

FROST & SULLIVAN

沙利文



头豹
LeadLeo

2024年 中国AI代码生成市场观测报告 (摘要版)

AI代码生成、人工智能、AI大模型

2024年09月

头豹研究院
弗若斯特沙利文咨询（中国）

研究框架

一、中国AI代码生成行业篇

- 定义、分类、发展历程及市场规模
- 行业现状及行业痛点
- 产业链图谱
- 成本结构及商业模式分析
- 收入主体及结构分析
- 中国及海外行业应用分析
- 行业应用分析-互联网&游戏

二、中国AI代码生成产品篇

- 产品开发底层技术
- 底层技术-数据&模型训练微调
- 产品价值、效能及系统设计分析
- 中国产品情况&国际产品情况

三、中国AI代码生成开发者篇

- AI代码生成对开发者的赋能
- AI代码生成对开发者的影响
- 开发者对AI代码生成的态度
- 开发者交互及心理模式
- 开发者面临的问题
- 开发者喜欢的产品特征
- 开发者未来职业影响

四、中国AI代码生成市场竞争篇

- 评价维度说明
- 综合竞争力表现
- 中国典型厂商代表



行业篇

AI 代码生成行业综述

- AI 代码生成定义
- AI 代码生成分类
- AI 代码生成发展历程
- AI 代码生成市场规模
- 行业现状及痛点
- 产业链图谱
- 成本结构分析
- 商业模式分析
- 收入主体及结构分析
- 中国行业应用分析
- 海外行业应用分析
- 行业应用分析-互联网
- 行业应用分析-游戏

中国AI代码生成行业综述——AI代码生成定义

关键发现

目前，AI代码生成工具在程序员编写代码的整个工作流程中起到了重要辅助作用，帮助程序员提高效率、减少错误、优化代码质量，并加速了软件的交付和迭代。这些工具能理解程序员的需求和意图，自动生成符合规范的代码片段或完整功能模块

AI代码生成赋能编程流程



■ 通过AI代码生成工具的全面集成，程序员们得以将宝贵的精力从繁琐的编码任务中解放出来，更加专注于解决复杂的技术难题和推动创新

通过自动生成代码和自动完成功能，可以有效减少手动编码时间，提高工作效率；智能检测代码中的错误和不足，并提供改进建议，有助于提升代码质量；同时，为初学者提供实时学习资源和编程指导，加速他们的技能提升过程；跨语言支持能够方便程序员在不同项目之间进行切换，提升工作灵活性；在团队协作环境中，智能代码助手能够提供统一的代码风格和最佳实践指导，有助于提升整体团队的效率和协作水平。智能代码助手的集成使得程序员可以将更多精力专注于解决复杂问题和创新，而不是花费在重复性的编码任务上。随着人工智能技术的不断进步，智能代码助手的功能和性能将会不断提升，为程序员的工作带来更多可能性。

来源：沙利文、头豹研究院

中国AI代码生成行业综述——AI代码生成分类

关键发现

AI代码生成功能的分类不是互斥的，很多工具会结合多种功能来提供更全面的编程辅助。随着AI技术的不断发展，这些功能将变得更加智能和高效，进一步推动软件开发行业的创新和生产力提升

