

工业大模型技术应用与发展报告1.0

2023.12

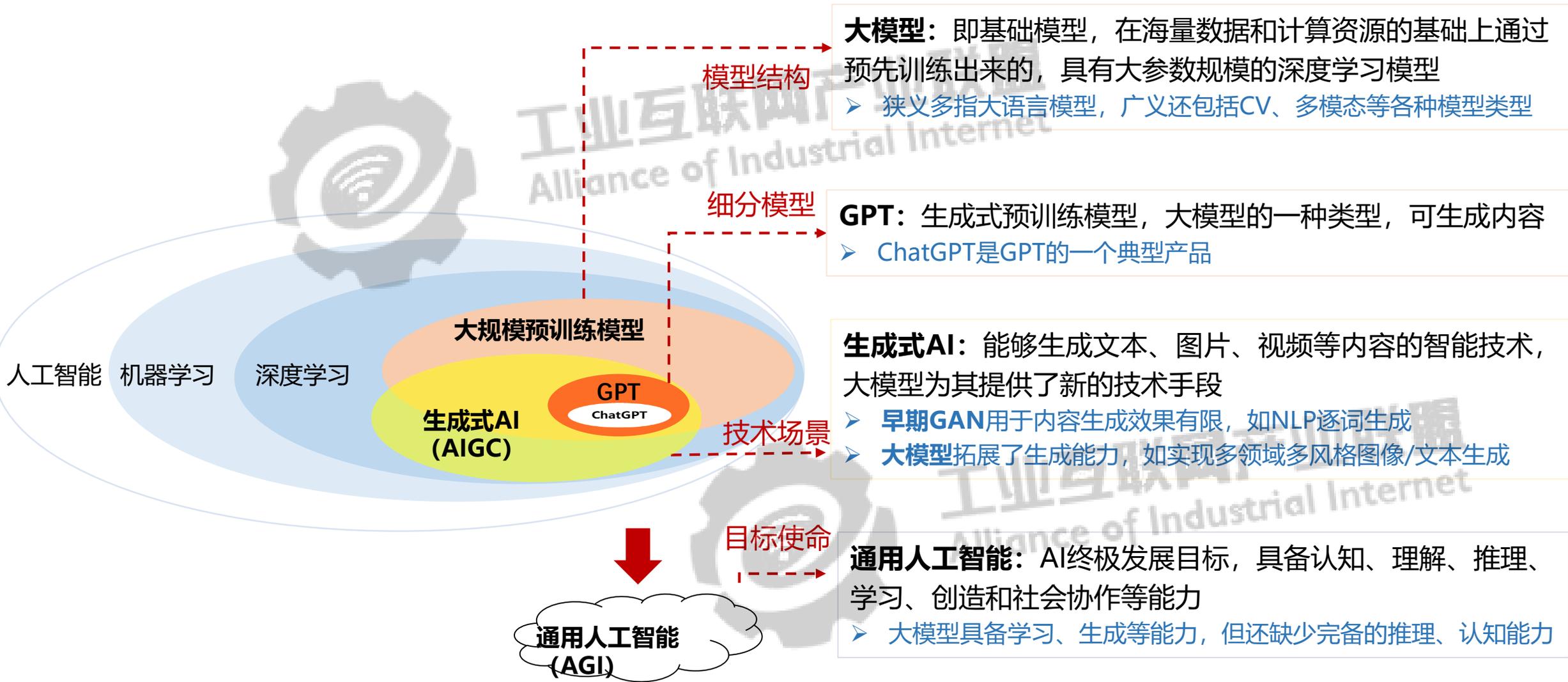
编写单位（排名不分先后）

牵头编写单位：中国信息通信研究院

参与编写单位： 百度在线网络技术（北京）有限公司
羚羊工业互联网股份有限公司
创新奇智科技集团股份有限公司
智昌科技集团股份有限公司
中国科学院自动化研究所
中工互联北京科技集团有限公司
西门子股份公司
航天云网科技发展有限责任公司
威派格智慧水务股份有限公司
卡奥斯COSMOPlat
阿里云计算有限公司
腾讯计算机系统有限公司

