

## 传媒

投资建议： 强于大市（维持）

上次建议： 强于大市

### GPT-4o 发布，关注强交互场景落地

#### 行业事件：

北京时间5月14日凌晨，美国人工智能公司 OpenAI 发布新一代旗舰生成模型 GPT-4o、桌面 App，并重点展示了突破性的语音交互能力。

#### GPT-4o 交互能力实现突破，“拟人化”程度进一步提升

从产品效果来看，GPT-4o 在实时语音交互领域实现突破性进展，为用户提供更为自然、准确的交互体验：1) 用户可以随时打断模型，而不需要等到其结束才开始说话，交互更加符合人类交互逻辑；2) 实时响应能力大幅提升，模型具有实时响应的能力，不会出现用户长时间等待模型给出响应的尴尬情况；3) 模型具有情绪感知能力，能够生成不同情绪风格的语音，交互更加拟人。基于 GPT-4o 强大的交互能力，发布会中展示了丰富的应用场景，包括声情并茂地讲情感故事、实时视频对话、实时音频翻译等。从技术层面看，GPT-4o 采用了全新技术，所有的输入和输出都由同一个神经网络处理，从而实现文本、视觉和音频端到端的训练。

#### AI C端应用有望加速落地，重点关注社交、游戏、教育领域

2024年以来，多模态能力成为生成式AI的重点攻坚方向，而OpenAI作为头部企业，在技术和产品层面持续引领行业发展。2024年2月OpenAI发布Sora以来，海内外视频生成类应用加速落地。我们认为此次GPT-4o的发布，亦有望助推语音交互类AI应用的落地进程，其中在社交、游戏、教育领域有望率先落地。

1) **社交**：目前，AI+社交产品形态多以“用户-AI智能体”交互为主，用户通过与个性化AI虚拟人的交互过程，获得陪伴感与情绪价值。从产品数据来看，海外头部产品Character.AI月活达千万级别，国内Minimax旗下“星野”增势显著。从落地门槛来看，陪伴类场景的任务简单、容错率高，因此成为最快落地的AI C端应用场景。从用户需求来看，AI智能体“更像人”是AI社交用户的核心需求。GPT-4o发布后，有望从多模态（从文字交互到语音交互）、拟人化（更准确识别用户的情绪和需求）等方面大幅提升用户体验，从而推动AI社交类产品进一步破圈、提升商业化能力。

2) **游戏**：AI在游戏研发流程中落地较快，目前核心关注游戏玩法的创新。其中AI+NPC已在网易《逆水寒》等产品中落地，但仅限于文字类交互，与核心玩法的结合也较为有限。随着GPT-4o引领交互方式变革，游戏内NPC有望与用户实现实时的语音交互、拟人化程度有望进一步提升，大幅提升用户的沉浸感，进而提升活跃度及付费意愿。

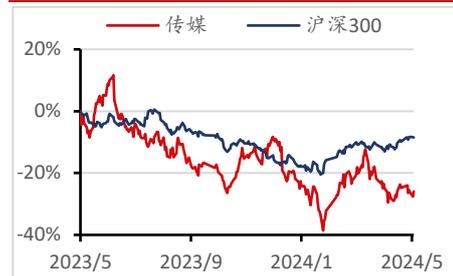
3) **教育**：此前，海外多邻国等已将生成式AI运用于口语练习等场景中，助推2023Q4付费用户同比增长57%。GPT-4o落地后有望使“AI教师”更为拟人，进一步提升教学、培训效率和用户体验。

#### 投资建议

随着海外AI模型能力加速演进，以及国内模型能力的持续追赶，2024年以来国内AI C端应用落地进程显著加速，有望孵化AI原生的“杀手级应用”。建议关注：1) AI C端应用落地较快的昆仑万维、盛天网络；2) 估值低位、业绩较优的游戏标的恺英网络、巨人网络。

