

# 全球半导体市场机遇

拥抱人工智能，扩大市场份额



# 前景乐观

当今科技创新迅猛发展，半导体行业有望持续增长。对半导体行业来说，2019年会相对疲软，但普华永道预计其将在2020年实现复苏并保持繁荣。2018年半导体行业销售总额为4,810亿美元。今后四年，即到2022年底，我们预计销售额将保持较慢但稳健的增长，复合年均增长率（CAGR）约为4.6%，达到5,750亿美元。

半导体行业由七类元件组成，即内存、逻辑元件、微型元件、模拟元件、光电元件、传感器和分离元件（OSD）。其中，内存产品销售额仍将是半导体收入的最大份额。然而，三星集团在2017至2018年对其半导体部门的巨额投入将使内存市场产能过剩，从而导致内存产品（尤其是3D NAND闪存产品）在2019年销量下滑，但该市场有望在2020年开始复苏。

此外，人工智能（AI）应用的快速增长带来的芯片需求，将极大促进该行业的整体增长。大部分需求来自汽车和工业市场，这两个领域增长最快。

由于电动汽车和混合动力汽车的普及率不断提高，再加上自动驾驶汽车的市场潜力巨大，汽车市场将是增长最快的市场。到2022年，其复合年均增长率将达到11.9%。与此同时，传统汽车芯片的需求依然强劲。工业市场继续受到安全和医疗领域对人工智能芯片及其实力的需求的推动。在此期间，整个工业市场的复合年均增长率预计将达到10.8%。

由于智能手机的更新换代、5G技术的引入以及新兴市场的增长，通信市场的复合年均增长率将达到2.2%。与此同时，到2022年底，消费类电子产品市场约有50%的收入将来自电视、视频游戏机、手持设备和数字机顶盒。

## 医疗市场机遇

随着医疗领域公司从大数据、数据分析和人工智能获得越来越多的益处，包括医疗器械制造商、制药公司和相关技术公司等企业将拥有前所未有的增长机会。就美国医疗器械市场来说，这一增长很大程度上来自于美国食品药品监督管理局出台的“2016年数字健康创新计划”。该计划鼓励传统医疗公司采用新技术，并与数字医疗初创公司合作。

苹果和谷歌等科技巨头已经对不断增长的医疗市场表现出相当大的兴趣。苹果的目标是通过提供综合平台（如ResearchKit和CareKit）来整合支离破碎的医疗应用市场，而谷歌正努力通过其DeepMind人工智能解决方案来改变这一行业。其他科技公司也开始参与其中，例如辉达（Nvidia）正在使用（GPU）驱动的深度学习技术，通过超声图像更快速地诊断癌症。