中科云谷科技有限公司
浪潮集团有限公司
中科视语科技有限公司
苏州海赛人工智能有限公司
北京通用AI研究院
树根互联股份有限公司
华为技术有限公司
IBM（国际商业机器（中国）有限公司）
美云智数科技有限公司
中科斯欧(合肥)科技股份有限公司

人工智能的几个相关概念



1.1 大模型成为通用AI新范式，引发工业应用变革讨论

大模型+大数据+大算力成为
发展主旋律 (ChatGPT)

AI产品渗透速度全球排名第2

引发产业领域应用的热烈讨论与憧憬

千亿参数基础模型

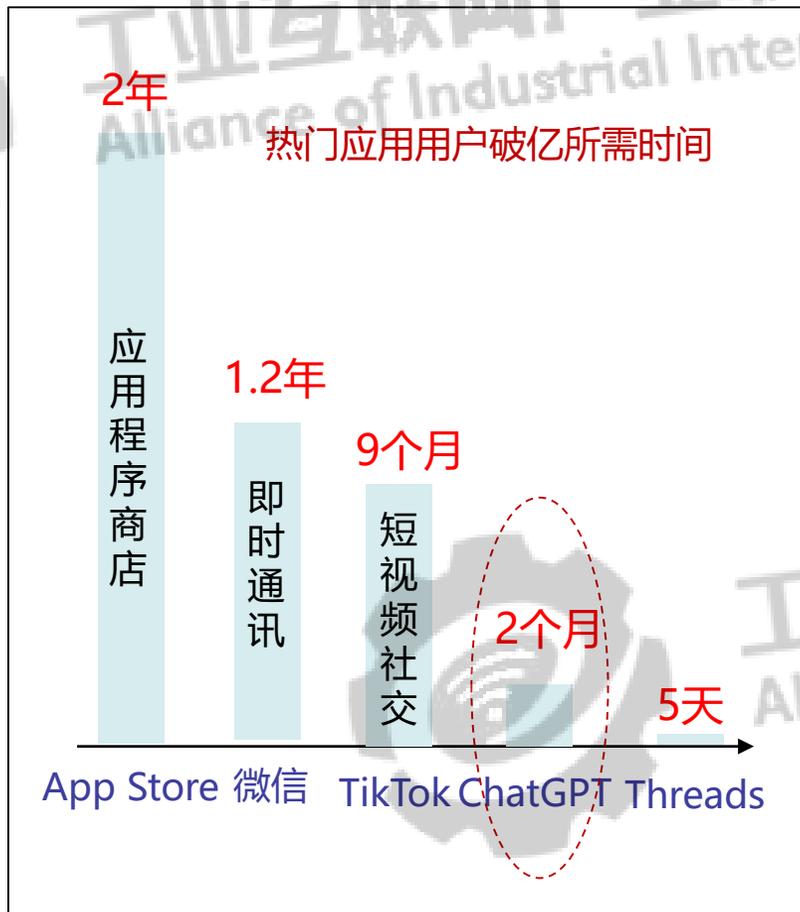
GPT-1 : 1.17亿
GPT-2:15亿
GPT-3:1750亿
GPT=4: 1.48万亿.....

TB级数据

超3000亿单词
830GB代码数据
1000个外包团队标注

E级智能超算总算力

超195PFlops我国最
快超算“神威·太湖之
光” 1.7倍



- 产业应用场景成为大模型最佳“练兵场”——科技日报
 - AI大模型落地背后，正带来一场智能制造的系统重构——百度
 - 工业大模型将会带来一场新的工业革命，它将来会成为工业领域的基础设施——中工互联
 - 制造业是AI大模型的重要战场，未来10年最大的机会——阿里
- ...

