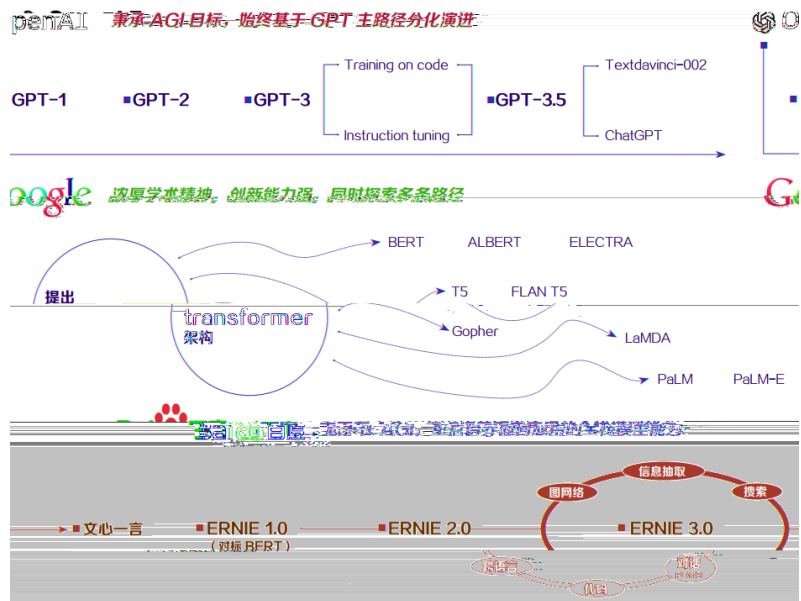
3-3

ChatGPT

Google

ChatGPT

AI



ChatGPT

API

Bing

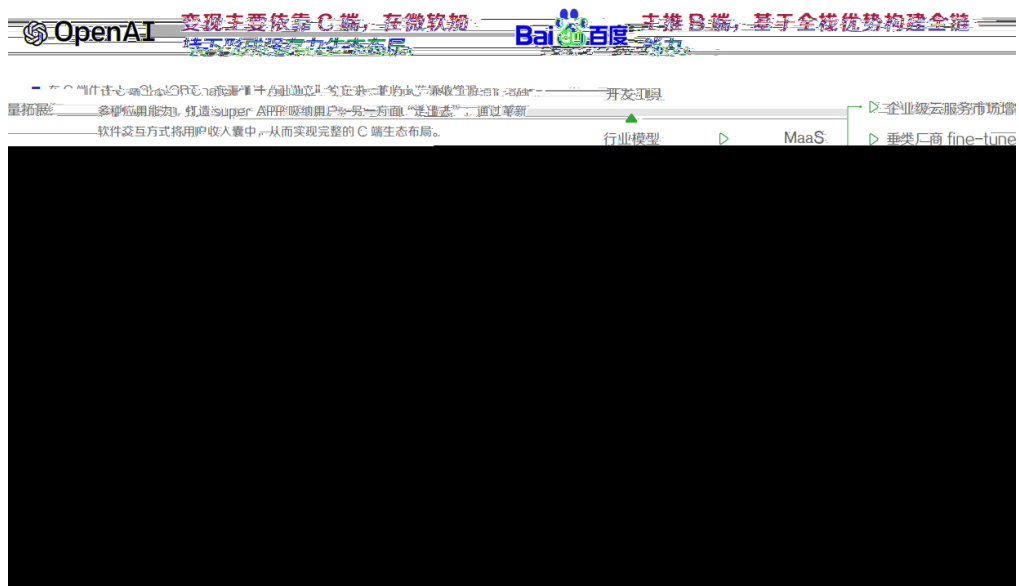
Office

3-4

Google

Google

B



C

ChatGPT

super APP

C

B

OpenAI

Azure

API

B

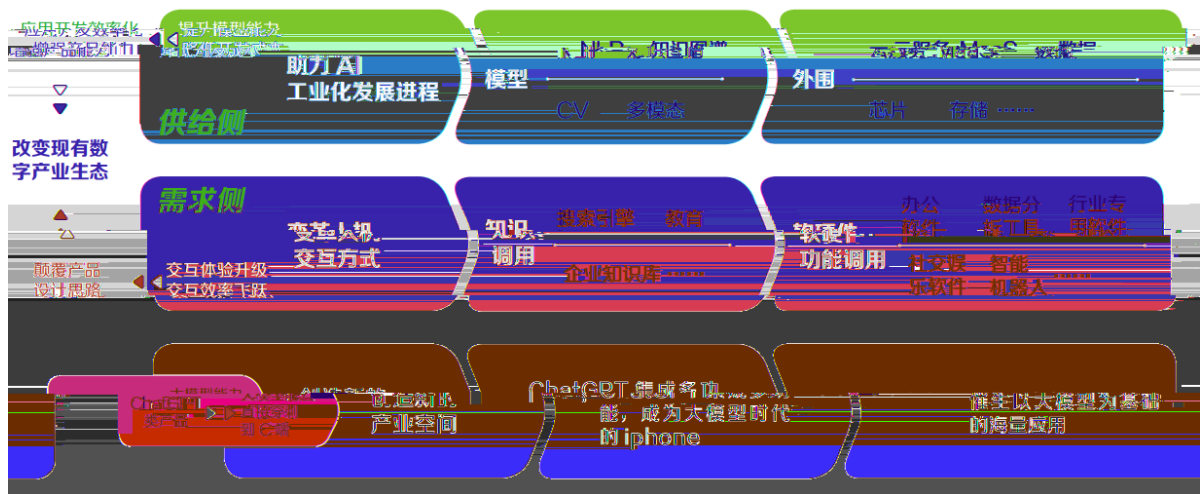
B

B

AI-first

AI

AI



ChatGPT

ChatGPT

ChatGPT

ChatGPT

360

ChatGPT

ChatGPT

ChatGPT

C

ChatGPT

ChatGPT

2023 3 17

Microsoft 365

Copilot

GPT-4

ChatGPT

Office

365

Business Chat

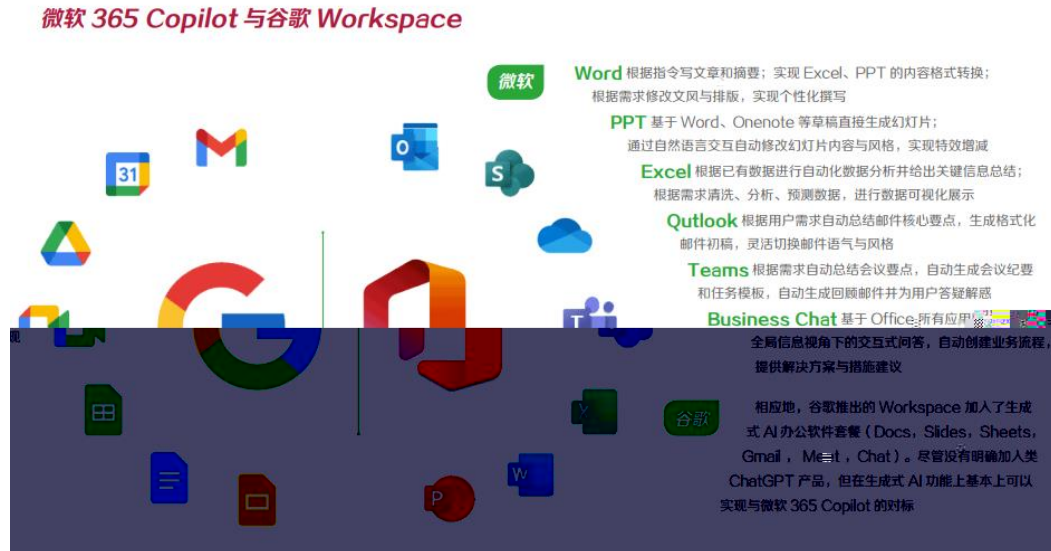
Office 365

Microsoft 365 Copilot

Workspace

AIGC LLMs

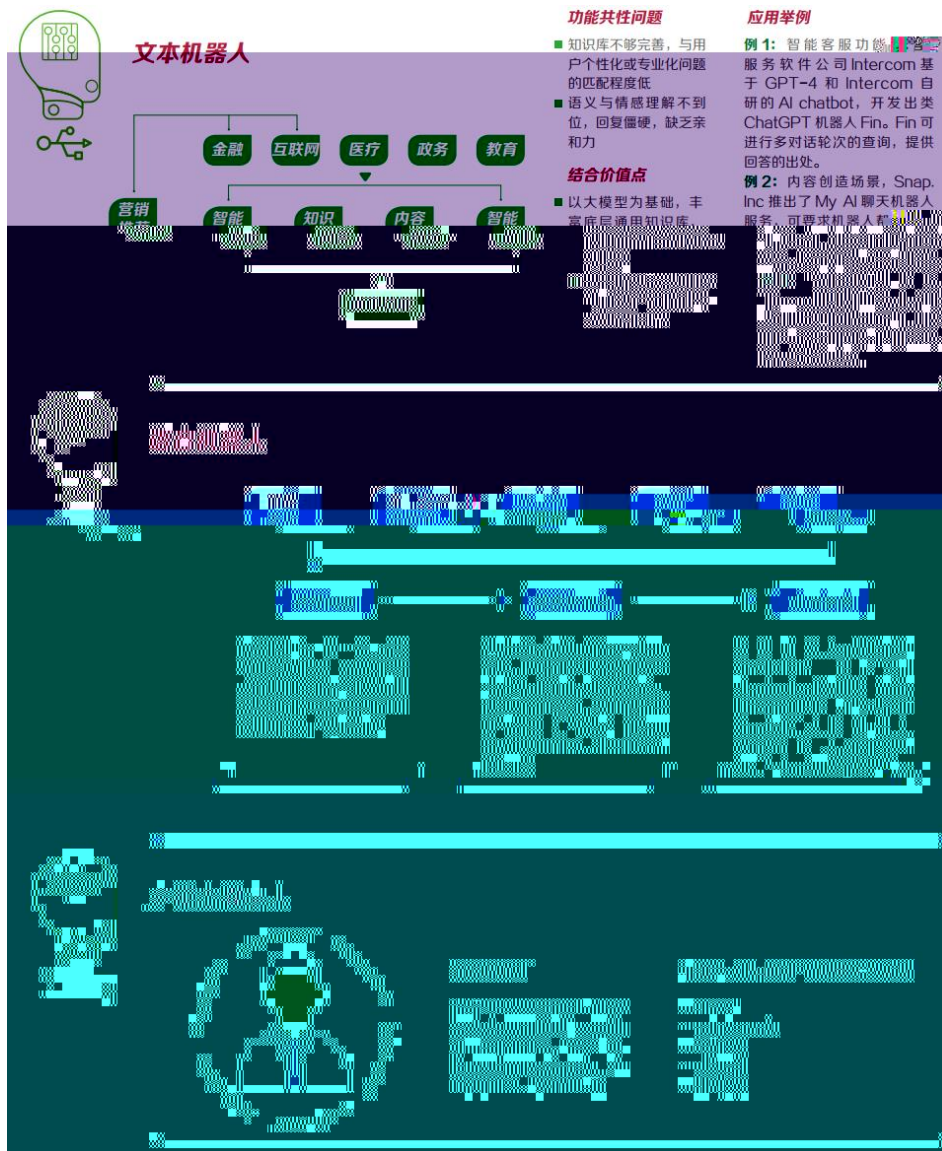
4-3



ChatGPT

AI

4-4



B

ChatGPT

G

AI

G

4-5



2023 3 24

OpenAI

ChatGPT Plugins

ChatGPT

4-6

ChatGPT

11

2

OpenAI

ChatGPT

ChatGPT

ChatGPT

ChatGPT