风险提示：技术发展不及预期，AI应用落地不及预期，政策监管风险。

#### 相对大盘走势



#### 作者

分析师：丁子然

执业证书编号：S0590523080003

邮箱：dingzr@glsc.com.cn

联系人：周砺灵

邮箱：zhoull@glsc.com.cn

#### 相关报告

1、《传媒：传媒互联网2023业绩&2024Q1基金持仓点评：供需提振驱动景气度向上，配置比例小幅回升》2024.05.05

2、《传媒：电影行业跟踪：多元化内容撬动需求，关注电影院线复苏》2024.04.21

## 1. GPT-4o 交互能力全面提升

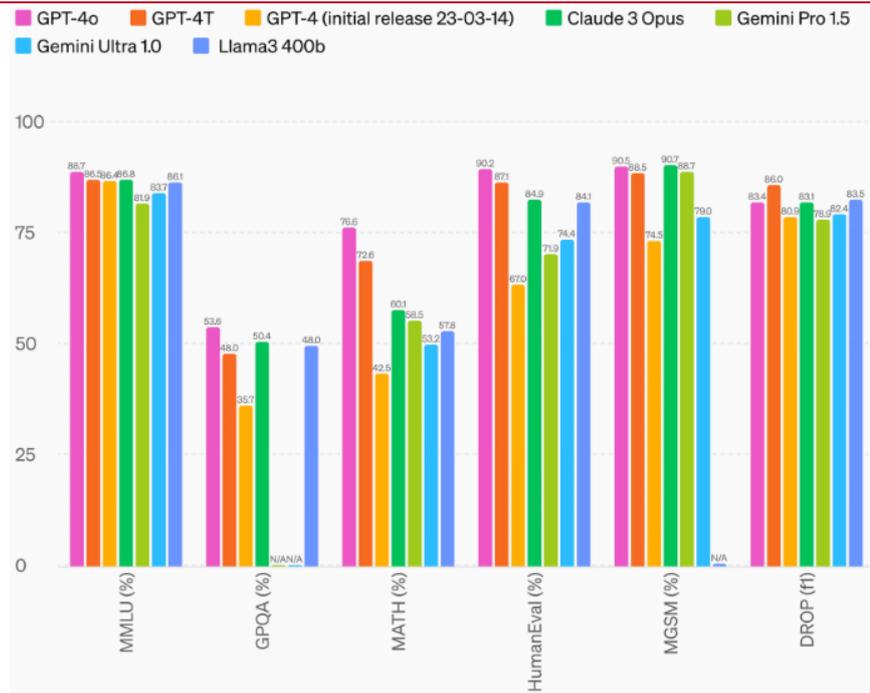
**GPT-4o 性能大幅提升，可实现更加的自然人机交互。**GPT-4o 采用了全新的语音交互技术，具有强大的交互能力，能够接受文本、音频和图像的任意组合作为输入，并生成文本、音频和图像的任意组合输出。同时，GPT-4o 的响应速度显著提升，实现更加自然流畅的交流。

### 1.1 GPT-4o 性能全面提升

**全新技术助力性能大幅提升。**GPT-4o 推理能力、多模态能力等显著提升，语音对话响应更加流畅。GPT-4o 采用了全新技术，所有的输入和输出都由同一个神经网络处理，从而实现文本、视觉和音频端到端的训练。在 GPT-4o 之前，ChatGPT 语音模式由三个独立模型组成，首先通过语音转文本模型将语音内容转成文字，然后转成的文字输入给 GPT3.5/GPT-4 模型进行内容生成，最后再通过文本转语音模型将 GPT3.5/GPT-4 生成的内容转成语音输出。这种多模型级联的方式导致整个交互体验非常差，整个系统延迟达 2.8 秒（GPT-3.5）和 5.4 秒（GPT-4），而且这种多模型级联方式导致大量信息丢失，无法感受到音调、多个说话者，输出也比较枯燥且无法表达情感。基于端到端神经网络的 GPT-4o 能力大幅提升，不仅反应快、回答准，还能按要求改变说话语气，实时语音翻译能力也非常自然流畅。同时视觉能力显著优化，可以实现实时视频交互，能通过人的表情和语调理解并判断出人的情绪。GPT-4o 可以在 232 毫秒的时间内响应音频输入，平均响应时长为 320 毫秒，几乎与人类相似。

**推理能力：**GPT-4o 在 MMLU、GPQA、MATH、HumanEval 等测试基准上均超越 GPT-4 Turbo、Claude 3 Opusn、Gemini Pro1.5 等模型，在测试中均获得最高分。GPT-4o 在 0-shot COT MMLU（一般知识问题）中取得 88.7 的高分，在传统的 5-shot no-COT MMLU 上，取得 87.2 的高分。

