

# Prompt Engineering (提示词工程技能) 之原理

讲师：李希沅

关注 玄姐谈AGI 视频号

# 目录

- 1 AIGC的发展和产业前景
- 2 大模型的类型和特点
- 3 大模型技术原理和发展

# **1 AIGC的发展和产业前景**

# 00、前言

## 什么是AI?

人工智能 (Artificial Intelligence) , 英文缩写为AI

## 什么是AGI?

AGI一般是Artificial general intelligence的简写 , 指通用人工智能

## 什么是AI2.0?

ChatGPT的横空出世代表着AI2.0时代, 也就是通用人工智能时代。

## 什么是大模型?

大模型是大规模语言模型 (Large Language Model) 的简称。语言模型是一种人工智能模型, 他被训练成理解和生成人类语言。

“大” 在 “大语言模型” 中的意思是指模型的参数量非常大。比如ChatGPT, 文心一言等。

A large, bold, black serif font spelling out the letters 'AGI'. The letters are highly stylized with thick strokes and sharp serifs, giving it a formal and authoritative appearance.

# 01、AIGC时代的到来

随着计算机硬件算力和算法模型的不断升级，出现了大量的内容生产力工具，比如 ChatGPT、Midjourney、Stable Diffusion等。这些工具大大提高了内容创作的效率，降低了创作门槛，各领域的内容创作者都在经历着一场名为AIGC的时代浪潮。

## 一、PGC

PGC 是指专业生成内容（Professional Generated Content）。它指的是由专业的内容创作者或团队进行创作、编辑和发布的内容。这种创作方式起源于传统媒体时代，如报纸、杂志、电视和电影等。在数字时代，PGC 已经应用到各种领域，如网站、应用程序、短视频和音乐等，比如CCTV、新华社等官媒，或是IT之家、36氪等垂直专业媒体都属于PGC范畴。

## 二、UGC

UGC 是指用户生成内容（User Generated Content）。它指的是由普通用户或受众参与创作、编辑和发布的内容。这种创作方式是由 Web 2.0 时代引起的，随着社交网络和博客的出现而流行起来。UGC 的应用场景包括社交网络、在线论坛、博客、知识共享平台等。UGC 可以是图片、视频、音乐、博客、评论等形式，它可以是个人的、团体的，甚至是全球性的。代表性的社区或应用有：小红书、抖音、百度贴吧等。

## 三、AIGC

AIGC 是指人工智能生成内容（Artificial Intelligence Generated Content）。它是一种新兴的创作方式，利用人工智能技术和自然语言处理技术来生成内容。目前，AIGC 主要应用大致可分为“写作、音频、图像、视频”四大类。