AI代码生成功能分类

功能分类	描述	应用场景	优点	缺点
代码补全	生成代码片段或建议，帮助完成当前编写的代码行	编写代码时自动补全函数调用、变量声明	提高编码速度，减少打字量	生成不完全符合上下文的代码
代码生成	根据描述或模型输入生成完整的代码片段或函数	根据业务逻辑描述生成特定功能的代码	加快开发流程，减少手动编写复杂代码的需求	生成的代码可能需要调整以满足特定需求
代码重构	自动识别代码中的改进点并提出重构建议	优化现有代码库，提高代码质量	提升代码可维护性和性能	重构建议可能不完全准确，需要人工审核
测试用例生成	生成测试用例和测试脚本	自动化测试，确保代码质量	减少测试用例编写工作，提高测试覆盖率	生成的测试用例可能需要根据实际情况调整
代码审查辅助	分析代码并提供改进建议	代码审查阶段，提高代码质量	帮助发现潜在问题，促进团队遵循最佳实践	可能产生误报，需要开发者判断
代码搜索与导航	根据自然语言查询提供代码定位和搜索结果	快速定位和理解大型代码库中的特定部分	提高代码理解和导航效率	搜索结果的准确性可能受限于索引和算法
交互式编程	通过自然语言与AI进行对话，获取编程帮助	解决编程问题，学习新技术	提供即时的编程指导和解决方案	可能无法理解复杂的编程问题或上下文

■ AI代码生成工具通过代码补全、代码生成、代码重构等功能提高了编码效率和质量，但也存在生成不完全符合上下文的代码等挑战

通过AI代码生成工具，程序员可以在编写代码时获得多种功能支持。其中，代码补全功能能够生成代码片段或建议，帮助完成当前编写的代码行，例如自动补全函数调用和变量声明，有效提高了编码速度，减少了打字量。代码生成功能则根据描述或模型输入生成完整的代码片段或函数，加快了开发流程，减少了手动编写复杂代码的需求，尽管生成的代码可能需要调整以满足特定需求。代码重构工具能够自动识别代码中的改进点并提出重构建议，优化现有代码库，提高了代码的可维护性和性能，尽管重构建议可能不完全准确，需要人工审核。其他功能还包括测试用例生成、代码审查辅助、代码搜索与导航以及交互式编程，这些功能各有优点和缺点，例如减少了测试用例编写工作、提高了代码质量、帮助发现潜在问题等，但也可能存在生成不完全符合上下文的代码、误报等问题。