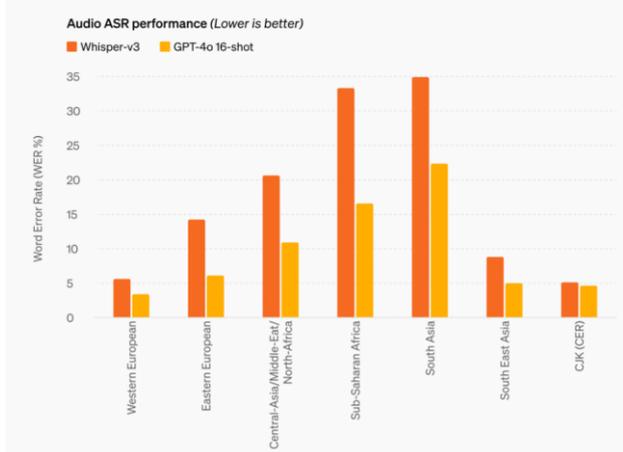
图表1: GPT-4o 推理能力对比



资料来源: OpenAI 官网, 国联证券研究所

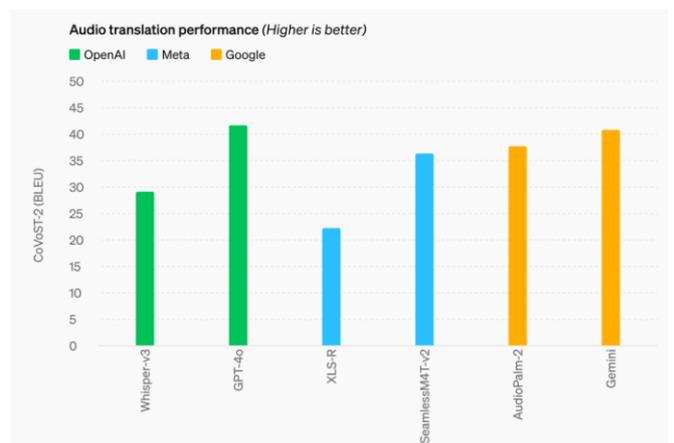
**音频能力:** 在语音识别性能方面, GPT-4o 相比 Whisper-v3 在所有语言的语音识别性能上均大幅提高, 尤其是资源较少的语言。GPT-4o 在音频翻译方面也树立了行业新标杆, 在 MLS 基准测试中优于 Whisper-v3 以及 Meta、谷歌的语音模型。