## 02、发展趋势和应用展望

### 1. 发展趋势：

**更强大的模型：** 未来AIGC将会使用更大规模、更强大的预训练模型。随着硬件技术的进步和模型设计的改进，模型的参数和能力将不断增加，使其能够处理更复杂的任务和生成更高质量的内容。

**多模态生成：** AIGC不仅限于文本生成，未来趋势将发展成为多模态生成，即同时生成文本、图像和音频等多种内容。这将为创意产业和多媒体应用带来更多的可能性。

**个性化生成：** 未来AIGC可能朝着更加个性化的方向发展，能够根据用户的偏好和需求生成内容，为用户提供定制化的体验。

### 2. 应用展望：

**内容创作和自动写作：** AIGC将在内容创作领域发挥重要作用，为新闻报道、文章、博客等提供自动生成的内容，辅助作家和编辑的工作。

**娱乐和游戏：** AIGC可以用于游戏剧情生成、虚拟角色对话，为游戏和虚拟现实应用增添更多内容和情节。

**广告和营销：** AIGC可以用于自动生成广告文案、营销宣传语，帮助品牌更有效地与用户进行互动。

**个性化助手和教育：** AIGC可以用于开发个性化的学习助手、虚拟导师等，为教育领域提供更个性化和有效的学习支持。

**自动化客服：** AIGC可以用于自动生成客服回复和答案，提供快速高效的客户支持。



### 03、发展趋势和应用展望

# 会不会是泡沫？

OpenAI创始人Sam Altman:  
AI是少有的被严重炒作之后，还被严重低估的东西

## ② 大模型的类型和特点



# 01、大模型对比

## 最新国产AI大模型大全



讯飞星火

科大讯飞



文心

百度 AI



MiniMax

MiniMax



通义千问

阿里巴巴



ChatGLM

清华&智谱 AI



混元

腾讯 AI



360智脑

360 AI



华为盘古

华为 AI



天工

昆仑万维

@Elena艾麗

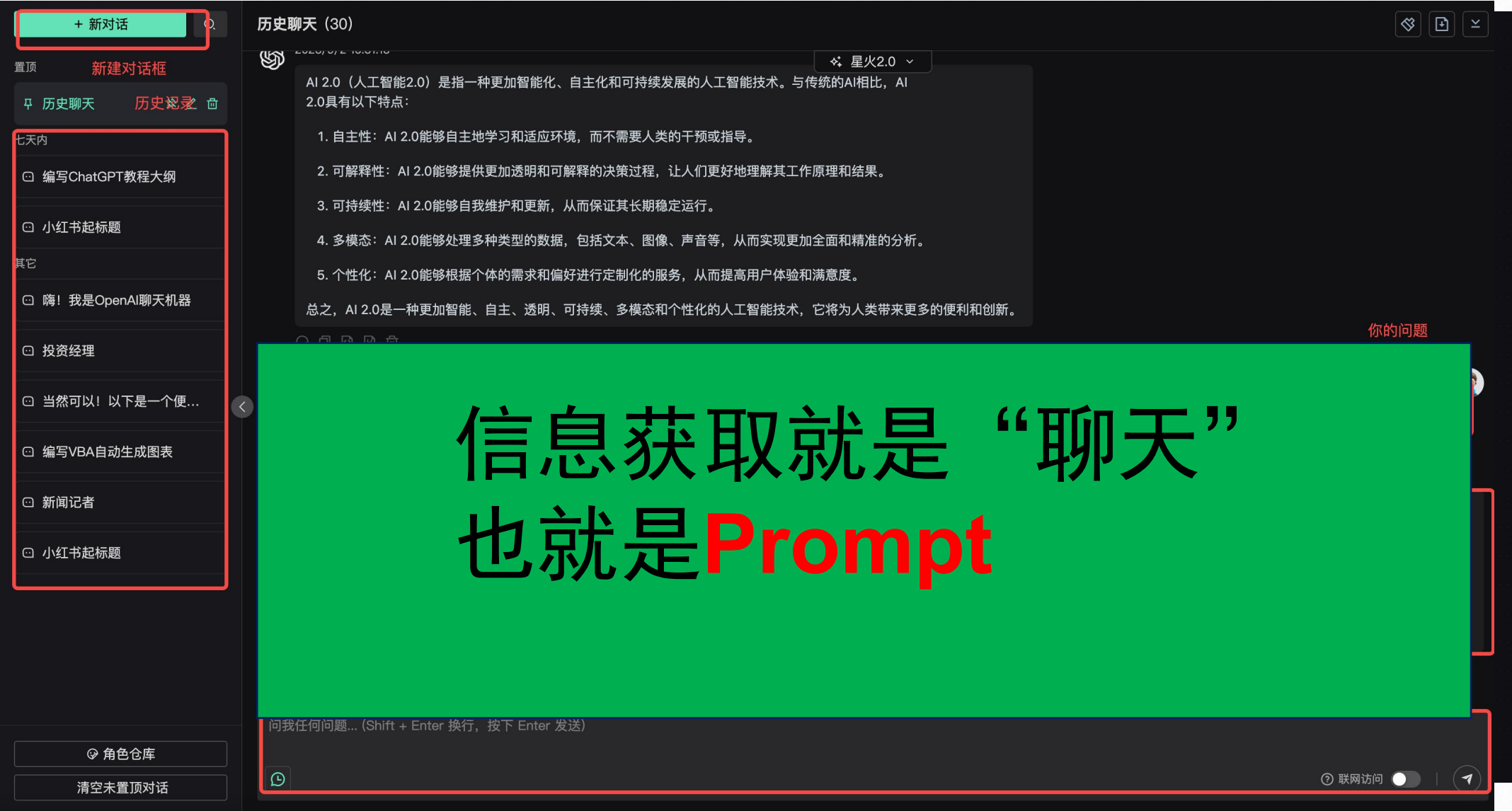
百度在AI方面布局多年，具有一定大模型先发优势。百度的文心一言API调用服务测试的企业已经上亿。在行业大模型上，已经与国网、浦发、人民网等有案例应用。

阿里通义大模型在逻辑运算、编码能力、语音处理方面见长，集团拥有丰富的生态和产品在线，在出行场景、办公场景、购物场景有广泛应用。

**ChatGPT是目前最先进的模型之一，是最流行的模型。**

对比：[https://articles.zsxq.com/id\\_l744s9pztge8.html](https://articles.zsxq.com/id_l744s9pztge8.html)

