

证券研究报告  
产业研究  
2022年12月22日



# AIGC：内容生产力的革命 —Web3.0系列专题研究之一—

国海证券研究所

杨仁文(证券分析师)	马川琪(联系人)
S0350521120001	S0350121090021
yangrw@ghzq.com.cn	macq@ghzq.com.cn

## 一. 人工智能生成内容：技术进步+海量需求，驱动AIGC快速发展

- 1.1.1 AI规模：研发+数据+人力成本驱动中国AI市场高速发展
- 1.1.2 AI技术：细分赛道持续创新+变革在即，本报告重点关注AIGC领域
- 1.2.1 AIGC发展：需求+技术迭代促发展，从降本增效走向额外价值转移
- 1.2.2 AIGC技术：数据+算法+算力为核，NLP技术突破推动商业落地进程
- 1.2.3 AIGC算法：深度学习模型+开源模式为普及加速器，大模型为核心趋势
- 1.2.4 AIGC技术场景：个性&自动化为核心壁垒，关注渠道把控及商业路径
- 1.2.5 AIGC应用场景：集中在数字化程度高+内容需求丰富的行业

## 二. 人工智能文本生成：大数据+大算法+大算力，交互&个性化生成技术前景广阔

- 2.1.1 AI文本生成：底层架构不断精进，交互+个性化生成技术前景广阔
- 2.2.1 对话式AI发展：技术突破催化产品落地，从规则匹配走向深度学习
- 2.2.2 对话式AI技术：智能语音+NLP+知识工程底层技术变革下的新机遇
- 2.3.1 Chatbot全球视角：技术驱动增长，亚太地区或成最大市场
- 2.3.2 Chatbot中国视角：百舸争流，软件产品收费为主，细分赛道需求向好
- 2.4.1 ChatGPT：大数据+RLHF技术路径下的“类人交流”模型
- 2.4.2 ChatGPT：实验证明InstructGPT优于GPT-3模型，但仍有改进空间（1/2）
- 2.4.3 ChatGPT：实验证明InstructGPT优于GPT-3模型，但仍有改进空间（2/2）
- 2.4.4 ChatGPT：AIGC发展里程碑，超级基础设施下的大模型+大算力
- 2.4.5 ChatGPT：仍需调优及训练，进而实现商业化落地

## 三. 人工智能生成内容市场机会：中国AIGC如日方升，建议关注垂直赛道初创公司及下游AI应用场景

3.1.1 产业链：关注垂直赛道初创公司及下游AI应用场景

3.2.1 商业路径：中国企业底层技术发展不足，变现聚焦产出内容及软件服务

3.3.1 海外公司：2020年开始涌现出一批代表企业，集中文字/图像/音频领域

3.3.2 国内公司：建议关注垂直赛道初创企业

3.4.1 行业展望：关注技术/产品的迭代、核心场景的确定及产业接纳的态度

## 四. 附录 - ChatGPT试用：回答基础问题逻辑尚可，暂时难以摆脱知识整合和逻辑推理困境

4.1.1 测评背景及范围：聚焦功能验证、性能测试及用户体验

4.1.2 测评结论：语义理解较准，暂时难以摆脱知识整合和逻辑推理困境

4.2.1 用ChatGPT写篇AIGC报告（1/6）

4.2.2 用ChatGPT写篇AIGC报告（2/6）

4.2.3 用ChatGPT写篇AIGC报告（3/6）

4.2.4 用ChatGPT写篇AIGC报告（4/6）

4.2.5 用ChatGPT写篇AIGC报告（5/6）

4.2.6 用ChatGPT写篇AIGC报告（6/6）



- AI核心技术包括集成AI、通用人工智能、知识图谱、合成数据、自动驾驶及AIGC等；其中，AIGC技术场景包括文本生成、音频生成、图像生成、视频生成等；Chatbot为AI文字生成中应用较为广泛的领域。
- 2022年11月底，OpenAI上线智能对话系统ChatGPT，引爆全球。ChatGPT作为高效的AIGC文字内容生产工具，被认为将在智能客服、虚拟人、游戏等领域得到更大落地应用，也对很多算力、数据标注、自然语言处理等底层技术公司利好。正如OpenAI的CEO Sam Altman在推特上写道：“Trust the exponential. Flat looking backwards, vertical looking forwards.” **ChatGPT或许正让我们处在某个起飞的点上。**