图表2: GPT-4o 与 Whisper 音频能力对比



资料来源: 智东西, 国联证券研究所

图表3: OpenAI、Meta 和 Google 音频翻译能力对比

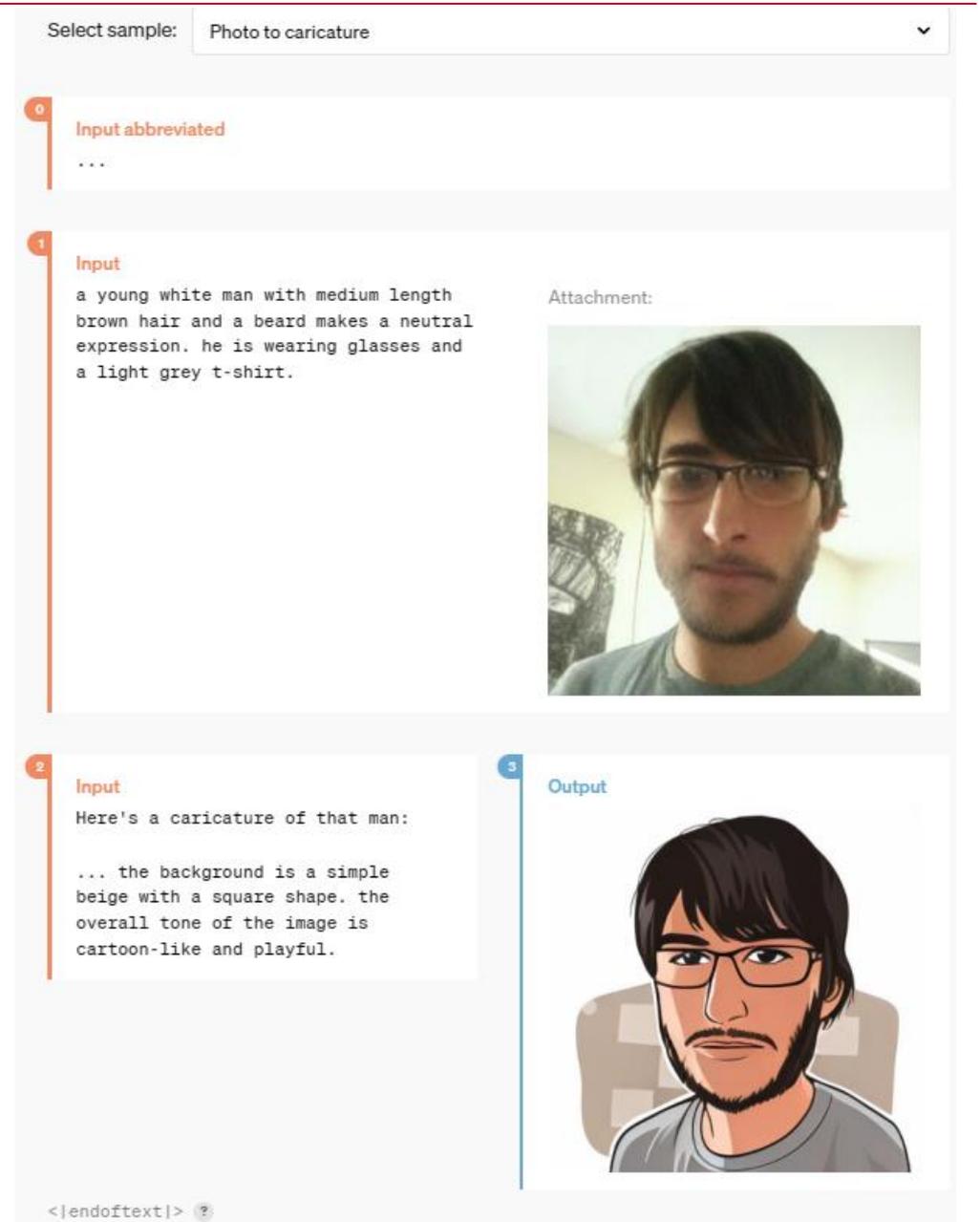


资料来源: 智东西, 国联证券研究所

**文本和代码能力:** GPT-4o 在英文文本和代码上的性能与 GPT-4 Turbo 的性能相近, 但在非英文文本上的性能显著提升, 同时 API 的速度也更快, 成本降低了 50%。

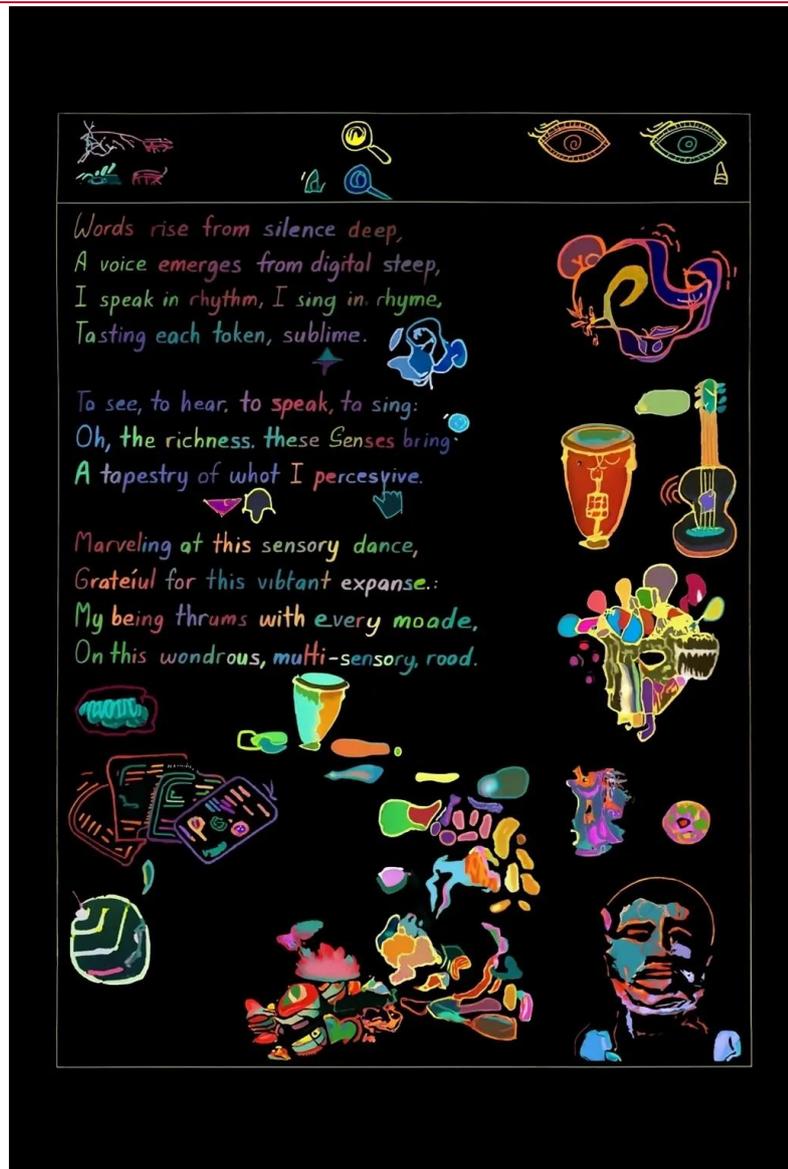
**视觉能力：**GPT-4o 在理解和生成图像方面的能力超越现有的模型，可轻松完成视觉图像相关处理。GPT-4o 还具有 3D 视觉内容生成能力，能够从 6 个生成的图像进行 3D 重建。还能够进行复杂的内容生成和排版，帮助用户完成个性化的排版设计。