# 02、上手



### 03、特点

知识丰富

道德感强，法律边界清晰

有记忆力，可联系上下文

知错就改

知人晓事

### **③ 大模型技术原理和发展**



# 01、成语接龙和暴力穷举



床\_\_

后

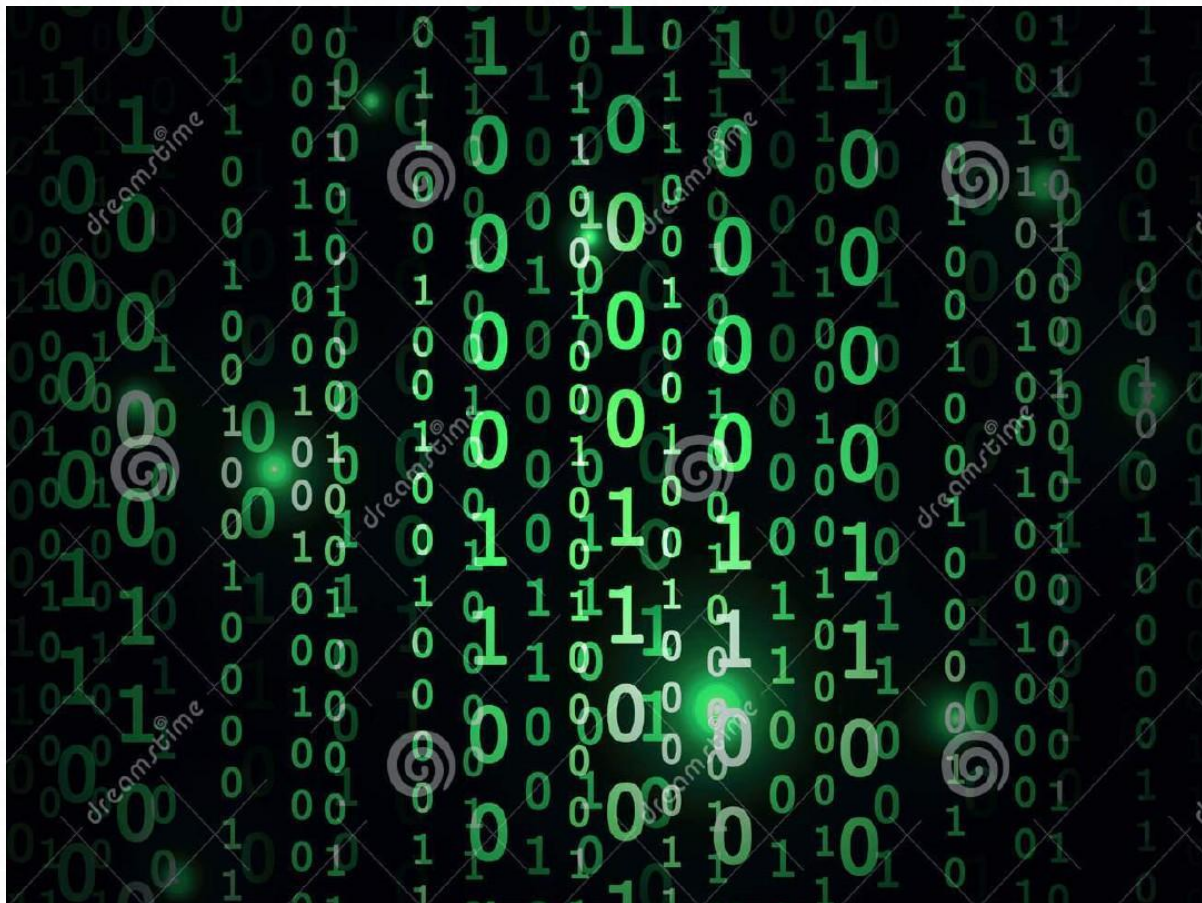
左

前

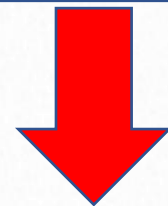
刘慈欣在《诗云》讲了一个非常有意思的故事：一个宇宙超级生命来到地球，有人问它，你能做出超越李白的诗吗？这个生命消耗了无数能量，将所有汉字的每一种排列组合都罗列出来，说：“这里面一定存在超越李白的诗句，可是我找不出来。”这可真是一个绝妙而悲伤的故事。