如同蒸汽时代的蒸汽机、电气时代的发电机、信息时代的计算机和互联网，人工智能正成为推动人类进入智能时代的决定性力量。全球产业界充分认识到人工智能技术引领新一轮产业变革的重大意义，纷纷转型发展，抢滩布局人工智能创新生态。人工智能细分赛道持续创新下变革在即，本报告重点关注AIGC领域。

- **AIGC颠覆传统内容产出模式，或为web3.0内容创造新引擎。**AIGC本质上是一种AI赋能技术，能够通过其高通量、低门槛、高自由度的生成能力广泛服务于各类内容的相关场景及生产者。随着人工智能生成能力的突破进展，内容生产已经从专业生成内容（PGC）、用户生成内容（UGC），进入到人工智能生成内容（AIGC，AI generated content）时代，AIGC被认为是web3.0的重要基础设施。AIGC的快速兴起源于深度学习技术的快速突破和日益增长的数字内容供给需求；应用价值层面，AIGC有望成为数字内容创新发展新引擎，为数字经济发展注入新能量。**数据+算法+算力三大核心要素，决定AIGC产出质量。**①**数据**，海量优质的应用场景数据是训练算法精确性关键基础。②**算法**，神经网络、深度学习等算法是挖掘数据智能的有效方法。与传统机器深度机器学习算法不同，神经网络在学习范式+网络结构上的迭代提升了AI算法的学习能力，未来多模态大模型或为核心趋势，赋能产业空间及实践潜力。③**算力**，计算机、芯片等载体为AIGC提供基本的计算能力。
- **AIGC技术场景中，个性化及自动化内容产出为核心价值。**①**技术成熟度较高结构化领域大部分是在和人力生成内容进行竞争。**其中的存量价值来源于同类内容的降本增效，而增量价值则来源于跨模态的内容生成以及AI本身带来的科技感。对内容渠道的把控将成为核心竞争力。发行商、内容最终消费渠道具有强的产业链话语权。②**底层技术基本明确/仍待完善的原创性创作领域，本质为AI下的个性化数字内容的自动化构建。**该领域重点关注和其配套数据或底层原理是否清晰、商业化路径。目前AIGC整体影响仍十分有限，主要是中国市场供给端仍处于起步阶段。
- **AIGC应用场景中，数字化程度高及内容需求丰富的领域有广阔应用空间。**随着AIGC技术快速迭代，其可高效生成不同模态的信息产出（包括文字、音频、视频及跨模态），以真实性、多样性、可控性及综合性等特征，有望帮助企业提高内容生产的效率，以及为其提供更加丰富多元、动态且可交互的内容，或将率先在**传媒、电商、影视、娱乐**等数字化程度高、内容需求丰富的行业取得重大创新发展。
- **深度学习模型+开源模式加速AIGC普及，海外AIGC已到了“快速发展阶段”。**①随着深度学习模型不断迭代，人工智能生成内容百花齐放，产出效果逐渐逼真直至人类难以分辨。2018年，人工智能生成的画作在佳士得拍卖行以43.25万美元成交，成为首个出售的人工智能艺术品；2019年，DeepMind发布DVD-GAN模型用以生成连续视频；2022年11月，OpenAI上线了智能对话系统（聊天机器人）ChatGPT，引发全球热潮。ChatGPT的成功离不开参数竞赛时代下的“大模型”，显卡等硬件优化带来的“大算力基础”与基于“大数据”的RLHF训练模式。但由于训练数据的缺乏及训练数据的偏差，ChatGPT仍需要高成本的调优及持续训练，进而实现商业化落地。②“开源模式”加速AIGC产业发展。以深度学习模型CLIP为例，开源模式加速CLIP模型的广泛应用，使之成为当前最为先进的图像分类人工智能，并让更多机器学习从业人员将CLIP模型嫁接到其他AI应用。
- **中国AIGC仍处“萌芽期”，技术能力与产品形态的成熟、核心场景的确定及产业的接纳态度为行业关键发展节点。**据量子位预测，AIGC在中国发展可分为三个阶段：助手阶段（摸索磨合期，2021年~2026年）：AIGC辅助人类进行生产，优先变现的关键在于编辑优化功能，行业创新关键能力为素材模块分拆+个性化推荐；协作阶段（推广应用期，2026年~2028年）：人机共创，主要价值为降本增效及提供创意，预计互联网大厂将普遍布局，竞争热度提升；原创阶段（价值增长期，2028年之后）：AIGC将独立完成内容创作，产生附加价值。中国AIGC企业均在初创阶段，机会也许藏在垂直应用领域中，对赛道的选择十分关键。
- **风险提示：**政策监管风险；行业市场增长的不确定性；竞争环境不确定性；ChatGPT生成报告信息不准确性；AIGC企业商业化路径不确定性；技术创新不及预期。