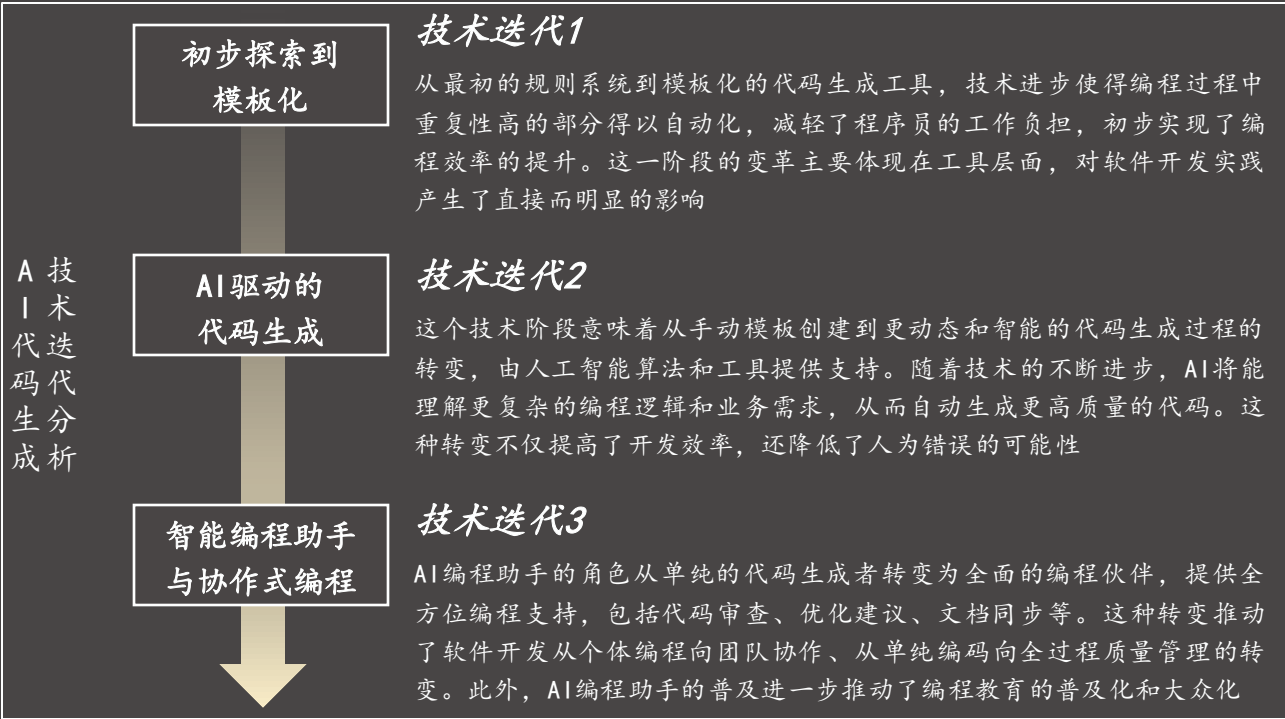
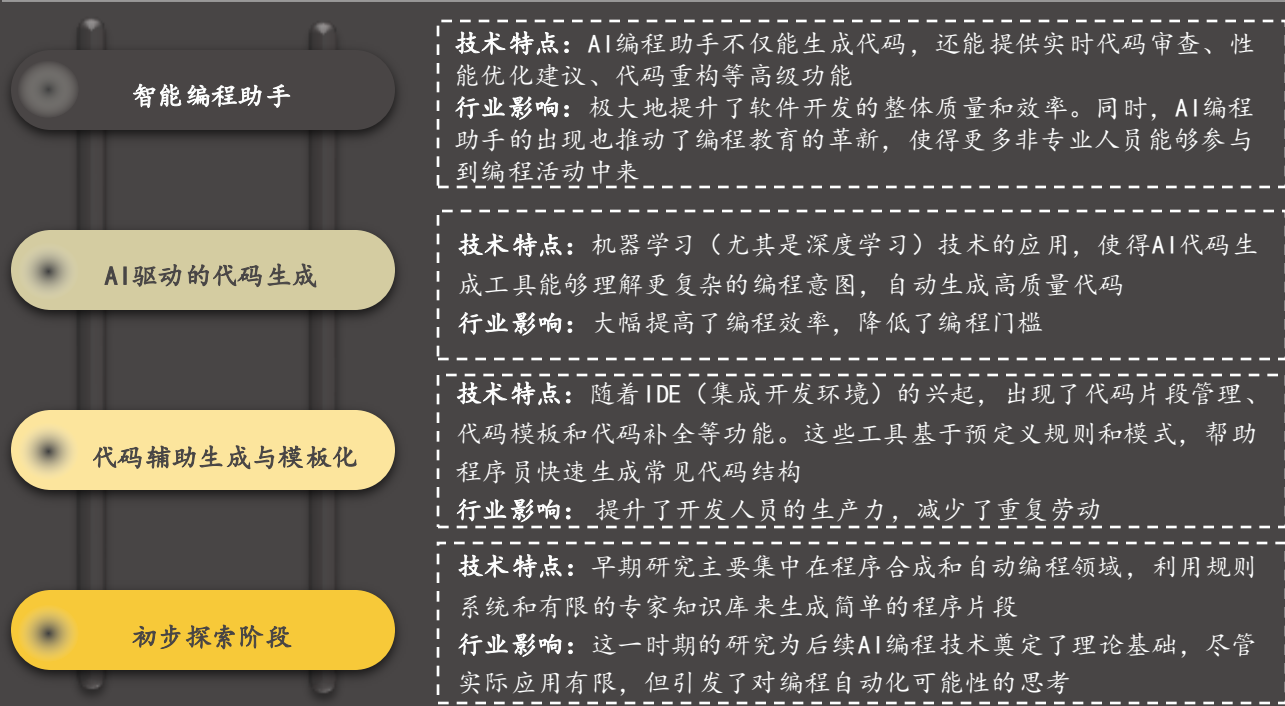
来源：沙利文、头豹研究院