## 02、ChatGPT如何理解人类语言

计算机底层：二进制

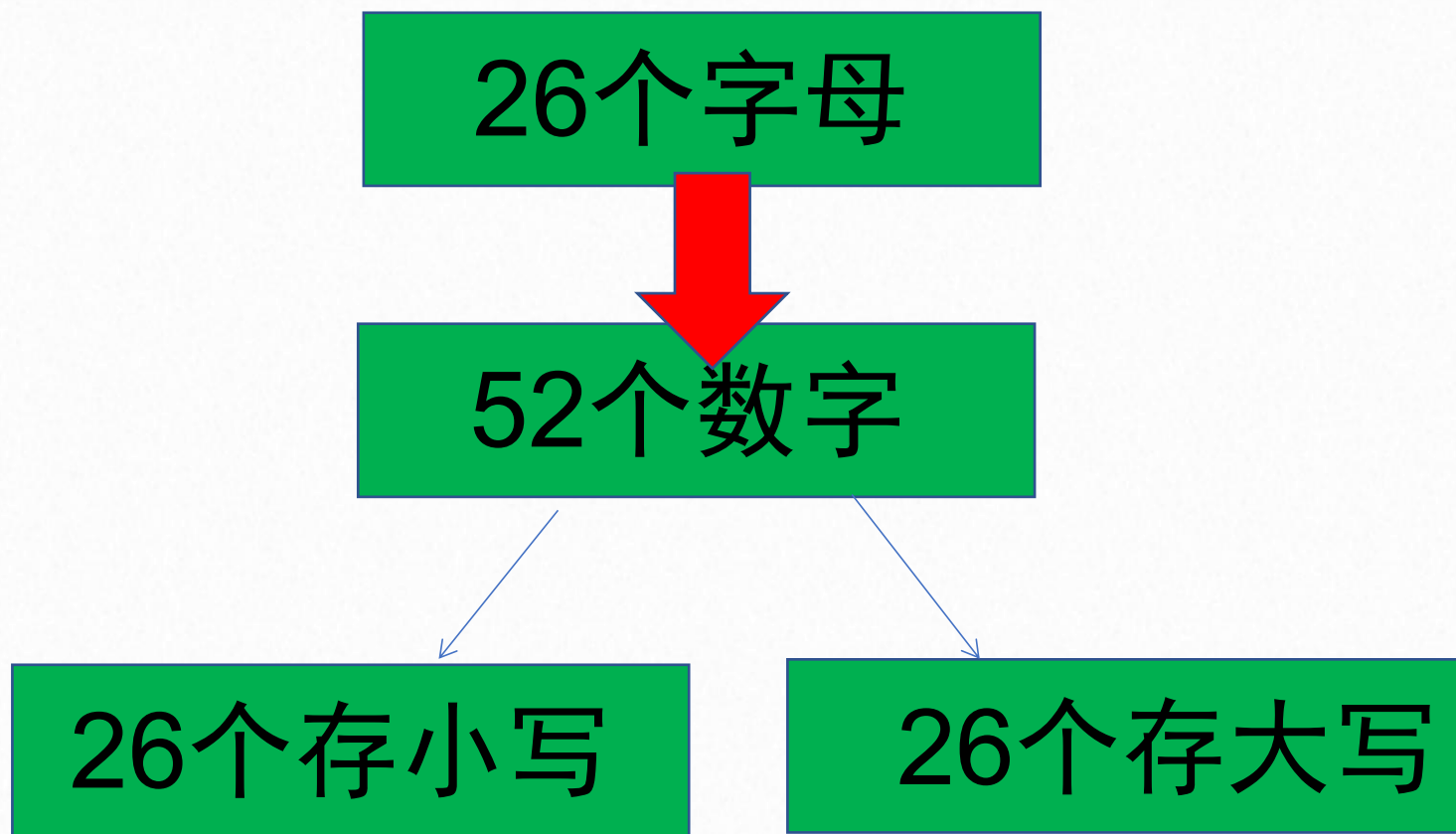


现实问题



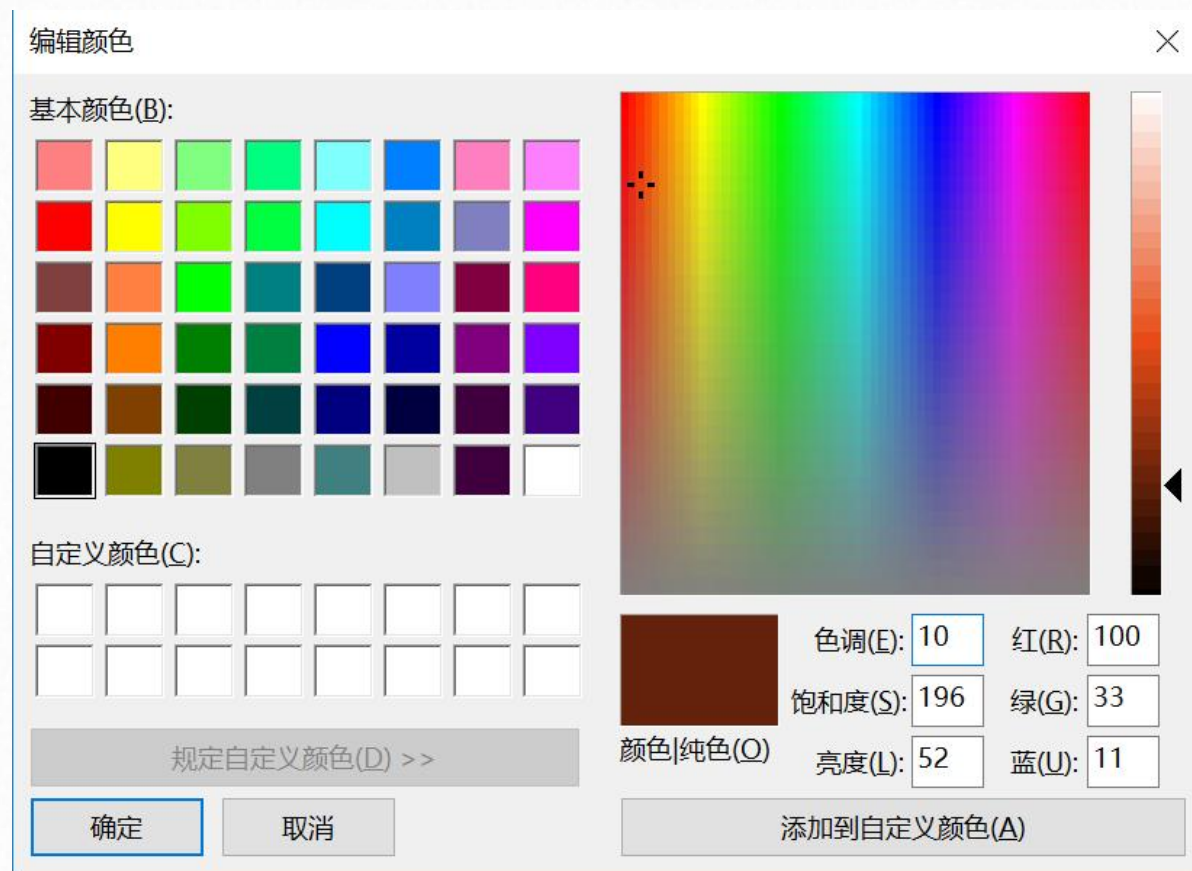
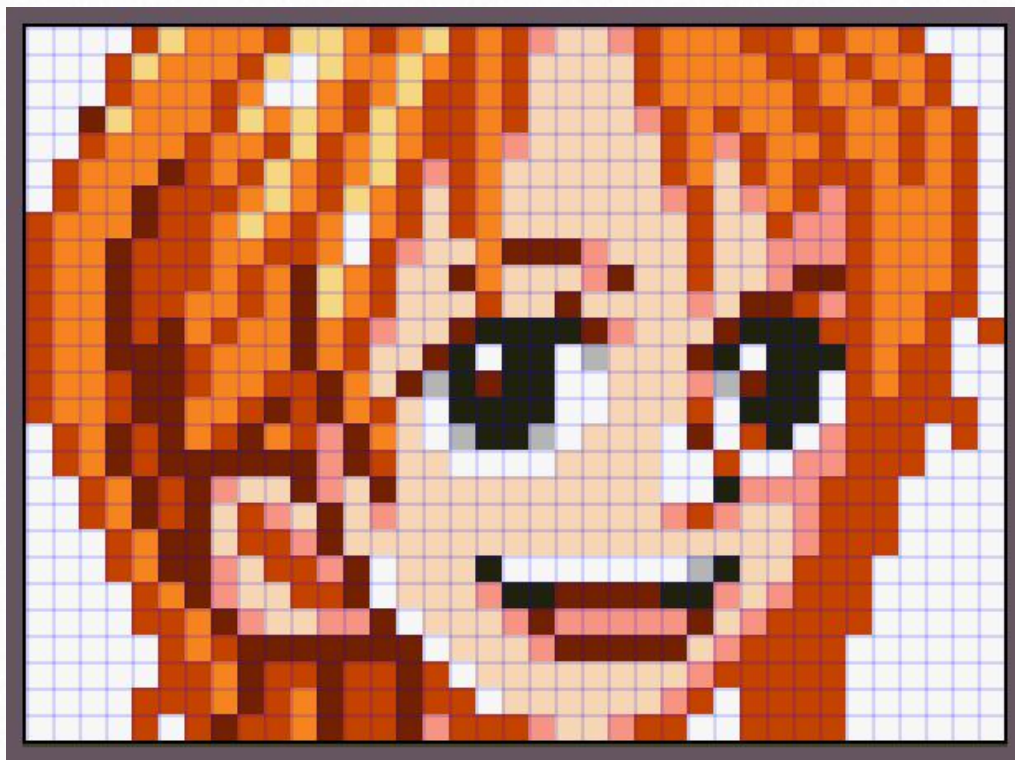
数学问题