1.2 工业大模型与专用小模型成为工业AI发展的两条协同路径

工业大模型=工业+大模型

1 满足大模型技术基本特征

- Transformer为基础框架
- 在大量通用数据上进行预先训练，以实现良好的通用性
- 模型参数一般达十亿以上（最大模型参数已达万亿级）

2 具备在工业各环节进行应用的能力，或与工业装备软件等融合赋能

工业大模型
Transformer结构

工业专用小模型
传统模型结构

应用层面

- **新场景**：代码生成、CAD生成等新应用...
- **泛化性强**：单模型应对多任务，更适合长尾落地
- **特定任务**：点状场景应用效果更好

工程层面

- **低成本运维**：低成本开发+维护
- **模型更新快**：参数量少，可进行快速迭代
- **轻量化部署**：所需存储空间和算力更小

1.3 大模型初步形成赋能工业的核心方式与产品形态

3类主要赋能方式

当前(可用于)工业领域大模型超30个

4类模型产品形态

基于通用底座直接赋能行业

- OpenAI ChatGPT
- SIEMENS PLC编程
- 用友 YonGPT
- 科大讯飞 星火大模型
- Paradigm 式说大模型

大模型API调用或软件方案

- 基于ChatGPT直接开展智能客服等应用

基于通用底座进行场景化适配调优或形成外挂插件工具

- HUAWEI 制造、矿山等行业大模型
- Baidu 百度 航天、能源等行业大模型
- salesforce Einstein GPT
- AUTHENTISE 3D打印GPT