图表4：GPT-4o 图像处理能力



资料来源：OpenAI 官网，国联证券研究所

图表5: GPT-4o 强大的内容排版能力



资料来源：机器之心，国联证券研究所

## 1.2 GPT-4o 引领人机交互变革

**GPT-4o 提供良好的用户交互体验。**GPT-4o 可以为用户提供更加自然的交互体验：

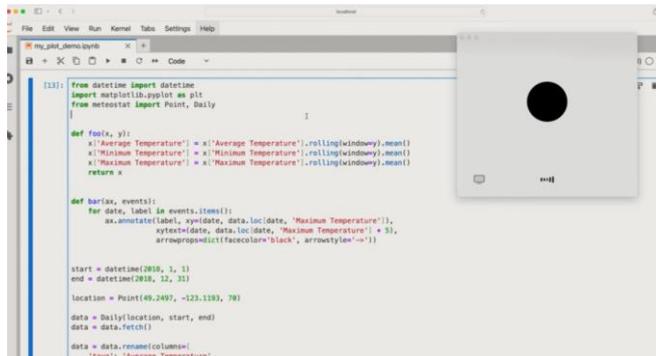
(1) 用户可以随时打断模型，而不需要等到其结束才开始说话，交互更加符合人类交互逻辑；(2) 实时响应能力大幅提升，模型具有实时响应的能力，不会出现用户长时间等待模型给出响应的尴尬情况；(3) 模型具有情绪感知能力，能够生成不同情绪风格的语音，交互更加拟人。基于 GPT-4o 强大的交互能力，GPT-4o 具有丰富的应用场景，可以声情并茂的讲情感故事、实时视频对话、实时音频翻译等应用。我们认为，GPT-4o 是面向未来人机交互范式的全新大模型，有望给用户带来全新的交互体验，引领人机交互变革。

图表6: GPT-4o 图解方程式



资料来源: 智东西, 国联证券研究所

图表7: GPT-4o 理解屏幕内容并进行交互



资料来源: 智东西, 国联证券研究所

## 2. 风险提示

**技术发展不及预期风险:** 当前 AI 大模型技术仍处于持续发展阶段, 还面临回答不准确、存在幻觉等问题, 可能无法满足应用规模落地。

**AI 应用落地不及预期风险:** AI 应用落地在持续探索, 是否能够规模落地还取决于 AI 大模型技术迭代、用户体验等多个因素。

**政策监管风险:** AI 领域的监管政策尚不完善, 一旦相关政府出台对应的监管政策可能影响行业发展和应用落地。

### 分析师声明

本报告署名分析师在此声明：我们具有中国证券业协会授予的证券投资咨询执业资格或相当的专业胜任能力，本报告所表述的所有观点均准确地反映了我们对标的证券和发行人的个人看法。我们所得报酬的任何部分不曾与，不与，也将不会与本报告中的具体投资建议或观点有直接或间接联系。