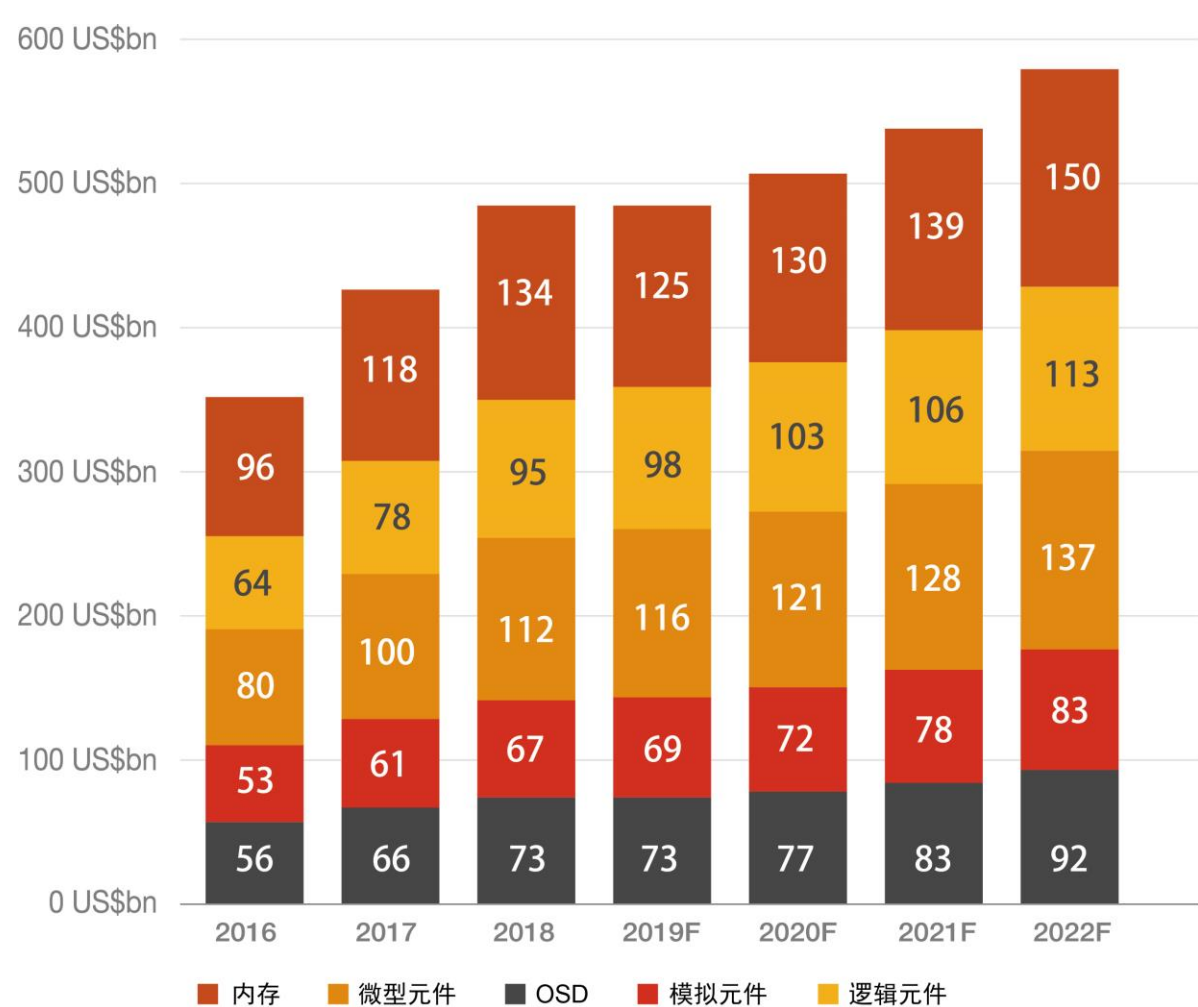
到2022年，消费类电子产品市场的复合年均增长率将达到6.0%。预计可穿戴设备的复合年均增长率高达21.0%，但仅占通信市场份额的10%左右。

数据处理市场的复合年均增长率为2.1%，主要来自服务器和存储设备销售。虽然预计销售额会在2019年同比下降2.8%，但从2020年开始有望回升。尽管我们预计个人电脑市场份额会下降，到2022年的复合年率下降5.2%，但这一降幅将被物联网（IoT）、机器学习以及服务器和数据中心领域其他形式人工智能的增长所抵消。

## 半导体行业各类元件的市场划分

在半导体行业生产的七类元件中，内存芯片元件的市场份额在2022年前将继续占据首位；不过，如前所述，其增长可能在2019年转为负值，然后在2020年实现回升。在整个预测期内，逻辑和微型元件芯片的销售额将持续占据半导体行业总收入的第二大份额（见图1）。

图1 各类元件的市场增长



资料来源：普华永道研究 单位：十亿美元

- **内存**。这一领域的很大一部分增长将由持续的技术进步推动，如云计算技术和智能手机等终端设备上的虚拟现实技术。动态随机存储器（DRAM）和NAND闪存芯片的平均销售价格大幅提高，也在推动收入增长方面发挥显著作用。一般而言，闪存和DRAM的新产能将会抵消预期内的价格下跌，从而更好地实现这类设备的供需平衡，以支持企业固态驱动器（SSD）、增强和虚拟现实、图形、人工智能和其他复杂的实时工作负载功能等新应用。然而，三星在2017至2018年对半导体部门的巨额资本投入将使内存市场产能过剩，尤其是3D NAND闪存市场。

产能过剩将导致市场供应过剩，进而拉低内存元件的市场价格。因此，这类芯片产生的收入将在2019年下降，并对整个半导体市场产生不利影响。

• **逻辑元件**。通信、数据处理和消费类电子产品行业的需求将在很大程度上推动这一市场发展。在预测期内，特殊用途专用集成电路（ASIC）和专用信号处理器（ASSP）逻辑芯片将占据绝大部分市场。