### 03、如何存储数据



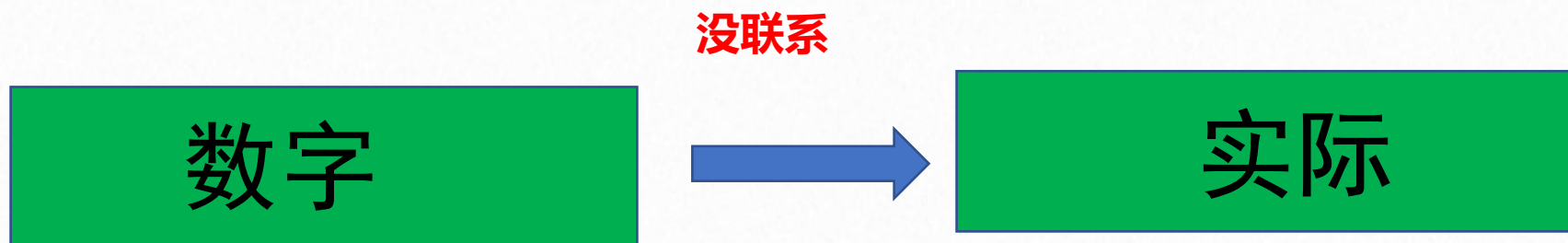


## 04、图像存储





## 05、电脑存数字



## 06、如何让电脑知道存的什么？

向量化

信息压缩与特征提取

## 07、向量化



【0, 180, 75, 35】

0,性别	180,身高
75,体重	35,年龄

## 08、更多纬度

【性别， 年龄， 身高， 体重， 胸围， 肤色， 腰围， 体脂率， 爱好， 语言， 城市， 收入等】

【0, 22, 170, 170, 170, 11, 22, 33, 65, 345, 67, 8888】

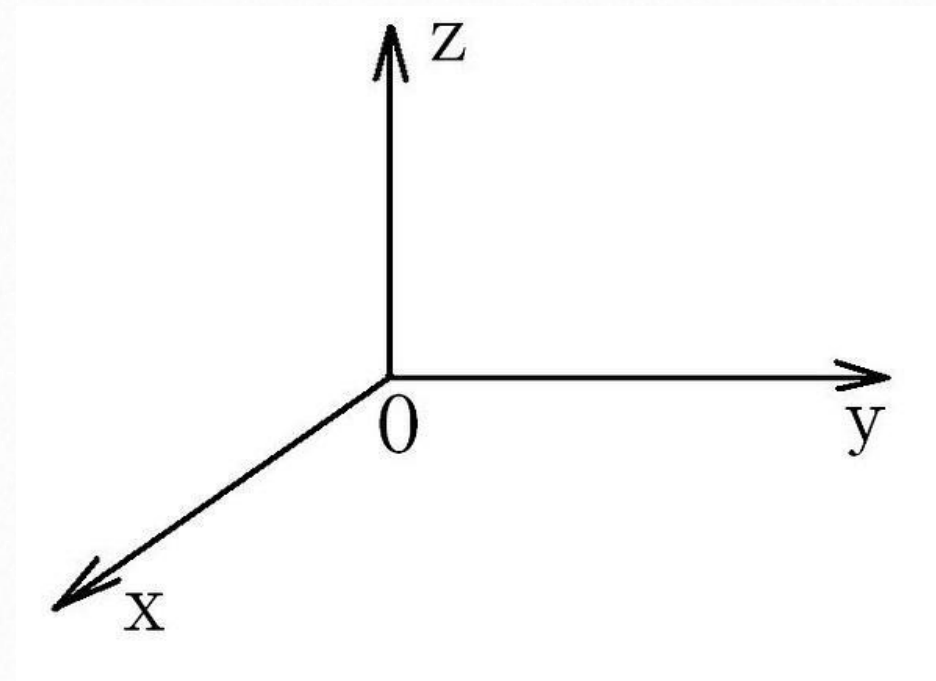
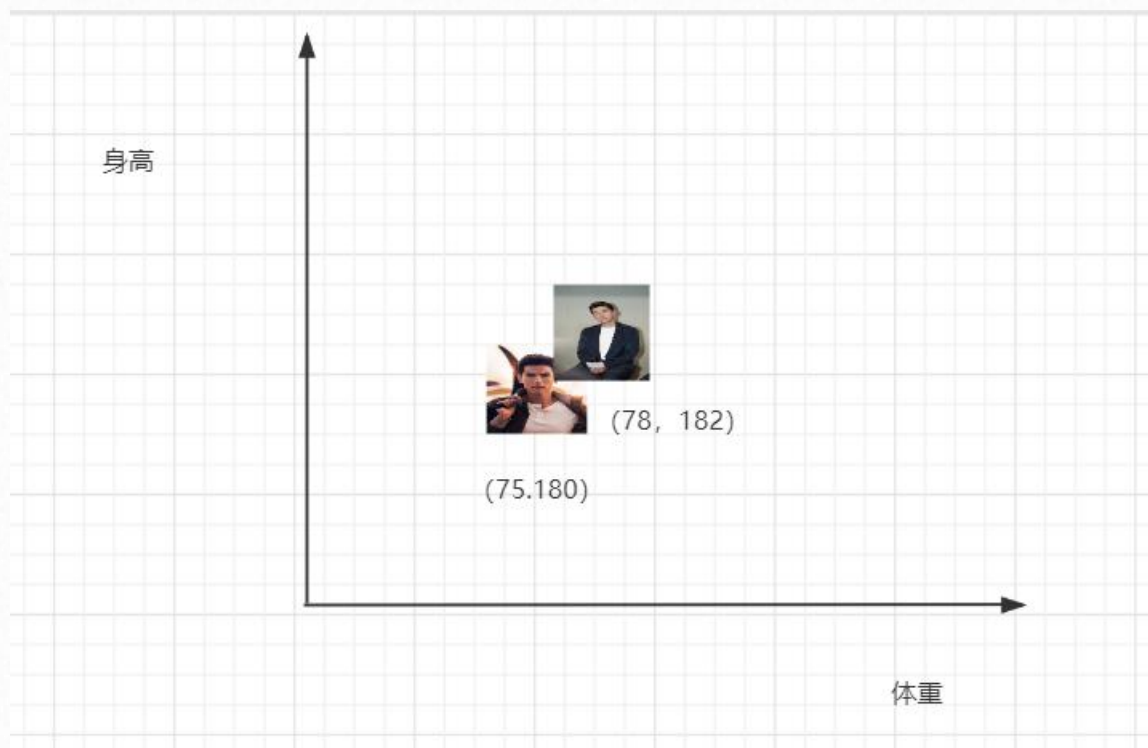