中国AI代码生成行业综述——AI代码生成发展历程

关键发现

AI代码生成的发展历程中，每一次技术迭代都伴随着行业变革，从提升编程效率、扩大开发者群体、改进软件质量到重塑软件开发模式和教育体系，AI技术正在深刻改变软件行业的面貌

AI代码生成发展历程



来源：沙利文、头豹研究院

中国AI代码生成行业综述——行业现状及痛点

关键发现

AI代码生成工具在发展过程中面临的挑战主要集中在用户体验、技术发展、技术提升的难点以及市场适应性等方面。解决这些痛点需不断的技术创新、市场反馈和经验积累。通过持续优化算法模型、增强用户交互设计等，可逐步克服这些挑战

中国AI代码生成行业发展痛点

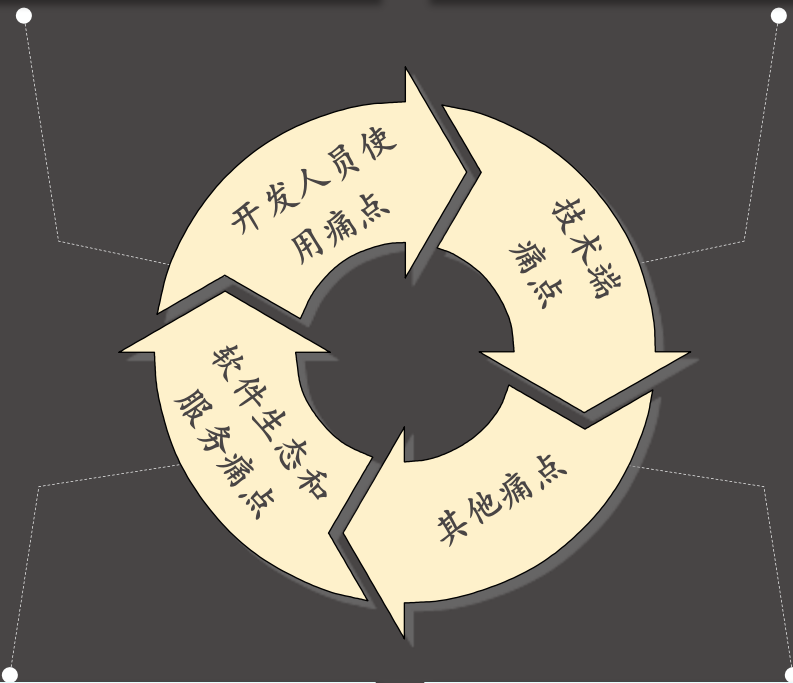
- 易用性

 - 开发人员在使用智能代码生成工具时，面临易用性问题
 - 工具通常是半自动化的，效率提升有限，不能完全替代人工
 - 需要找到自动化和人工之间的平衡
- 开发能力增值

 - 除了提升效率，工具是否能在更多场景和开发能力维度上提供增值价值是待解决的问题
- 跨语言跨平台兼容性

 - 开发原理和编程语言不断升级迭代，跨语言跨平台开发需要更好的兼容性
 - 调试兼容环境是技术测试中需要解决的一部分
- 代码生成技术

 - 未来代码生成可能通过脚本方式完成整个代码的生成
 - 目前技术在复杂任务方面仍有提升空间



- 软件生态融合

 - 目前的软件生态需要进行大量融合工作，以提高兼容性
 - 需要时间和用户反馈来不断完善和迭代
- 经验累积和场景丰富度

 - 相对于国外成熟的市场，中国AI代码生成工具起步较晚，需要时间积累经验和丰富开发场景
- 降低使用门槛

 - 智能代码生成器仍需要一定技术基础，进入门槛有待降低
 - 需要考虑如何使初级用户更容易上手和使用这些工具
- 适用人群扩展

 - 目前的工具主要面向有一定开发基础的用户

来源：沙利文、头豹研究院

中国AI代码生成产业链分析——产业链图谱

关键发现

中国AI代码生成产业链上游包括服务器、AI芯片、存储器、IDE和AI开发平台；中游分为国际AI代码生成厂商和本土AI代码生成厂商；下游主要由软件集成商（40%）和终端用户（60%）两个主体构成