2022

2023

3

AI

AI

ChatGPT

Midjourney

ChatGPT

AI

GPT-4 AI

(Reinforcement learning from

human feedback RLHF)

OpenAI

ChatGPT

AIGC

AIGC

AI

OpenAI

AIGC

AGI

GPT-4

AGI

AGI

AGI

AGI

AGI

AGI

AGI

AGI

Natural Language Processing NLP

NLP

NLP

Pre-trained Language Model

NLP

Reinforcement Learning with

Human Feedback RLHF

RLHF

Prompt

1. X. B. U. ! ! ! ! ! -! Y !
preprint arXiv:2302.07730, 2023.
2. B! S ! L! O ! U! T ! ! J! T ! J !
language understanding by generative pre- - !31 29/
3. J. Devlin, M.-W. Chang, K. L. ! ! L! U ! C ! Q -training of
! ! ! ! ! -! Y ! !
arXiv:1810.04805, 2018.
4. A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, and I. Sutskever,
M ! ! ! ! ! -!P B! ! !
1, no. 8, p. 9, 2019.
5. Q! M ! X! Z ! K! G ! [! K ! I ! I ! ! H! O ! Q -train,
prompt, and predict: A systematic survey of prompting methods in natural
! -!BDN!D !T ! !66! !: , pp. 1 35,
2023.
6. X!C!L ! !Q!T ! U !U ! ! ! ! ! !
-! !3119!8 !JFFF! ! ! ! !
and learning, IEEE, 2008, pp. 292 297.
7. L! T! L ! B! S ! ! T! T ! Bmmus: A survey of
transformer- ! ! ! ! ! ! -! Y !
preprint arXiv:2108.05542, 2021.
8. H! X ! O! X ! W! M ! ! Q! T ! E ! !
Interactive agent shaping in high- ! ! -! !Qroceedings
of the AAAI conference on artificial intelligence, 2018.
9. K! N H ! ! /! J ! ! ! -dependent human
-! !J !D ! !N !M !QN! !31 28!
pp. 2285 2294.

10.E /N /[! ! /! G -t ! ! ! ! - !

arXiv preprint arXiv:1909.08593, 2019.

/

1

ChatGPT

GPT-3.5

ChatGPT

2

ChatGPT

ChatGPT

3

ChatGPT

Google

4

/



TSINGHUA PBCSF
清华五道口