• **微型元件**。这类芯片是所有电子设备的关键组成部分，市场增长将与这些设备的销量成正比。由于标准台式机、笔记本电脑和平板电脑的出货量疲软，微型元件在2019年的增长将停滞不前。2022年之前，微型元件的市场增长来自于汽车行业。汽车制造商正在将大量微型元件集成到智能汽车的动力传动系统、下一代底盘和安全系统中，用于在安全和防撞系统中处理复杂的实时传感器功能。此外，物联网的日益普及也带来对高性能电子产品的需求，从而催生对高性能处理器的需求。本地处理能力通常由微控制器、混合微控制器或微处理器以及集成微控制器设备提供，这些设备可以提供实时嵌入式处理，这是大多数物联网应用的首要要求。

• **模拟元件**。我们预计模拟元件的强劲增长主要受到通信行业的需求推动，尤其是汽车行业。产生需求增长的用例包括电源管理（延长手机电池寿命）、信号转换（用于数据转换器、混合信号设备等）和汽车专用模拟应用（自动驾驶汽车、电动汽车及电子系统）。

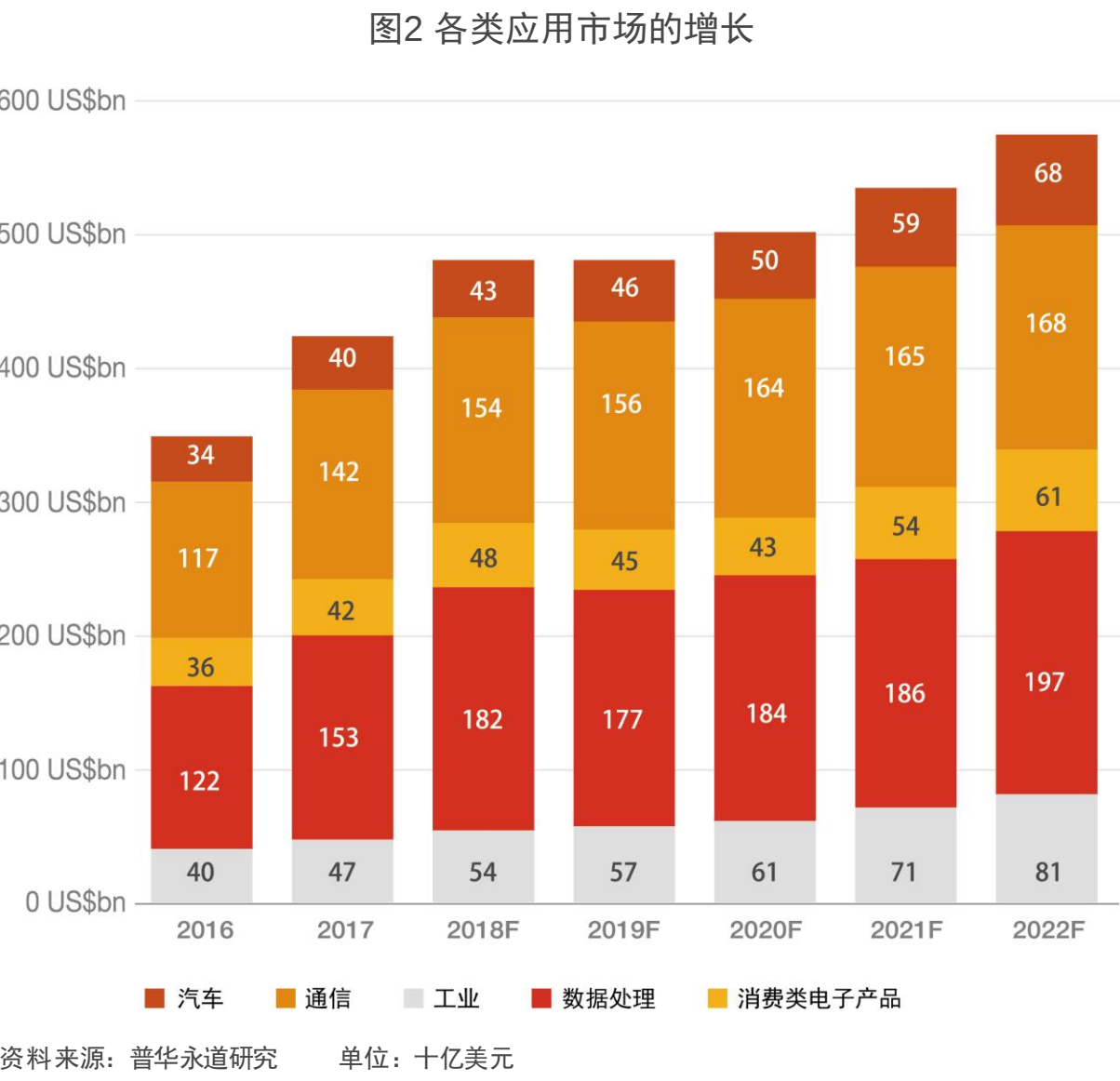
• **光电元件、传感器和分离元件（OCD）**。这三类元件与集成电路相邻。目前大量投产的新兴技术设备将推动这些芯片的需求增长。其中包括固态照明、机器视觉、图像识别、智能电网能源、物联网和智能便携式系统中的“融合”多传感器。