中国AI代码生成行业产业链图谱



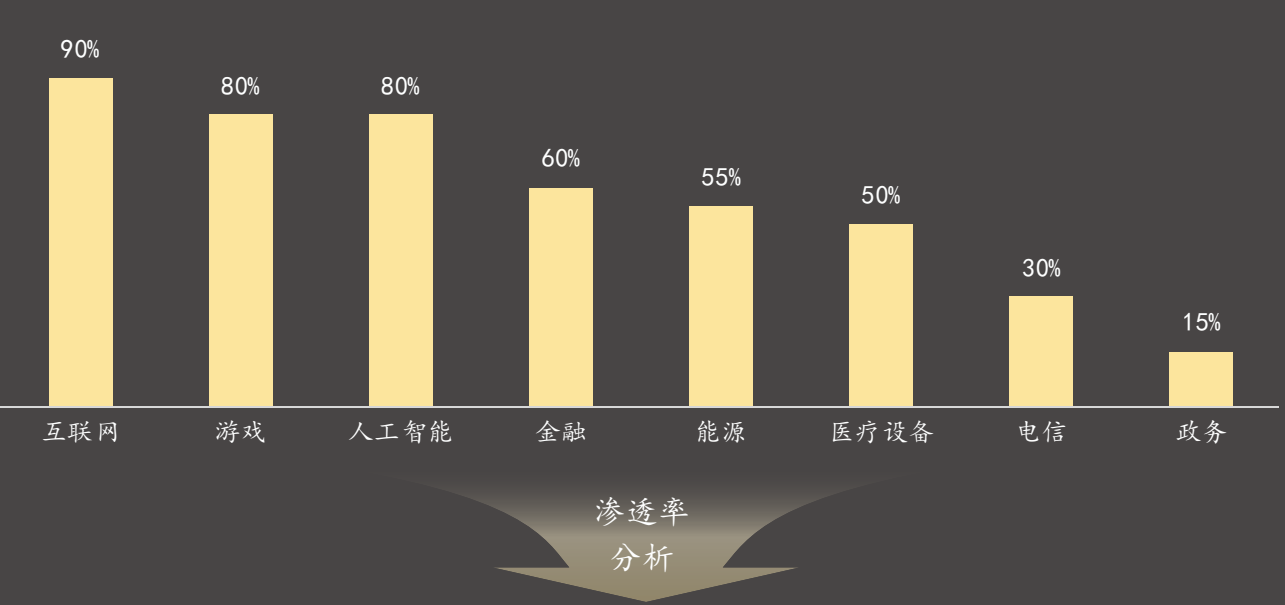
来源：沙利文、头豹研究院

中国AI代码生成产业链分析——中国行业应用分析

关键发现

AI代码生成在互联网和游戏行业的渗透率极高，主要因为这些行业的开发特性（密集、规范、项目制）与AI代码生成的高度契合。此外，这些行业对标准化和效率提升的强烈需求，进一步推动了AI代码生成的应用