成熟工业产品叠加基础模型能力

- 倍福将大模型融入 TwinCAT XAE客户端，实现基于对话辅助编程

外挂插件工具

- 工业管理软件企业Authentise推出插件，用户可查询最大的增材制造知识库
- 浙大开发用于表格处理的TableGPT

面向工业或具体任务针对性开发

- Google AlphaFold2
- Meta ESMFold
- DPTechnology 深势科技 Uni-Mol
- HUAWEI 科学大模型

用于私有化部署的一体机

- 科大讯飞推出星火一体机

2.1 大模型赋能工业领域的适用边界与核心能力

适用问题：大模型并非万金油

● 大场景

- 工业场景具备一定通用性
- 涉及关联复杂的智能任务
- 数据边界对决策效果有直接影响

● 大语料

- 工业场景的基础数据/语料/规则约束充足

● 问题边界清晰

- 结果存在于封闭信息环境，不依赖语料外的信息

核心能力

工业领域应用变革

✓ 语言理解

预置型对话 → 与设备/工业系统的自然交互与推理

✓ 生成创作

规则式生成 → 工业代码/图文内容的“涌现式”生成

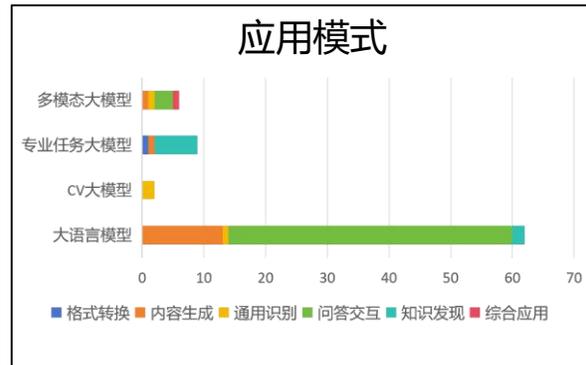
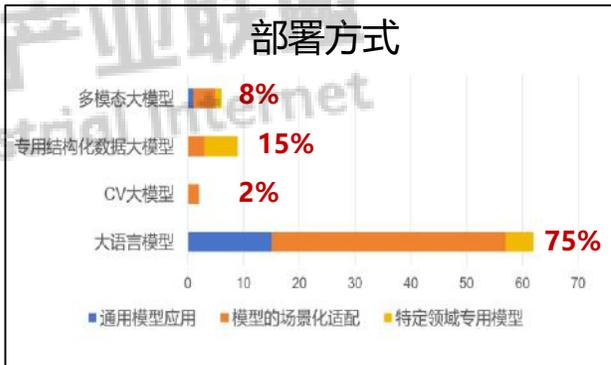
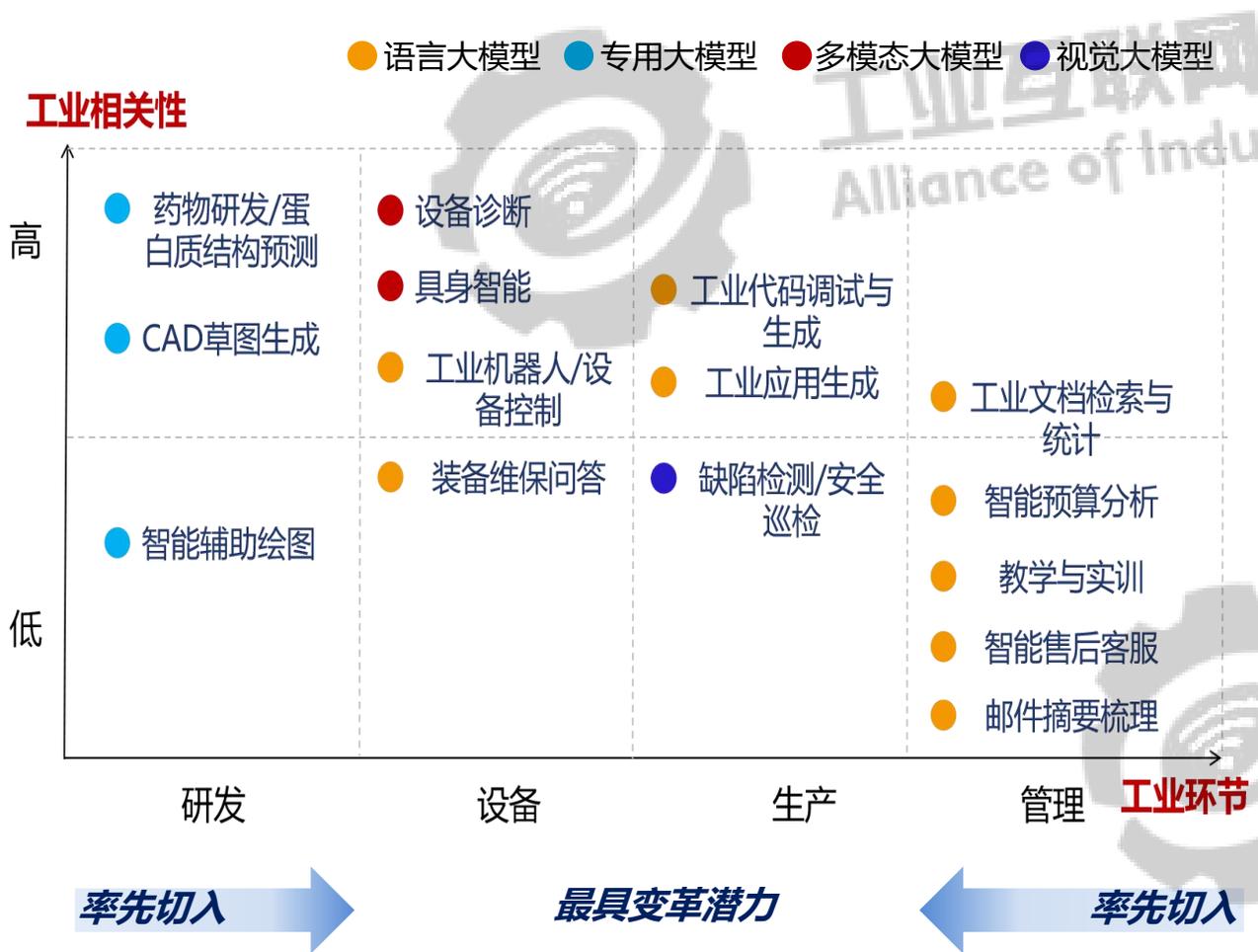
✓ 识别/模拟/预测