由于光电芯片在嵌入式摄像机的互补金属氧化物半导体（CMOS）图像传感器、汽车安全、固态照明应用的视觉自动化和更高性能的LED中使用越来越广泛，预计其将继续保持强劲的增长势头。总体而言，LED照明解决方案正在迅速改变各种住宅、商业和工业应用的市场。推动其增长的因素包括：采用更节能的照明解决方案、LED价格不断降低、基础设施的现代化改造和新发展。就能效、寿命、多功能性、色彩质量和成本而言，LED照明具备很多超越传统照明技术的优势。

我们预计传感器市场也将实现快速增长。虽然近年来，由于新型自动化控制和物联网应用的单位出货量增加，传感器价格有所下降。随着功率晶体管和其他设备需求的稳定增多，分离元件市场有望得到增长。

半导体应用市场的增长

经济增长预期表明，以汽车和数据处理市场为主导的应用市场将持续扩大（见图2）。





• **汽车**。我们预计在所有市场中，汽车市场增长最快，复合年均增长率将达到11.9%。主要原因是：电动汽车和混合动力汽车的普及率不断提高，其半导体需求量大约为传统汽车的两倍；此外，自动驾驶汽车有着巨大的市场潜力。随着汽车变得更加自动化，每辆汽车对半导体的需求量增加，先进驾驶辅助系统（ADAS）、光探测和测距（LiDAR）、信息娱乐系统以及安全和便利功能由此受到越来越多的关注。据IC Insights统计，每辆全自动驾驶汽车的半导体需求量将是驾驶辅助系统汽车的5倍。然而，传统汽车仍然是半导体销量的重要催化剂。2018年，传统汽车销量占汽车市场总收入近95%。

• **通信**。通信市场对半导体近80%的需求量由手机驱动。虽然目前手机市场高度饱和，但5G的引入、智能手机持续的高更新率以及新兴市场对手机的需求增加，通信市场的复合年均增长率将保持在2.2%。尽管预计高端手机需求会下降，但普通手机的强劲增长将抵消这种影响。

• **消费类电子产品**。得益于智能电视、4K超高清电视、3D编程、视频点播、对大屏显示器的偏爱以及曲面OLED的日益普及，电视设备将促进消费类电子产品应用的半导体收入增长。游戏技术和机顶盒也将成为收入的强力助推器。因此，该市场的复合年均增长率将达到2.2%。尽管可穿戴设备市场仍然相对较小，但其在消费类电子产品应用中增长最快，复合年均增长率将达到6.0%。然而，随着美国Netflix和Amazon Prime等更具吸引力的替代品的流行，数字播放器芯片的收入正在下降，复合年均增长率为2.3%。此外，越来越多的消费者开始转向手机游戏，导致游戏机市场在2018年达到饱和。

• **数据处理**。到2022年，数据处理市场（包括个人电脑、超便携设备、平板电脑、服务器和存储设备）中的半导体销售额将达到2.1%的复合年均增长率。由于终端设备智能功能需要更多的半导体，来自存储设备的市场增长预计非常可观，复合年均增长率为12.3%。这一增长很大程度上来自新兴的固态驱动器技术，该技术克服了传统数据驱动器的长周期、易过热和高能耗等缺点。智能手机和其他连接设备的强劲销售会加速对存储卡和存储设备的需求。由于该市场也存在优化服务器性能的压力，每台设备的半导体需求量将会增加。

