

# 工业大模型应用报告

#M

2024年3月



# 目录

<b>1. 大模型为工业智能化发展带来新机遇</b>	<b>1</b>
1.1. 大模型开启人工智能应用新时代	1
1.2. 大模型有望成为驱动工业智能化的引擎	3
1.3. 大模型应用落地需要深度适配工业场景	4
<b>2. 大模型和小模型在工业领域将长期并存且分别呈现 U型和倒U 型分布态势</b>	<b>6</b>
2.1. 以判别式 AI为主的小模型应用呈现倒U 型分布	6
2.2. 以生成式 AI为主的大模型应用呈现U 型分布	7
2.3. 大模型与小模型将长期共存并相互融合	9
<b>3. 工业大模型应用的三种构建模式</b>	<b>11</b>
3.1. 模式一：预训练工业大模型	11
3.2. 模式二：微调	12
3.3. 模式三：检索增强生成	13
3.4. 三种模式综合应用推动工业大模型落地	14
<b>4. 大模型应用探索覆盖工业全链条</b>	<b>16</b>
4.1. 大模型通过优化设计过程提高研发效率	16
4.2. 大模型拓展生产制造智能化应用的边界	19
4.3. 大模型基于助手模式提升经营管理水平	23
4.4. 大模型基于交互能力推动产品和服务智能化	25
<b>5. 工业大模型的挑战与展望</b>	<b>28</b>
5.1. 工业大模型应用面临数据质量和安全、可靠性、成本三大挑战	28
5.2. 工业大模型应用将伴随技术演进持续加速和深化	30

## 1. 大模型为工业智能化发展带来新机遇

### 1.1. 大模型开启人工智能应用新时代

**大模型引领人工智能技术创新和应用。**自1956年达特茅斯会议 (Dartmouth Conference) 提出人工智能的概念以来, 人工智能技术经历了多个发展高峰和低谷。在这一长期的发展过程中, 人工智能技术不断演进, 逐步朝着更高的智能水平和适应性发展。2022年11月30日, OpenAI 发布了ChatGPT, 引发了行业热潮, 直至今日, 业界普遍认为, 大模型时代已经到来, 也象征着人工智能开启了迈向通用人工智能 (AGI, Artificial General Intelligence) 的新阶段。在大模型出现之前, 人工智能通常需要针对特定的任务和场景设计专门的算法, 这种方法虽然在特定领域有效, 但人们对“智能”的期望是能够适应多种任务和场景的智能系统, 单一任务的人工智能系统已经无法满足这些更广泛的需求。大模型能够跨越传统人工智能的局限性, 理解和推理的能力有了极大的飞跃, 同时也提高了复用的效率, 为人工智能技术在更多领域的应用提供了坚实的基础, 推动人类社会迈向通用人工智能 (AGI) 的新阶段。

**通用性和复用性是大模型的关键价值。**2017年, Google Brain (谷歌大脑) 团队在其论文《Attention Is All You Need》中创造性地提出 Transformer 架构, 凭借注意力机制, 极大地改善了机器学习模型处理序列数据的能力, 尤其是在自然语言处理 (NLP) 领域。Transformer 架构的出现, 为后续的大模型如 ChatGPT 等奠定了技术基础。ChatGPT、Bert 等大模型通过海量数据和庞大的计算资源支持, 使得模型具备了强大的通用性和复用性。大模型可以被广泛应用于自然语言处理、计算机视觉、语音识别等领域的各种任务, 能够为各种应用和开发人员提供共享的基础架构, 并进一步通过微调、零样本学习等方式, 直接在一系列下游任务上使用, 取得一定的性能表现, 支持不同行业、不同场景的应用构建。

**大模型展现出三大基础特征。** 目前大模型并没有明确的定义，狭义上指大语言模型，广义上则指包含了语言、声音、图像等多模态大模型。如李飞飞等人工智能学者所指出，这些模型也可以被称为基础模型 (Foundation Model)。我们认为，大模型主要具备以下三大特征：