局部建模预测 → 基于全局信息高效高精度预测优化

✓ 多模态

单一格式工业数据处理 → 多格式数据综合转换分析

2.2 应用总体视图：4类核心模型、15+应用场景，目前处于初步探索阶段



*信通院统计的全球79个大模型工业应用案例

- **工业各环节围绕语言、专用、多模态和视觉四类大模型开展探索**
- **当前以大语言模型为主，4类模型应用占比：75%、15%、8%和2%**
- **通用模型的场景化适配调优是主要部署方式，问答交互为主要应用模式**

2.2 (1) 大语言模型：主要应用于工业问答交互、内容生成，以提升任务处理效率为主，暂未触及工业核心环节

有望形成具有认知智能的数字员工及超级自动化链路，实现从需求理解到规划、自动化执行及结果交付的全链条能力

工业相关性

高

低

研发

生产

设备

管理

工业代码生成：基于输入文本实现PLC代码生成与辅助编程

设备控制与维保助手：从基于对话实现设备指令识别与维保知识获取

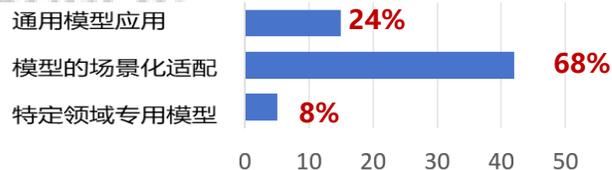
工业管理助手：实现BI数据查询与图表绘制