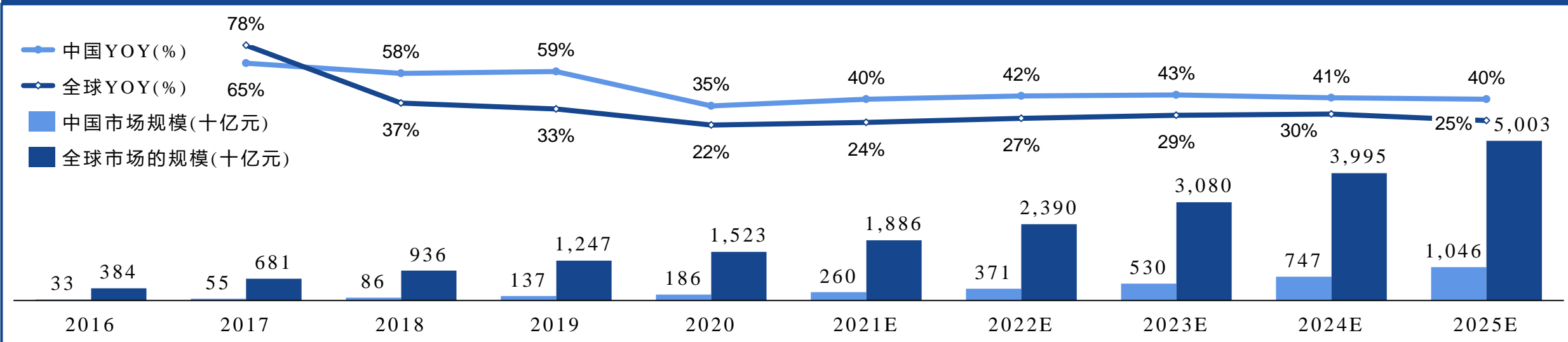
# 1. 人工智能生成内容

技术进步+海量需求，驱动AIGC快速发展

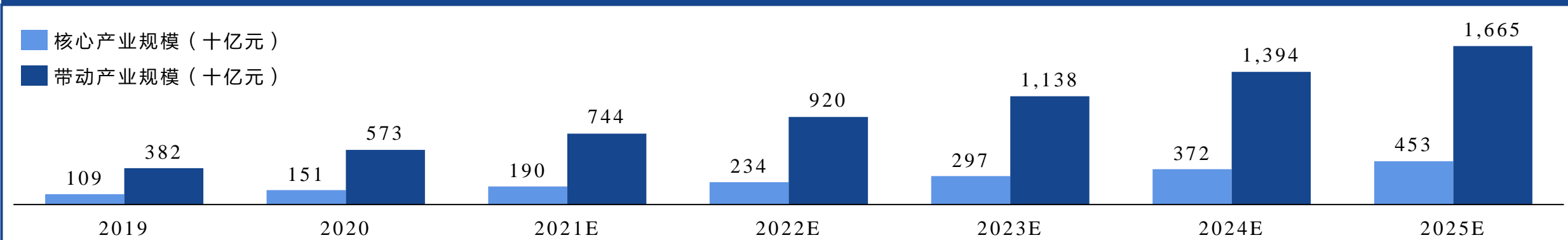


# 1.1.1 AI规模：研发+数据+人力成本驱动中国AI市场高速发展

## 人工智能（AI）全球&中国市场规模及增速：研发实力+海量数据+相对较低的人员成本将共同推动中国AI市场发展



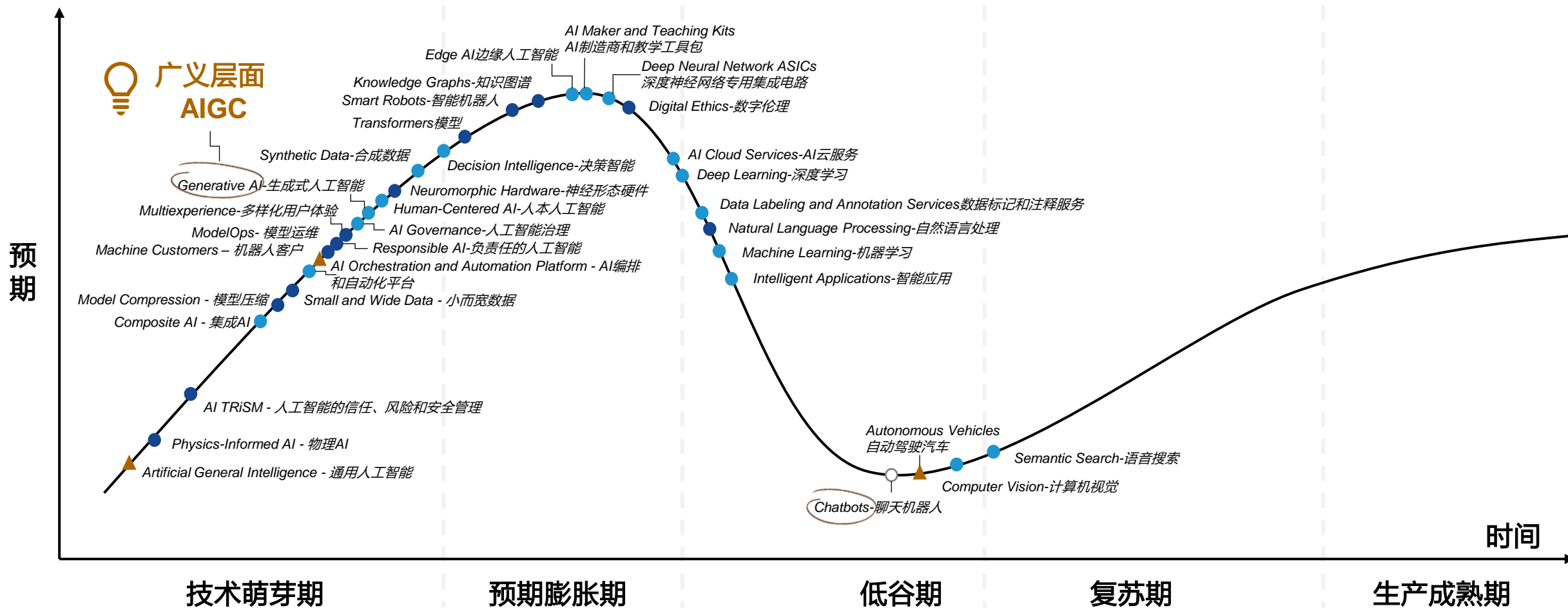
## 中国AI核心产业及带动产业规模及增长：AI赋能特性，带动相关产业发展，回馈社会经济



注：核心产业规模包括计算机视觉、智能语音、人机交互、机器学习、知识图谱与NLP、AI芯片等；带动产业规模包括AI核心产品所带动的工程服务、大数据平台与应用建设、计算机通信产品整机销售收入、甲方企业产值与效益提升的规模总和。

# 1.1.2 AI技术：细分赛道持续创新+变革在即，本报告重点关注AIGC领域

2021-人工智能（AI）技术成熟曲线：超过一半细分AI技术将为5年内成为主流技术



距稳定阶段需要：

○ 少于2年   ● 2~5年   ● 5~10年   ▲ 超过10年

注：AI为artificial intelligence缩写，即人工智能；AIGC全称为AI-Generated Content（人工智能生成内容）。