**参数规模大：** 大模型的参数规模远大于传统深度学习模型。大模型发展呈现“规模定律” (Scaling Law) 特征，即：模型的性能与模型的规模、数据集大小和训练用的计算量之间存在幂律关系，通俗而言就是“大力出奇迹”。不过“大”并没有一个绝对的标准，而是一个相对概念。传统模型参数量通常在数万至数亿之间，大模型的参数量则至少在亿级以上，并已发展到过万亿级的规模。如 OpenAI 的 GPT-1 到 GPT-3，参数量从1.1亿大幅拉升到1750亿，GPT-4 非官方估计约达1.8万亿。

**泛化能力强：** 大模型能够有效处理多种未见过的数据或新任务。基于注意力机制 (Attention)，通过在大规模、多样化的无标注数据集上进行预训练，大模型能够学习掌握丰富的通用知识和方法，从而在广泛的场景和任务中使用，例如文本生成、自然语言理解、翻译、数学推导、逻辑推理和多轮对话等。大模型不需要、或者仅需少量特定任务的数据样本，即可显著提高在新任务上的表现能力。如 Open AI 曾用 GPT-4 参加了多种人类基准考试，结果显示其在多项考试中成绩都超过了大部分人类 (80% 以上)，包括法学、经济学、历史、数学、阅读和写作等。

**支持多模态：** 大模型可以实现多种模态数据的高效处理。传统深度学习模型大多只能处理单一数据类型 (文本、语音或图像)，大模型则可以通过扩展编/解码器、交叉注意力 (Cross-Attention)、迁移学习 (Transfer learning) 等方式，实现跨模态数据的关联理解、检索和生成。多模态大模型 (LMMs, Large Multimodal Models) 能够提供更加全面的认知能力和丰富的交互体验，拓宽 AI 处理复杂任务的应用范围，成为业



界探索迈向通用人工智能的重要路径之一。典型如 OpenAI 的 Sora模型推出，掀起了全球多模态大模型的发展新热潮。

## 1.2. 大模型有望成为驱动工业智能化的引擎

**人工智能推动工业智能化发展进入新阶段。**工业发展是一个逐步演进的过程，经历了机械化、电气化、自动化、信息化的阶段后，当前正处于从数字化向智能化迈进的阶段。每个阶段都是工业与各类创新技术的融合，对传统制造业进行升级和改造，提高生产效率、降低成本、提升产品质量。当前阶段，工业领域积累了大量的数据、基础能力和场景需求，为工业场景与人工智能技术的融合提供了基础条件。而人工智能逐渐展现出类似人的理解和分析能力，这些能力与工业场景的融合，将智能化带入到工业生产、运营、管理等领域，不断提升感知、认知和决策等多个环节，有望推动工业发展走向“自适应、自决策、自执行”的智能化阶段。

**大模型为工业智能化提供新动力。**尽管人工智能在智能制造、工业4.0、工业互联网等方面有所应用，但这些应用往往受限于特定任务，难以实现跨领域、跨场景的融合创新。过去，人工智能在工业的应用主要聚焦于如质量检测、预测性维护等单一功能，这形成了人工智能应用上“一场景一训练一模型”的局限，尚未形成大规模的应用。然而，大模型的崛起有望带来“基础模型+各类应用”的新范式。大模型凭借其卓越的理解能力、生成能力和泛化能力，能够深度洞察工业领域的复杂问题，不仅可以理解并处理海量的数据，还能从中挖掘出隐藏在数据背后的规律和趋势。此外，区别于传统的人工智能模型只能根据已有数据进行预测和推断，大模型则能够生成新的知识和见解。最后，大模型的泛化能力能够在更广泛的工业场景发挥作用。

**大模型为工业智能化拓展新空间。**人工智能在工业领域的应用，尽管已经取得了一些显著的成果，但整体来看，其应用的普及率仍然处于相对较低的水平。据凯捷

(Capgemini) 统计数据显示,即便是欧洲顶级的制造企业,其AI应用的普及率也仅超过30%,而日本和美国制造企业的AI应用率分别达到了30%和28%。相较于这些发达国家,中国制造企业AI普及率尚不足11%,显示出这一领域巨大的发展潜力和广阔的空间。相较于以往的小模型,大模型有望挖掘工业领域人工智能应用的新场景,提升人工智能应用的普及率。例如在研发设计领域,大模型能够深度挖掘和分析海量数据,为产品设计提供更为精准和创新的思路。在经营管理领域,大模型能够实现对生产流程、供应链管理等各个环节的监控和智能优化,从而提升企业的运营效率和市场竞争力。