工业文档外挂与快速检索：实现设备诊断、知识查询、员工培训、设计合规检查

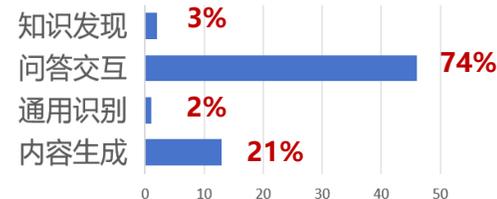
通用文档生成：基于输入文本实现设计方案、报告与邮件编写

● 内容生成 ● 问答交互

部署方式



应用模式



问答交互：应用相对广泛，行业与设备的(半)开放数据是主要语料

运维助手



罗克韦尔：将数字孪生与AIGC相结合，实现设备状态及设备故障原因咨询

管理助手



C3：通过自然语言对话，以文字+统计图完成业务指标的分析洞察

检索助手



中工互联：基于智工大模型实现跨行业、跨领域和跨岗位工业专用知识检索

内容生成：已能实现通用内容和PLC基本控制逻辑代码的生成

工业代码



西门子：与微软合作，基于GPT开发代码生成工具，可通过NLP输入生成PLC代码

通用内容



BACANCY：基于RPA+GPT实现自动邮件回复等功能

2.2 (2) 专业任务大模型：围绕研发形成辅助设计、药物研发两个重点方向，进一步增强研发模式的创新能力

面向工业设计、蛋白质结构预测及药物研发创新等场景，扩展创新边界、降低创新成本与时间

工业相关性

高

需求导向的辅助设计