## 09、为什么需要向量化

方便电脑处理

寻找规律

## 10、寻找规律



### 13、每个字都是用数字代替

床

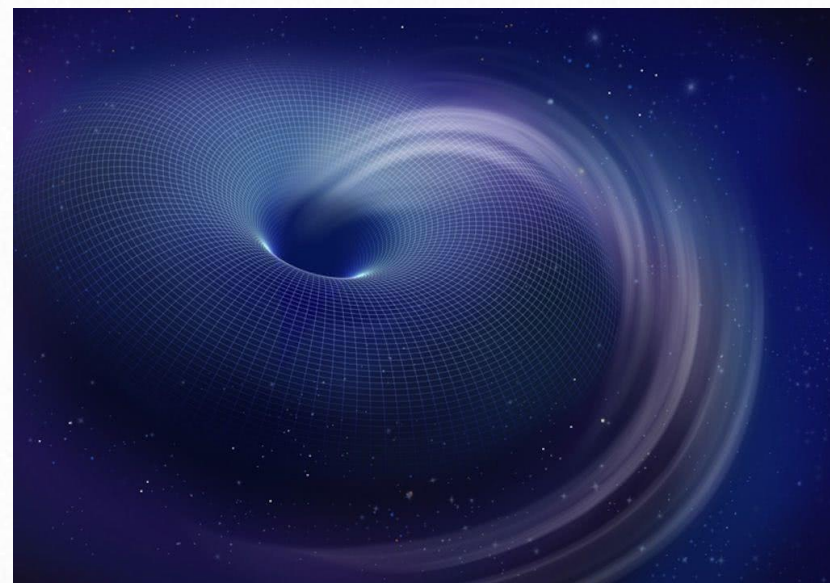
【0, 11, 2, 33, 44, 54, 66, 75, 89, 12】

大

【0, 11, 2.234, .....】

美

【0, 11, 2.222, .....】

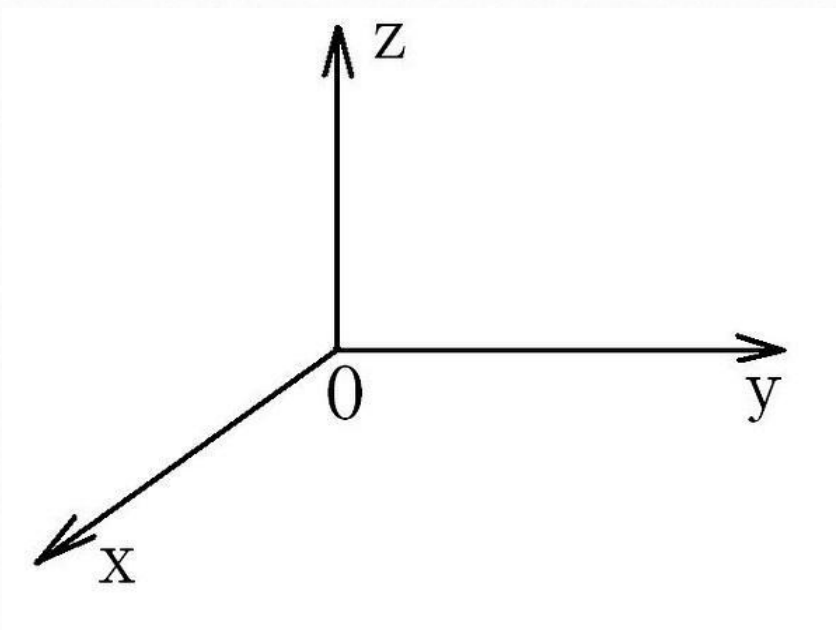


## 14、向量可以计算

相加

相减

相乘



国王减去男人加上女  
人与女王/王后的位置  
非常接近



## 15、出现了难题



坐标要多少维度

每个维度代表什么

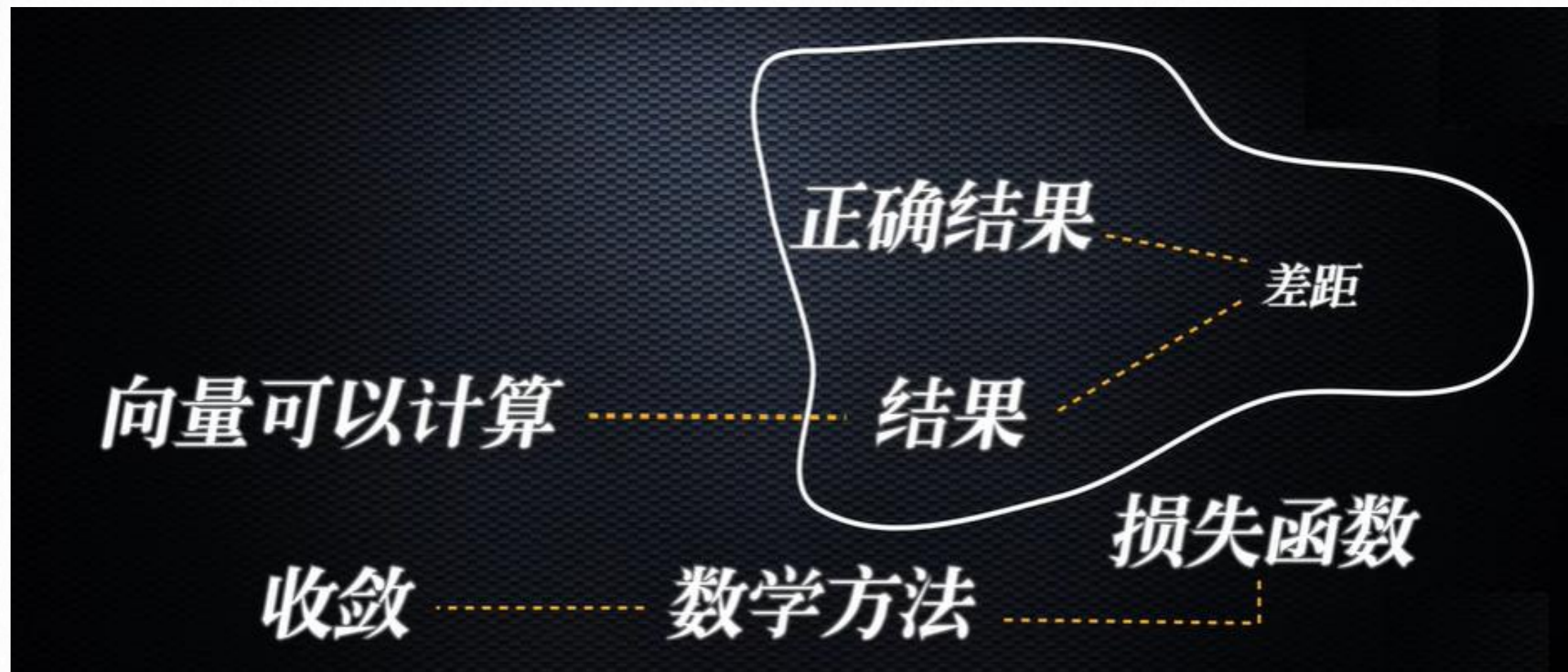
## 16、调整位置

[illegible]