# 1.2.1 AIGC发展：需求+技术迭代促发展，从降本增效走向额外价值转移

## 定义

- AIGC全称为AI-Generated Content（人工智能生成内容），指基于生成对抗网络GAN、大型预训练模型等人工智能技术，**通过已有数据寻找规律，并通过适当的泛化能力生成相关内容的技术<sup>1</sup>**。AIGC既是从内容生产者视角进行分类的一类内容，又是一种内容生产方式，还是用于内容自动化生成的一类技术集合

## AIGC发展复盘



①**驱动因素**：AIGC的兴起源自深度学习技术的快速突破和日益增长的数字内容供给需求，市场潜力逐渐显现

- 一方面，技术进步驱动AIGC可用性不断增强；
- 另一方面，大量需求牵引AIGC应用落地。

②**发展预判**：AIGC在中国发展可分为三个阶段

- 助手阶段**（摸索磨合期，2021年~2026年）：AIGC辅助人类进行生产，**优先变现的关键在于编辑优化功能，行业创新关键能力为模块分拆+个性化推荐**；
- 协作阶段**（推广应用期，2026年~2028年）：人机共创，**主要价值为降本增效及提供创意，预计互联网大厂将普遍布局，竞争热度提升**；
- 原创阶段**（价值增长期，2028年之后）：**AIGC将独立完成内容创作，产生附加价值**。