智能辅助设计：自动生成大量符合需求的设计，并与既有零部件及加工能力结合优化

药物/材料研发突破

高效研发：整合基因库/结构库，加速材料发现、药物分子结构发现、性能预测效率

低

研发

设备、生产、管理

● 格式转换 ● 知识发现

部署方式

模型的场景化适配

29%

特定领域专用模型

71%

应用模式

知识发现

77%

内容生成

13%

格式转换

10%

智能辅助设计：基于图像或文本进行2D-CAD草图构建



DeepMind：基于图像或文本进行2D-CAD草图构建，受样本数量+生成规范的限制，仅个别企业开展验证性探索

- ✓ 基于470万CAD草图数据训练
- ✓ 每个CAD草图对应构建草图生成规范

药物材料研发：聚焦蛋白质/药物的性质、结构与匹配能力的预测优化



Meta：E SMFold模型能够基于序列输入，实现蛋白质结构和序列的预测，模型参数已达150亿，仅2周完成包含罕见物质的6亿+蛋白结构预测

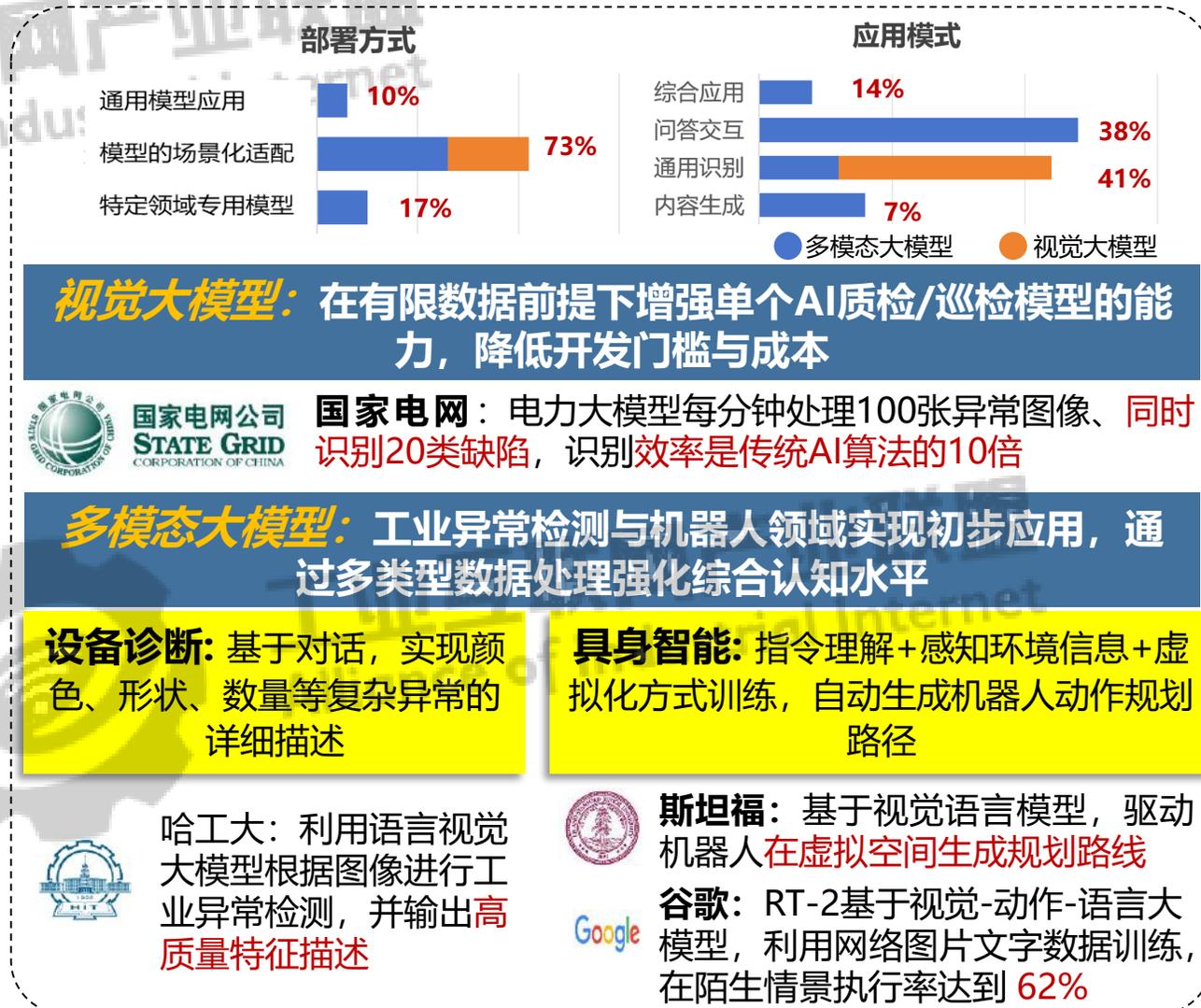
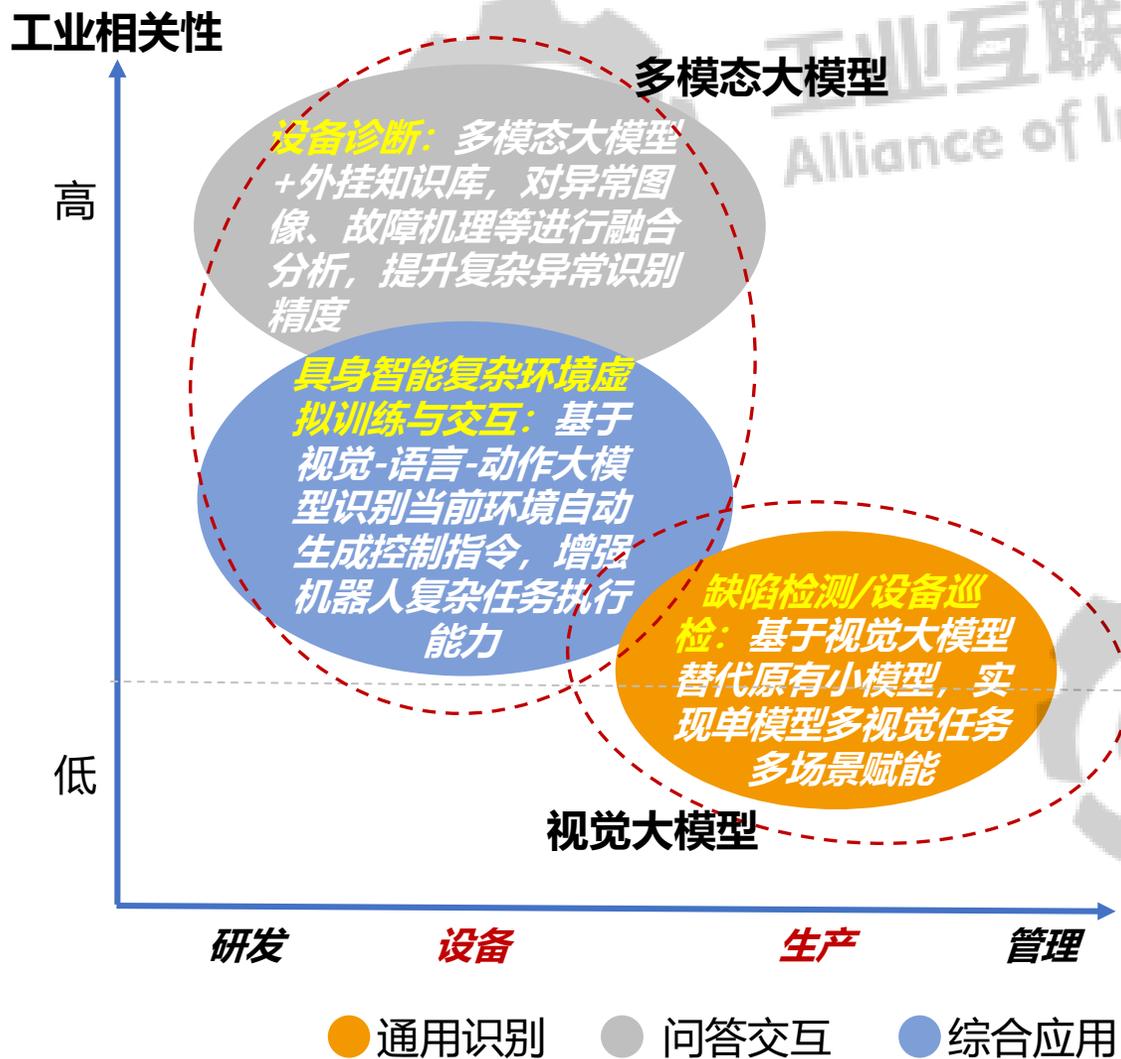