# 50维度 10000个词

# 通过计算距离，不断调整位置

## 17、科学的方式





# 18、Word2vec算法

## Word2vec

两个语言模型：CBOW 和 Skip-Gram

CBOW：根据上下文的词语预测当前词语的出现概率的模型

Skip-Gram：已知当前词语，预测上下文

两个降低复杂度的近似方法：Hierarchical Softmax 和 Negative Sampling

Hierarchical Softmax：负例是二叉树的其他路径

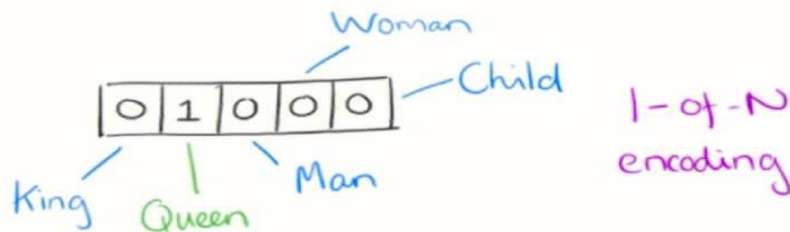
Negative Sampling：负例是随机挑选出来的

### (一) word2vec原理：CBOW与Skip-Gram模型基础

#### 1. 词向量基础

##### One hot representation

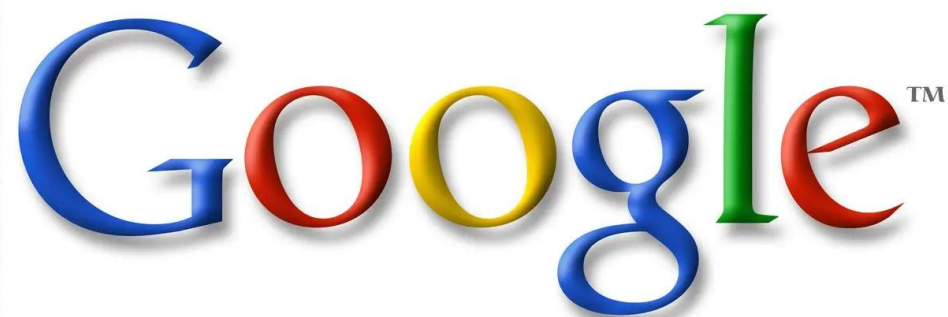
最早的词向量是很冗长的，它使用是词向量维度大小为整个词汇表的大小，对于每个具体的词汇表中的词，将对应的位置置为1。比如有5个词组成的词汇表，词“Queen”的序号为2，其词向量就是(0,1,0,0,0)。而词“Woman”的词向量就是(0,0,0,1,0)。这种词向量的编码方式我们一般叫做1-of-N representation或者one hot representation.



但这样的向量其实除了一个位置是1，其余的位置全部都是0，表达的效率不高，因此产生了Distributed representation

##### Distributed representation

比如下图将词汇表里的词用“Royalty”, “Masculinity”, “Femininity”和“Age”4个维度来表示，King这个词对应的词向量可能是(0.99,0.99,0.05,0.7)(0.99,0.99,0.05,0.7)





