

📅 0001年1月1日 ⌚ 1 分钟阅读

## 整体概况

产品线分为： 模型， AI平台(开发和部署,比如Vertex AI， AI Studio)， AI 应用（ Gemini app, NotebookLM, Learn about, DeepResearch， Astra)

## 大模型

多模态大模型： Gemini， Gemma 图像大模型： Imagen， PaLM-E  
视频大模型： Veo

## 开发环境

Google AI Studio  
Google Vertex AI  
Google Cloud AI

## 应用

Google Learn about  
Google AI Research  
Google NotebookLM  
Google Gemini App

## NotebookLM

<https://notebooklm.google/>

我主要用它来阅读论文，科技报告，长篇技术文章等等。 它的一些新功能不错，比如思维导图，生成双人播客（用户可以随时插入讨论）， 视频生成

## 目录

## 文章信息

字数

阅读时间

发布时间

### Astra (Project Astra)

还在内测！

谷歌的Astra是由 DeepMind 开发的先进人工智能代理，旨在通过多模态功能无缝融入日常生活。Astra 处理并响应各种输入——文本、图像、视频和音频——使其成为高度交互和直观的助手。Astra 的突出特点包括用于上下文理解的实时记忆、高级工具使用（例如谷歌搜索、地图和镜头）以及协助完成诸如识别物体或提供建议等任务的能力。例如，用户可以将手机对准书架并询问 Astra 识别评分最高的书，展示了其连接数字世界和物理世界的独特能力。随着将其部分功能集成到其他谷歌产品的计划，Astra 将在 2025 年及以后重新定义个人人工智能。目前，Project Astra正在由有限数量的受信任用户进行测试，以完善其功能并为通用人工智能助手发现新用途。

人工智能代理结合多模态功能来处理文本、图像、视频和音频，创造出由谷歌 Gemini AI 模型支持的自然对话体验。

主要特点：

- 多模态处理（文本、图像、视频、音频）
- 实时对象识别和鉴定
- 具有对过去交互记忆的情境感知响应
- 与各种设备（手机、原型眼镜）集成
- 创造性的问题解决和讲故事能力
- 多语言支持及混合语言理解
- 工具使用（谷歌搜索、Lens 和地图集成）
- 改进的会话中及长期记忆

优点：

- 先进的人工智能能力，可实现自然、情境感知的交互
- 在各种任务和场景中有多重应用
- 通过可信测试人员的反馈持续改进
- 与谷歌生态系统（搜索、Lens、地图）集成
- 有潜力彻底改变日常协助和问题解决

缺点：

- 可用性有限（目前处于研究原型阶段）
- 大量数据处理可能引发隐私担忧

可能引发关于人工智能在日常生活中作用的伦理问题  
依赖谷歌生态系统和互联网连接

## project-mariner

还在内测！ <https://deepmind.google/technologies/project-mariner/>

基于 Gemini 2.0 框架构建，这款实验性的 Chrome 扩展程序可以导航网站、理解屏幕内容并代表用户执行操作，标志着朝着更直观的人机交互迈出了重要一步。

### 主要特点：

- 导航网站、移动光标、点击按钮并填写表单。
- 支持在 Chrome 标签页中实时执行任务。
- 将复杂任务分解为可管理的步骤。
- 处理屏幕截图以解释网页内容。
- 通过聊天窗口接受用户指令。
- 采取顶级安全措施——对敏感操作请求用户确认。
- 与谷歌生态系统集成，利用 Gemini AI 模型和 Chrome 浏览器。

### 优点：

- 简化基于网络的任务，节省时间和精力。
- 适应各种场景，如在线购物、旅行预订和研究。
- 为其操作提供逐步解释以提高透明度。
- 仅在活动标签页中操作并寻求权限，确保用户控制。

### 缺点：

- 目前速度较慢，操作之间有明显延迟。
- 仅局限于 Chrome 浏览器和活动标签页。
- 不能执行某些操作，如接受 cookie 或完成购买。
- 由于屏幕截图处理，存在潜在的隐私问题。

## 技术规范

[AI Principles](#) [AI Safety](#) [AI Ethics](#) [AI Governance](#)

学术

DeepMind

<https://deepmind.google/discover/events/>

分享这篇文章