中国AI代码生成在各行业的渗透率分析



AI代码生成在各行业的渗透率存在显著差异，这主要源于各行业开发活动的特点、基础代码数量以及迭代速度的不同

- 行业软件开发活动特点
- 互联网和游戏行业由于项目数量多，积累了大量代码样本，为AI代码生成工具提供了丰富的学习材料，从而在这些行业中实现了较高的效率提升。
 - 相比之下，项目数量较少或代码样本不足的行业，AI代码生成工具的学习基础较弱，导致其在辅助开发过程中的效率提升有限。
- 基础代码数量
- 互联网行业由于其庞大的项目数量和快速的迭代周期，积累了大量的基础代码，这为AI代码生成工具提供了丰富的参考和学习资源。
 - 在其他行业，尤其是那些开发活动相对较少或迭代周期较长的行业，基础代码数量有限，从而限制了AI代码生成工具的参考和辅助能力，进而影响效率提升效果。
- 迭代速度与更新频率
- 互联网、游戏以及人工智能行业的快速迭代要求开发工具能够迅速适应新的变化，AI代码生成工具的灵活性在这方面展现出优势。
 - 在迭代速度较慢的行业，AI代码生成工具的快速适应性和持续更新能力不会被充分利用，从而影响其提升效率的潜力。

来源：沙利文、头豹研究院



产品篇

AI代码生成产品洞察

- 产品开发底层技术
- 底层技术-数据
- 底层技术-模型训练微调
- 产品价值分析
- 产品效能分析
- 产品系统设计
- 中国产品情况
- 国际产品情况

中国AI代码生成产品分析——产品开发底层技术

关键发现

在AI代码生成产品技术架构中，关键组件包括芯片、编译器、算法库和深度学习框架。这些组件共同支持模型的训练与微调，确保模型在实际应用场景中表现出色并持续优化。同时，数据源的质量和数据维护工作对于保障模型性能和质量至关重要

AI代码生成产品底层技术分析

AI代码产品		
芯片	中国芯片：寒武纪、昆仑芯、曙光、摩尔线程、华为海思等；国际芯片：NVIDIA、AMD、Intel等	
编译器	针对不同硬件平台的编译器	RISC-V编译器、AI芯片编译器、多核DSP编译器等
	通用算子库	飞桨算子库、MindSpore算子库、MegEngine算子库、TNN算子库、Triton算子库等
算法	深度学习框架	国际AI算法框架：TensorFlow、PyTorch、Theano、Caffe2、Torch等 中国AI算法框架：飞桨（PaddlePaddle）、天元（MegEngine）、MindSpore等
模型训练微调	训练原则	融合与创新：建立动态评估体系，衡量模型在标准测试和实际场景中的表现，包括用户反馈、功能测试和长尾场景分析 产品驱动的增量训练：对于代码模型，需采用产品驱动的训练策略，根据需求设计增量训练，优先提升核心功能相关能力，如跨文件引用理解、特定语言特性掌握 边界问题的前瞻性分析与修复：边界问题通常涉及模型在处理复杂输入、罕见语法结构或异常逻辑时的表现。通过对这些问题进行深入的前瞻性分析，能够发现模型在特定条件下的潜在漏洞
	训练方法	训练方法深化：监督微调、对抗训练、迁移学习、联合学习、强化学习 动态预训练：持续收集最新开源代码，覆盖多种语言和框架，确保技术栈适应性 全方位模型评估：在代码模型的评估中，除了利用自动化测试集全面检验模型性能外，还应结合智能代理进行更贴近实际使用的评估 基于场景的训练目标设计：针对不同任务需求，代码模型训练需设定专门目标。对于需要上下文理解的任务，如代码自动补全，模型需要接受长序列输入的训练，并进行上下文相关性的优化
	数据来源	代码数据：开源代码库、企业专有的代码数据、开源数据集与代码示例网站 非代码数据：开发场景中各环节的自然语言对话、用户行为、用户反馈、AI合成数据
数据	数据维护	数据清洗与预处理：异常与缺失值处理、去噪、格式转换与标准化 特征工程：数据缩放、编码、构建特征和数据富集 数据标注与分类：自动化标注工具的开发和使用、数据的分类与标签分配 质量控制与合规性：严格的质量控制流程和合规性检查，确保数据符合标准和法律要求 数据评估与监控：多维度的数据价值评估、模型性能测试、偏差检测和持续监控