## 19、信息压缩与特征提取

图片  卷积神经网络（CNN）

找不到提取语言特征的方法

循环神经网络（RNN）

大家好，这里是AIGC\_\_\_\_

## 20、信息压缩与特征提取

听到今天的直播课，小明非常喜欢，抑制不住对李老师的喜爱，伸手点了个\_\_

A，赞

B，菜

按照顺序去寻找，无法找到正确的词向量

## 21、谷歌论文-自注意力机制

---

### 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

Lukasz Kaiser\*  
Google Brain  
lukaszkaizer@google.com

Illia Polosukhin\* ‡  
illia.polosukhin@gmail.com

ChatGPT  
T: transformer

解决了自然语言特征提取的问题

## 22、解决问题的思路

听到今天的直播课，小明非常喜欢，抑制不住对李老师的喜爱，伸手点了个\_\_

A，赞

B，菜

让句子的词跟句子里所有的词都做向量计算



## 23、Transformer算法演进



BERT

GPT

## 24、通过人工智能模型

细分领域

围棋

玩游戏

图像识别

设计模型

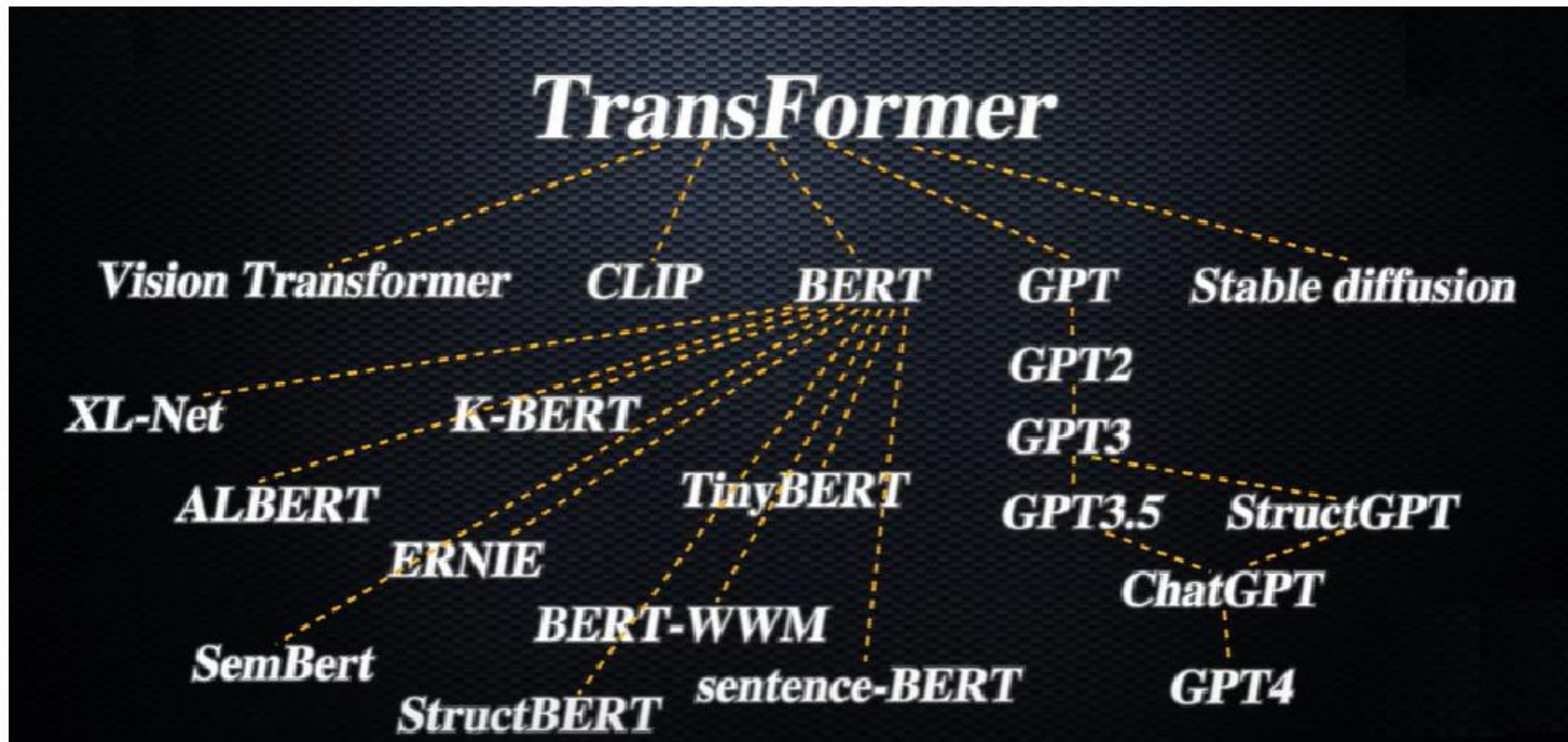
标注数据

自然语言

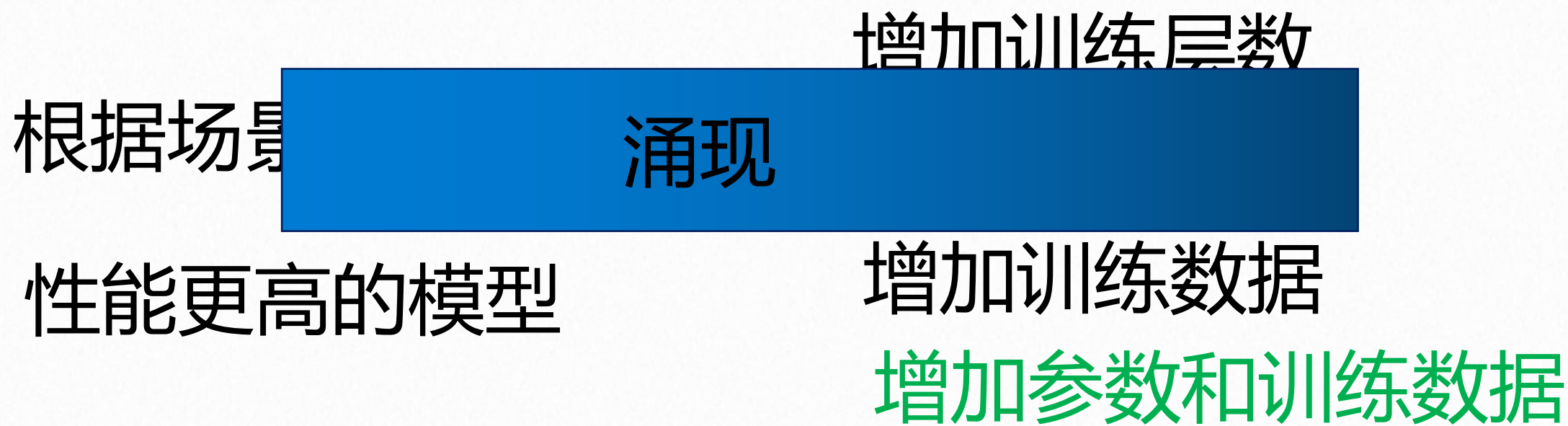
理解整个文明成果的能力，和人类无缝交流的能力



## 25、通用人工智能模型



## 26、通用人工智能模型





# 参数1000亿

## 语言模型具备了本来不应该具备的能力

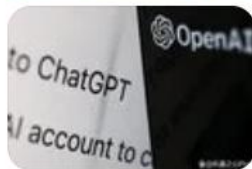
### 2个月,暴增1亿用户!全网刷屏的ChatGPT,有何魔力?



2023年2月13日 据瑞银集团报告,截至1月末,Open AI推出的生成式人工智能——ChatGPT月活用户已突破1亿,成为史上用户增长最快的应用程序。这是什么概念? 据World of Engineering整理,想要够到1亿月...

黑马公社

### 史上增速最快消费级应用,ChatGPT月活用户突破1亿



2023年2月3日 今日, 瑞士银行巨头瑞银集团的一份报告显示, 在 ChatGPT 推出仅两个月后, 它在 2023 年 1 月末的月活用户已经突破了 1 亿, 成为史上用户增长速度最快的消费级应用程序。根据 Sensor ...

机器之心Pro

### 上线两月用户破亿,出道即巅峰!ChatGPT在狂飙



2023年2月19日 讨论的热潮也帮助ChatGPT收获了不少用户。数据显示, 在去年11月30日发布后的5天时间, ChatGPT收获了100万用户; 推出仅仅2个月, 月活跃用户就成功过亿, 成为历史上增长最快的消费者...

湖南日报

# 关注视频号：玄姐谈AGI

## 助力数字化人才提升

### AIGC 能力



玄姐谈 AGI



扫一扫二维码，关注我的视频号