HUAWEI

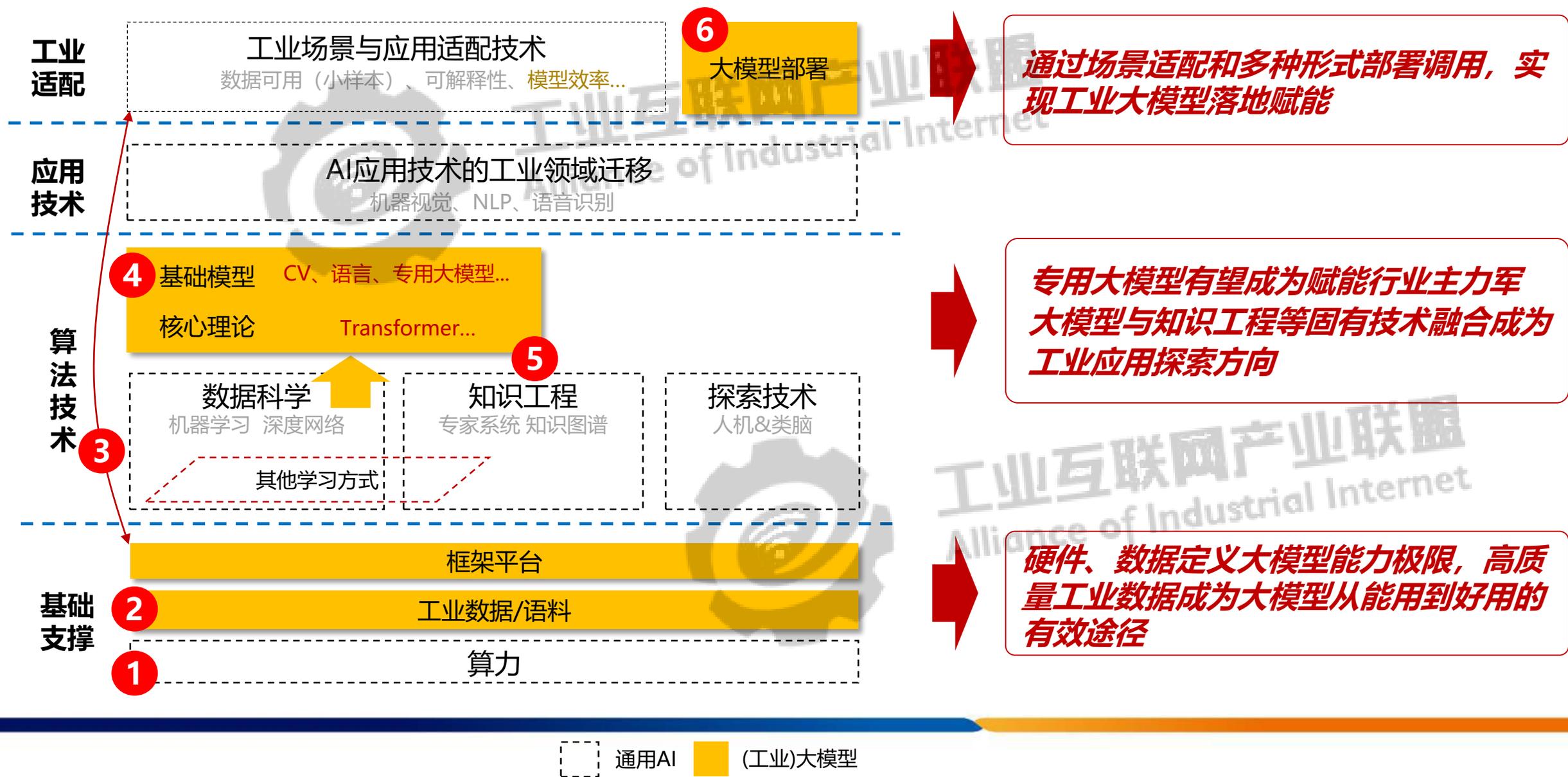
华为：盘古药物分子大模型，能够基于图结构药物分子输入，实现高效的药物分子生成和药物分子定向优化，生成1亿药物分子，新颖性达99.68%

2.2 (3) 多模态大模型与视觉大模型：在装备智能化和视觉识别领域应用获得初步尝试

结合视频、语义、执行等多类型数据综合分析，有望构建认知能力的装备、系统方案及智能工厂



3 技术体系：大模型是工业AI深度学习路径的深化与拓展



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/967135143022006025>