来源：沙利文、头豹研究院



市场竞争篇

AI 代码生成产品竞争分析

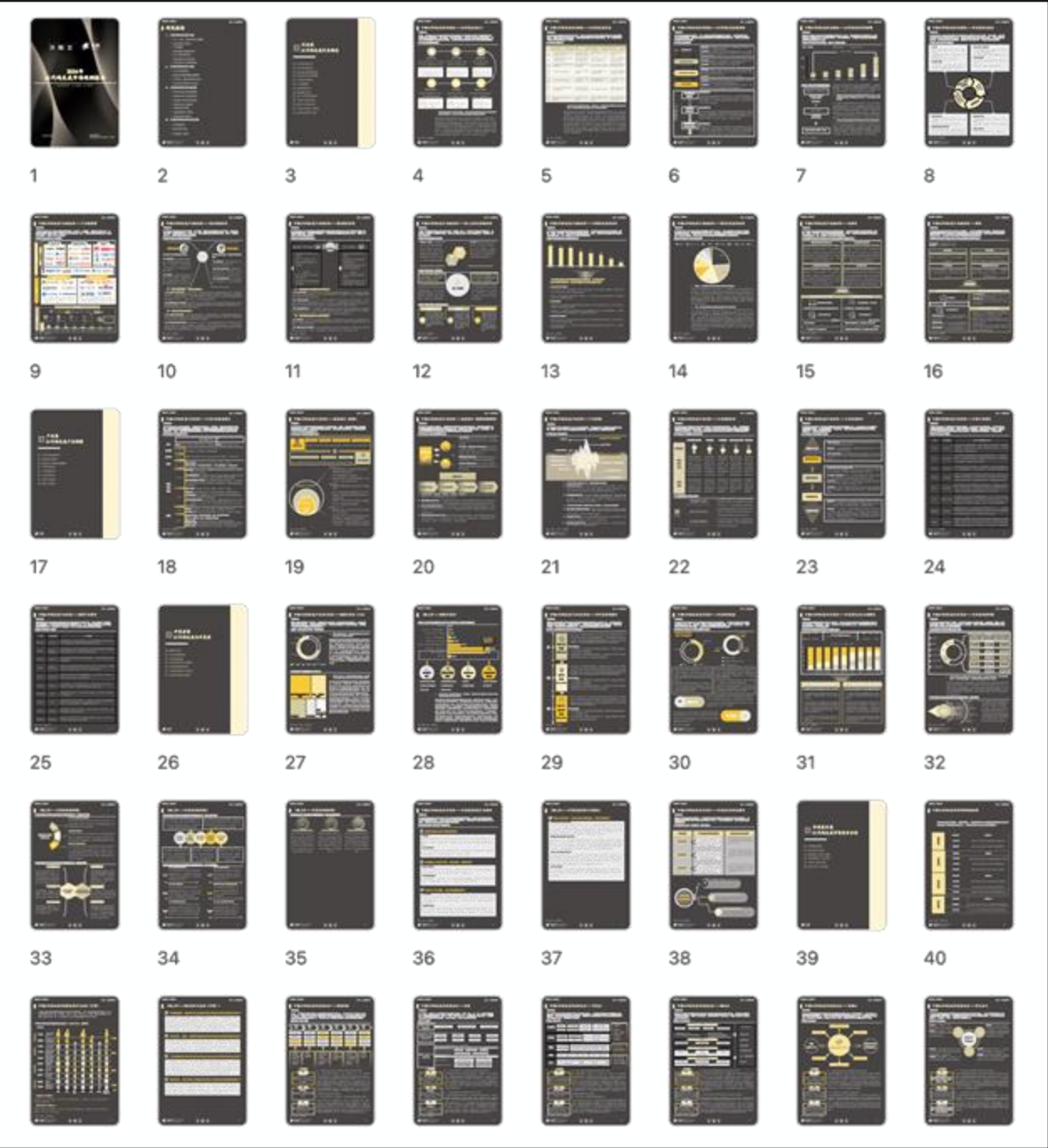
- 评价维度说明
- 综合竞争力表现
- 代表性厂商分析

中国AI代码生成综合竞争力表现

关键发现

沙利文联合头豹依据AI代码生成产品评价维度，从产业生态图谱中筛选出六家中国典型AI代码厂商：阿里、百度、科大讯飞、商汤、腾讯和智谱AI以及国际主流AI代码厂商，并对其AI代码生成产品的竞争力进行综合评估