### 1.3. 大模型应用落地需要深度适配工业场景

大模型的优势在于其强大的泛化能力,可以在不同的领域和任务上进行迁移学习,而无需重新训练。但无法充分捕捉到某个行业或领域的特征和规律,也无法满足某些特定的应用场景和需求,在真正融入行业的过程中,需要适配不同的工业场景,其核心就是要解决以下三个问题。

**不懂行业:** 大模型在处理特定行业任务时,往往表现出对行业知识、术语、规则等的理解,导致生成的解决方案与实际需求存在偏差,这主要是由于大模型在训练过程中缺乏特定行业的数据和知识,难以覆盖各个行业的专业细节。这种行业知识的匮乏使得大模型在应对工艺流程优化、设备故障预测等专业问题时有所缺陷,难以提供精确、可靠的解决方案,无法满足工业现场的个性化要求。

**不熟企业:** 当大模型接入企业系统时,往往难以全面理解企业的业务流程、数据结构和运营模式,导致生成的解决方案与企业实际需求不匹配。每个企业都有其独特的运营环境和业务需求,而大模型往往缺乏对企业特定环境的深入理解。此外,企业内部的数据孤岛和碎片化的信息系统也增加了大模型理解企业环境的难度。这种不熟

企业的问题使得大模型难以真正融入企业的运营流程，无法平滑地嵌入到现有的各类系统中。

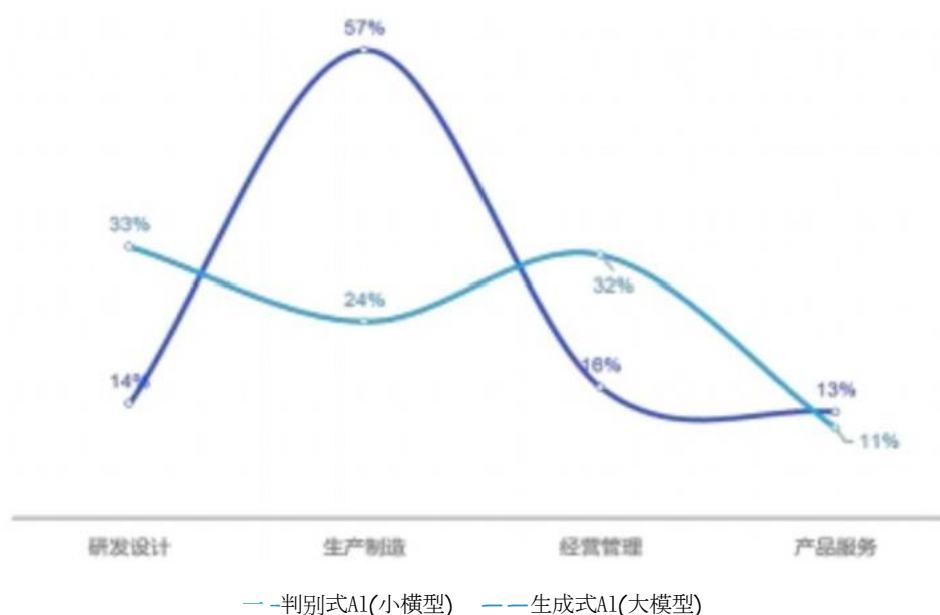
**存在幻觉：** 在某些情况下，大模型会产生与实际情况不符的错误输出，即“幻觉”现象。这主要是由于模型在训练过程中受到了噪声数据、偏差样本等因素的影响，或者由于模型的复杂性和数据维度过高导致过拟合。这种幻觉现象对工业领域的影响是全方位的，无论是生产过程中的质量控制、设备维护，还是供应链管理、市场预测等环节，错误的输出都可能导致严重的决策失误和经济损失。特别是在对安全性、可靠性要求极高的工业场景中，如航空航天、核能等领域，幻觉现象可能带来灾难性的后果。