注：1. Gartner提出了相似概念Generative AI，也即生成式AI。生成式AI是指该技术从现有数据中生成相似的原始数据。相较于量子位智库认为的AIGC，这一概念的范围较狭窄。据Harvard Business Review，生成AI能够生成文本和图像，涵盖博客文章、程序代码、诗歌和艺术作品，2017年谷歌Brain提出LLMs（large language model），此后领先的科技公司对于大型语言和文本-图像模型的应用激增

# 1.2.2 AIGC技术：数据+算法+算力为核，NLP技术突破推动商业落地进程

## AIGC核心要素

### 数据 - “饲料”

- 数据是AI算法的“饲料”，数据包括语音、文本、影像等
- 海量优质的应用场景数据是训练算法精确性关键基础

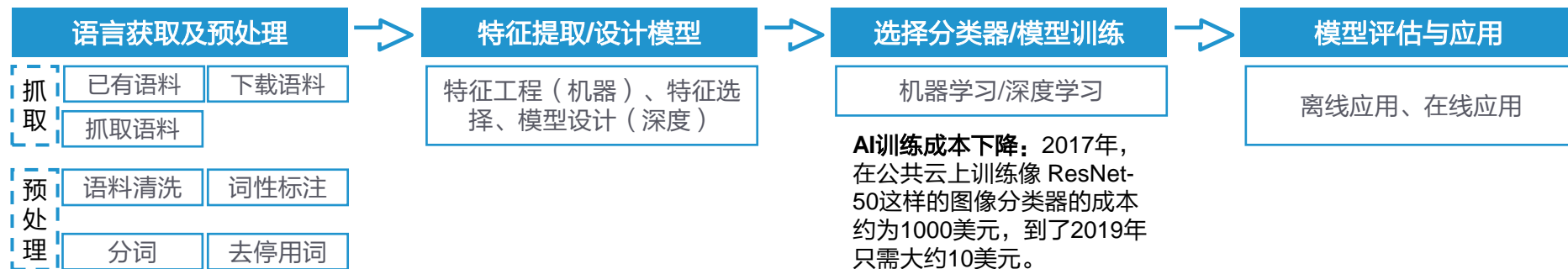
### 算法 - “推手”

- 算法是AI的背后“推手”
- 神经网络、深度学习等算法是挖掘数据智能的有效方法

### 算力 - “基础”

- 算力是基础设施，AI算法模型对于算力的巨大需求，推动了今天芯片业的发展
- 计算机、芯片等载体为人工智能提供基本的计算能力
- AI芯片包括通用/半定制化/定制化芯片

自然语言处理（NLP）是研究人与计算机交互中的语义理解问题的技术，解决的是人机交互中“听得懂”的问题。是在机器语言和人类语言之间沟通的桥梁，以实现人机交流的目的。以下为NLP应用流程：



## NLP - 自然语言处理

### NLP核心任务

#### 1 NLU - 自然语言理解

#### 2 NLG - 自然语言生成

## NLG的6个步骤

### 内容确定

- NLG系统需要决定哪些信息应该包含在正在构建的文本中，哪些不应该包含。通常数据中包含的信息比最终传达的信息要多。

### 文本结构

- 确定需要传达哪些信息后，NLG系统需要合理的组织文本的顺序。

### 句子聚合

- 将多个信息合并到一个句子里表达可能会更加流畅，也更易于阅读。

### 语法化

- 确定每句内容后，将这些信息组织成自然语言。

### 参考表达式生成

- 跟语法化很相似，不过与语法化的本质区别在于参考表达式生成需要识别出内容的领域，然后使用该领域（而不是其他领域）的词汇。

### 语言实现

- 当所有相关的单词和短语都已经确定时，需要将它们组合起来形成一个结构良好的完整句子。

欢迎访问：电子书学习和下载网站 (<https://www.shgis.com>)

人工智能行业Web3.0系列专题研究之：AIGC内容生产力的革命.pdf

请登录 <https://shgis.com/post/1379.html> 下载完整文档。

手机端请扫码查看：