报告完整版登录www.leadleo.com 搜索《2024年AI代码生成市场观测报告》



方法论

- ◆ 头豹研究院布局中国市场，深入研究19大行业，532个垂直行业的市场变化，已经积累了近100万行业研究样本，完成近10,000多个独立的研究咨询项目。
- ◆ 研究院依托中国活跃的经济环境，从纵深防御、快速响应、轻量化部署等领域着手，研究内容覆盖整个行业的发展周期，伴随着行业中企业的创立，发展，扩张，到企业走向上市及上市后的成熟期，研究院的各行业研究员探索和评估行业中多变的产业模式，企业的商业模式和运营模式，以专业的视野解读行业的沿革。
- ◆ 研究院融合传统与新型的研究方法，采用自主研发的算法，结合行业交叉的大数据，以多元化的调研方法，挖掘定量数据背后的逻辑，分析定性内容背后的观点，客观和真实地阐述行业的现状，前瞻性地预测行业未来的发展趋势，在研究院的每一份研究报告中，完整地呈现行业的过去，现在和未来。
- ◆ 研究院密切关注行业发展最新动向，报告内容及数据会随着行业发展、技术革新、竞争格局变化、政策法规颁布、市场调研深入，保持不断更新与优化。
- ◆ 研究院秉承匠心研究，砥砺前行的宗旨，从战略的角度分析行业，从执行的层面阅读行业，为每一个行业的报告阅读者提供值得品鉴的研究报告。

法律声明

- ◆ 本报告著作权归头豹所有，未经书面许可，任何机构或个人不得以任何形式翻版、复刻、发表或引用。若征得头豹同意进行引用、刊发的，需在允许的范围内使用，并注明出处为“头豹研究院”，且不得对本报告进行任何有悖原意的引用、删节或修改。
- ◆ 本报告分析师具有专业研究能力，保证报告数据均来自合法合规渠道，观点产出及数据分析基于分析师对行业的客观理解，本报告不受任何第三方授意或影响。
- ◆ 本报告所涉及的观点或信息仅供参考，不构成任何证券或基金投资建议。本报告仅在相关法律许可的情况下发放，并仅为提供信息而发放，概不构成任何广告或证券研究报告。在法律许可的情况下，头豹可能会为报告中提及的企业提供或争取提供投融资或咨询等相关服务。
- ◆ 本报告的部分信息来源于公开资料，头豹对该等信息的准确性、完整性或可靠性不做任何保证。本报告所载的资料、意见及推测仅反映头豹于发布本报告当日的判断，过往报告中的描述不应作为日后的表现依据。在不同时期，头豹可发出与本报告所载资料、意见及推测不一致的报告或文章。头豹均不保证本报告所含信息保持在最新状态。同时，头豹对本报告所含信息可在不发出通知的情形下做出修改，读者应当自行关注相应的更新或修改。任何机构或个人应对其利用本报告的数据、分析、研究、部分或者全部内容所进行的一切活动负责并承担该等活动所导致的任何损失或伤害。

报告完整版登录 www.leadleo.com

搜索《2024年AI代码生成市场观测报告》

首席分析师

袁栩聪

☎ 15999806788

✉ oliver.yuan@frostchina.com

研究总监

李庆

☎ 13149946576

✉ livia.li@frostchina.com

 www.frostchina.com ; www.leadleo.com

 <https://space.bilibili.com/647223552>

 <https://weibo.com/u/7303360042>

©弗若斯特沙利文咨询（中国）

©头豹研究院