## 2. 大模型和小模型在工业领域将长期并存且分别呈现U型和倒U型分布态势

从工业智能化的发展历程可以看出，在大模型出现之前，人工智能技术在工业领域已有较多应用。在前期阶段，工业人工智能的应用主要是以专用的小模型为主，而大模型开启了工业智能化的新阶段。结合两者不同的技术特点和应用能力，目前在工业领域形成了不同的分布态势。

图表1生成式AI(大模型)和判别式AI(小模型)在工业主要领域分布情况<sup>1</sup>



### 2.1. 以判别式 AI 为主的小模型应用呈现倒U 型分布

根据中国信通院<sup>2</sup>对507个AI小模型应用案例的统计分析，这些应用主要集中在生产制造领域，占比高达57%，而在研发设计和经营管理领域的应用则相对较少。这种分布呈现出明显的倒U型。

小模型的核心特点是学习输入与输出之间的关系。小模型通过学习数据中的条件概率分布，即一个样本归属于特定类别的概率，再对新的场景进行判断、分析和预测。它的优点是通常比大模型训练速度更快，而且可以产生更准确的预测结果，尤其适用

<sup>1</sup> 1507个小模型应用数据引用自中国信息通信研究院《工业智能白皮书(2022)》，99个大模型应用数据由本文编写组收集、整理、统计分析所得

<sup>2</sup> 中国信息通信研究院《工业智能白皮书(2022)》





于对特定任务进行快速优化和部署的场景。以工业质检领域为例，小模型能够从海量的工业产品图片数据中，学习到产品的外观特征、质量标准和缺陷模式等关键信息。当面对新的样本时，小模型能够迅速判断样本是否合格，从而实现对产品质量的快速检测。同样在设备预测性维护方面，小模型通过对设备运行数据的分析，能够学习到设备正常运行的模式和潜在的故障特征。一旦监测到异常情况，小模型能够及时发出预警，提醒工作人员进行检修或维护。

**小模型的能力更适合工业生产制造领域。**首先，小模型能够基于有限数据支撑精准的判别和决策，而生产过程需要针对不同场景进行精准的分析 and 决策，这两者间的契合使得小模型在生产制造领域具有独特的优势。其次，生产制造过程对准确性和稳定性有着极高的要求，任何微小的误差都可能导致产品质量下降或生产线停工。小模型在训练过程中，能够针对具体场景进行精细化的调整和优化，从而确保模型的准确性和稳定性，这使得小模型在生产制造领域的应用更为可靠和有效。最后，小模型在成本投入方面相对较低，使得其在生产制造现场的应用更具经济性，并在有限的硬件条件下，能够稳定运行。

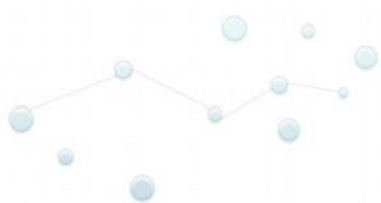
**小模型的定制化需求制约了其进一步渗透。**尽管小模型在生产制造领域表现出色，但其应用过程中也面临着一些挑战。以判别式 AI 为代表的小模型通常需要依靠个性化的业务逻辑进行数据采集、模型训练与调优，往往只能处理单一维度的数据。这一过程不仅消耗大量的算力和人力，而且训练后的模型往往无法在多行业通用。例如，工业缺陷检测领域的视觉模型往往需要针对一个产品或者一个设备训练一个模型，产品或设备不同，就要对模型进行重新训练，这种定制化的需求在一定程度上制约了小模型在工业领域的进一步渗透。

## 2.2. 以生成式AI 为主的大模型应用呈现U 型分布

根据对99个工业大模型应用案例的统计分析，大模型在研发设计和经营管理领域的应用相对更多，整体上呈现出 U 型分布。这表明大模型当前的能力更适配于研发设计和经营管理，在生产制造环节的能力和性能还需进一步提升。