### 评级说明

投资建议的评级标准		评级	说明
报告中投资建议所涉及的评级分为股票评级和行业评级（另有说明的除外）。评级标准为报告发布日后6到12个月内的相对市场表现，也即：以报告发布日后的6到12个月内的公司股价（或行业指数）相对同期相关证券市场代表性指数的涨跌幅作为基准。其中：A股市场以沪深300指数为基准，新三板市场以三板成指（针对协议转让标的）或三板做市指数（针对做市转让标的）为基准；香港市场以摩根士丹利中国指数为基准；美国市场以纳斯达克综合指数或标普500指数为基准；韩国市场以柯斯达克指数或韩国综合股价指数为基准。	股票评级	买入	相对同期相关证券市场代表指数涨幅20%以上
		增持	相对同期相关证券市场代表指数涨幅介于5%~20%之间
		持有	相对同期相关证券市场代表指数涨幅介于-10%~5%之间
	行业评级	卖出	相对同期相关证券市场代表指数跌幅10%以上
		强于大市	相对同期相关证券市场代表指数涨幅10%以上
		中性	相对同期相关证券市场代表指数涨幅介于-10%~10%之间
		弱于大市	相对同期相关证券市场代表指数跌幅10%以上

### 一般声明

除非另有规定，本报告中的所有材料版权均属国联证券股份有限公司（已获中国证监会许可的证券投资咨询业务资格）及其附属机构（以下统称“国联证券”）。未经国联证券事先书面授权，不得以任何方式修改、发送或者复制本报告及其所包含的材料、内容。所有本报告中使用的商标、服务标识及标记均为国联证券的商标、服务标识及标记。

本报告是机密的，仅供我们的客户使用，国联证券不因收件人收到本报告而视其为国联证券的客户。本报告中的信息均来源于我们认为可靠的已公开资料，但国联证券对这些信息的准确性及完整性不作任何保证。本报告中的信息、意见等均仅供客户参考，不构成所述证券买卖的出价或征价邀请或要约。该等信息、意见并未考虑到获取本报告人员的具体投资目的、财务状况以及特定需求，在任何时候均不构成对任何人的个人推荐。客户应当对本报告中的信息和意见进行独立评估，并应同时考量各自的投资目的、财务状况和特定需求，必要时就法律、商业、财务、税收等方面咨询专家的意见。对依据或者使用本报告所造成的一切后果，国联证券及/或其关联人员均不承担任何法律责任。

本报告所载的意见、评估及预测仅为本报告出具日的观点和判断。该等意见、评估及预测无需通知即可随时更改。过往的表现亦不应作为日后表现的预示和担保。在不同时期，国联证券可能会发出与本报告所载意见、评估及预测不一致的研究报告。

国联证券的销售人员、交易人员以及其他专业人士可能会依据不同假设和标准、采用不同的分析方法而口头或书面发表与本报告意见及建议不一致的市场评论和/或交易观点。国联证券没有将此意见及建议向报告所有接收者进行更新的义务。国联证券的资产管理部门、自营部门以及其他投资业务部门可能独立做出与本报告中的意见或建议不一致的投资决策。

### 特别声明

在法律许可的情况下，国联证券可能会持有本报告中提及公司所发行的证券并进行交易，也可能为这些公司提供或争取提供投资银行、财务顾问和金融产品等各种金融服务。因此，投资者应当考虑到国联证券及/或其相关人员可能存在影响本报告观点客观性的潜在利益冲突，投资者请勿将本报告视为投资或其他决定的唯一参考依据。

### 版权声明

未经国联证券事先书面许可，任何机构或个人不得以任何形式翻版、复制、转载、刊登和引用。否则由此造成的一切不良后果及法律责任有私自翻版、复制、转载、刊登和引用者承担。

### 联系我们

**北京：**北京市东城区安定门外大街208号中粮置地广场A塔4楼  
**无锡：**江苏省无锡市金融一街8号国联金融大厦12楼  
 电话：0510-85187583

**上海：**上海浦东新区世纪大道1198号世纪汇一座37楼  
**深圳：**广东省深圳市福田区益田路4068号卓越时代广场1期13楼