**大模型通过构建庞大的参数体系来深入理解现实世界的复杂关系。**大模型的核心在于学习数据中的联合概率分布，即多个变量组成的向量在数据集中出现的概率分布，进而通过使用深度学习和强化学习等技术，能够生成全新的、富有创意的内容。与传统的数据处理方法不同，大模型并不简单地区分自变量与因变量，相反，它致力于在庞大的知识数据库中提炼出更多的特征变量。这些特征变量不仅数量庞大，而且涵盖了多个维度和层面，从而更全面地反映现实世界的复杂关系。以自然语言处理为例，大模型通过学习大量的文本数据，能够掌握语言的规律和模式。当给定一个句子或段落时，大模型能够基于联合概率分布生成与之相关的新句子或段落。这些生成的内容不仅符合语法规则，而且能够保持语义上的连贯性和一致性。此外，大模型还能够根据上下文信息理解并回答复杂的问题，展现出强大的推理和创造能力。

**大模型更适合综合型和创造类的工业场景。**在综合型工业场景中，由于涉及到多个系统、多个流程的协同工作，需要处理文档、表格、图片等多类数据，变量之间的关系往往错综复杂，难以用传统的分析方法进行精确描述。大模型通过深度学习和复杂的网络结构，可以捕捉并模拟这些关系，从而实现对复杂系统的全面理解和优化。在创造类工业场景中，大模型的优势体现在其强大的内容生成能力上。例如，在产品外观设计方面，传统的设计方法往往依赖于设计师的经验和灵感，设计周期长且难以保证设计的创新性。而大模型通过学习大量的设计数据，能够掌握设计领域的规律和模式，进而生成符合要求的全新设计内容，提升产品设计的效率和质量。



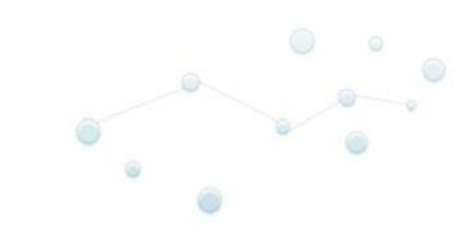
**大模型在工业领域的应用潜力仍有待释放。** 首先，大模型技术本身正处于快速发展的阶段，尽管已取得了显著进步，但在成本、效率和可靠性等方面仍有待进一步提升，以适应工业领域日益复杂的需求。其次，工业场景众多且各具特色，大模型作为新技术，需要逐步与各个工业场景紧密结合，在逐步提升技术渗透率的过程中，挖掘可利用的场景，并根据行业特定需求提供定制化的解决方案。最后，工业领域自身的数据分散且缺少高质量的工业数据集，同时在实际生产中如何确保工业数据的隐私和安全也是企业关注的重点，这些现实问题也限制了大模型的推广应用。

### 2.3. 大模型与小模型将长期共存并相互融合

**目前大模型在工业领域还未呈现出对小模型的替代趋势。** 尽管以生成式 AI 为代表的大模型被视为当前 AI 的热点，但在工业领域的实际应用中，大模型的能力和成本问题导致其尚不能完全取代以判别式 AI 为代表的小模型。一方面，小模型在工业领域具有深厚的应用基础和经验积累，其算法和模型结构相对简单，易于理解和实现，其稳定性和可靠性得到了验证。另一方面，大模型在成本收益比、稳定性和可靠性等方面存在问题，其在工业领域的探索还处在初级阶段。小模型以其高效、灵活的特点，在特定场景和资源受限的环境中发挥着重要作用；而大模型则以其强大的泛化能力和处理复杂任务的优势，在更广泛的领域展现着巨大潜力，两者将长期共存。

**大模型与小模型有望融合推动工业智能化发展。** 对于小模型而言，利用大模型的生成能力可以助力小模型的训练。小模型训练需要大量的标注数据，但现实工业生产过程可能缺少相关场景的数据，大模型凭借强大的生成能力，可以生成丰富多样的数据、图像等。例如，在质检环节，大模型可以生成各种可能的产品缺陷图片，为小模型提供丰富的训练样本，从而使其能够更准确地识别缺陷和异常。此外，大模型可以通过 Agent 等方式调用小模型，以实现灵活性与效率的结合。例如，在某些场景下，大

模型可以负责全局的调度和决策，而小模型可以负责具体的执行和控制。这样既能保证系统的整体性能，又能提高响应速度和灵活性。






### 3. 工业大模型应用的三种构建模式

大模型的构建可以分为两个关键阶段，一个是预训练阶段，一个是微调阶段。预训练主要基于大量无标注的数据进行训练，微调是指已经预训练好的模型基础上，使用特定的数据集进行进一步的训练，以使模型适应特定任务或领域。针对工业大模型，一是可以基于大量工业数据和通用数据打造预训练工业大模型，支持各类应用的开发。二是可以在基础大模型上通过工业数据进行微调，适配特定工业任务。三是可以在不改变模型参数的情况下，通过检索增强生成（RAG）为大模型提供额外的数据，支持工业知识的获取和生成。这三种模式并不独立应用，往往会共同发力。

图表2工业大模型应用的三种构建模式对比

	预训练工业大模型	微调	检索增强生成
 数据需求	无标注及标注的工业数据，静态数据	标注的工业数据为主，静态数据	外挂行业数据源，动态数据
特点	具备组分工业领域的理解能力	适用于工业领域的特定任务	不改变模型快速接入行业信息
优点	对工业通用知识理解	精准执行工业特定任务	快速利用外部信息资源，减少幻觉
缺点	成本较高，缺乏对特定任务的优先能力	泛化能力较弱，可能过拟合	不同模型对行业的深度理解能力
适用场景	作为基础模型支撑多种工业应用的开发	针对高标注量的工业数据实现特定任务	快速结合数据源进行信息检索和输出

#### 3.1. 模式一：预训练工业大模型

无监督预训练主要利用大量无标注数据来训练模型，目的是学习数据的通用特征和知识，包括 GPT-3/GPT-4、LLaMA1/LLaMA2 等，都是通过收集大量无标注的通用数据集，使用 Transformer 等架构进行预训练得到。预训练之后的模型已经足够强大，能够使用在广泛的领域。例如，当无监督预训练技术应用于 NLP 领域时，经过良好训练的语言模型可以捕捉到对下游任务有益的丰富知识，如长期依赖关系、层次关系等。然而，另一方面完全基于互联网等通用数据训练的大模型缺乏对行业知识的理解，在应对行业问题上表现出的性能较差，因此在预训练阶段可以使用通用数据加行业数据进行模型训练，使得在基础模型的层面就具备了一定的行业专有知识。





无监督预训练工业大模型的优点是可以具备广泛的工业通用知识，最大程度地满足工业场景的需求，实现模型的最优性能与稳定性。但这一模式的缺点是需要大量的高质量工业数据集，以及庞大的算力资源，对成本和能力的要求较高，面临技术和资源的巨大挑战。在最终应用前，无监督预训练工业大模型与GPT3 类似，同样需要通过适当的指令微调、奖励学习、强化学习等阶段，形成面向最终场景的应用能力。

SymphonyAI 推出了基于无监督预训练的工业大语言模型，该模型的训练数据包含3万亿个数据点，12亿token, 能够支持机器状况诊断，并回答故障状况、测试程序、维护程序、制造工艺和工业标准相关的问题。

制造流程管理平台提供商 'Retrocausal' 发布的LeanGPTTM,也采用了无监督预训练的模式，是制造领域的专有基础模型。基于 LeanGPTTM这一基础模型，Retrocausal 还推出了 Kaizen CopilotTM的应用程序，可以帮助工业工程师设计和持续改进制造装配流程。

### 3.2 模式二：微调

微调模式是在一个已经预训练完成的通用或专业大模型基础上，结合工业领域特定的标注数据集进行进一步的调整和优化，从而使模型能够适应具体的工业场景需求，更好地完成工业领域的特定任务。在微调期间，需要使用特定任务或领域量身定制的标记数据集来训练，与模型预训练所需的巨大数据集相比，微调数据集更小，单个任务的微调通常只需要几千条到上万条有标注数据即可。通过微调，大模型可以学习到工业细分领域的知识、语言模式等，有助于大模型在工业的特定任务上取得更好的性能。在当前主流的行业大模型构建路线中，众多行业模型都是使用基础大模型+行业标注数据集来微调得到的。

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。  
如要下载或阅读全文，请访问：<https://d.book118.com/116241022124010144>