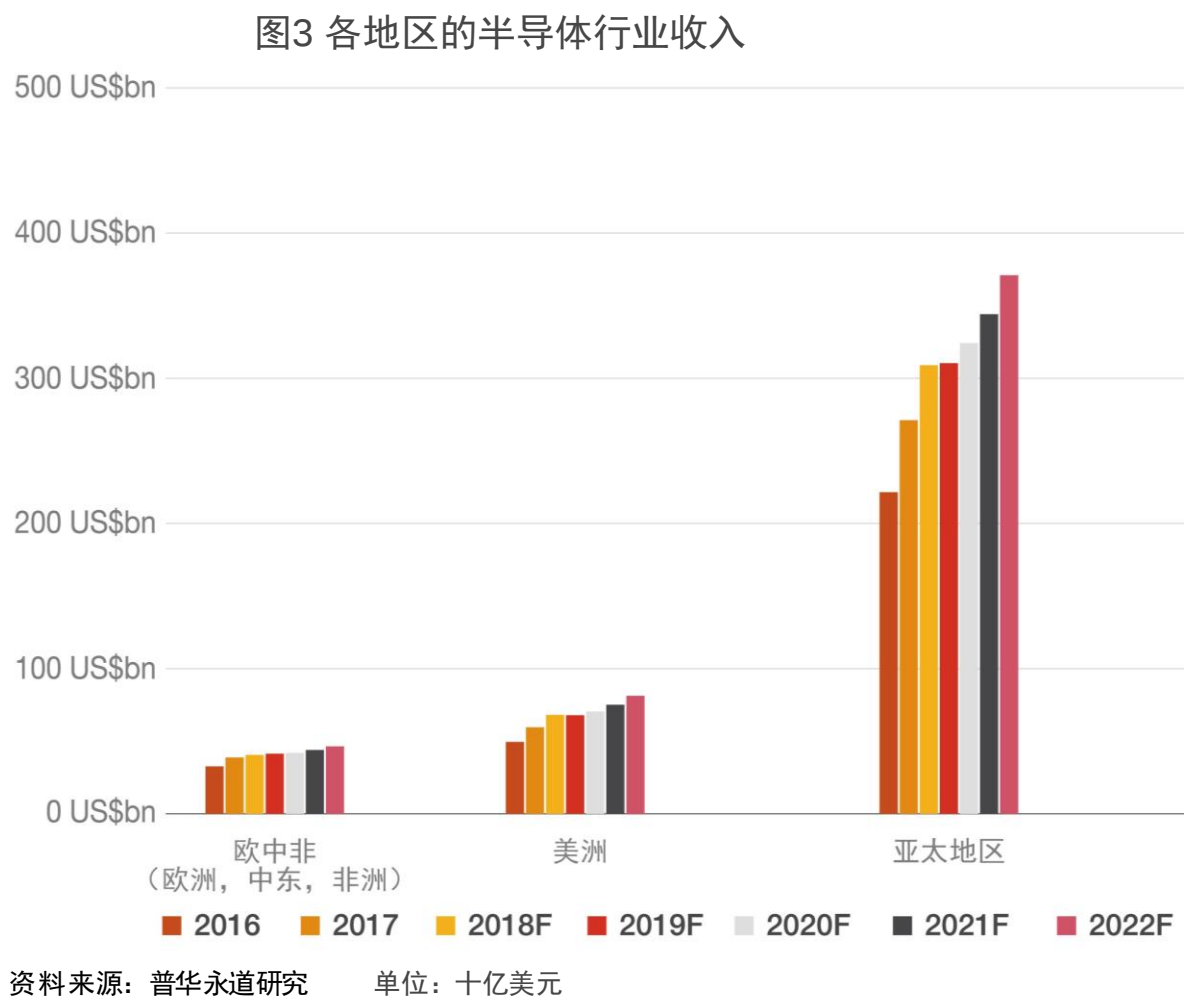
• **工业**。仅次于汽车，工业市场将是所有应用类型中增长最快的市场，预计到2022年其复合年均增长率将达到10.8%。这一增长的最大份额将来自对安全、自动化、固态照明和运输的需求。我们预计安全应用对半导体的需求增长最快，复合年均增长率将达到17.8%。这得益于持续推动更安全和更智能城市的建设，尤其是在亚太地区。对机场和火车站的恐怖袭击事件日益增多，激发了对先进周边安全和门禁控制系统的投资；此外，对舒适性和便利性的日益重视也正在促进指纹门禁系统、PIN和RFID门禁系统的普及。

“在人工智能的背景下，我们需要考虑安全性。一般来说，欧洲企业处于有利地位：一方面包括安全要素，另一方面包括处理环境条件和安全层，就像车辆领域的情况一样。这两个方面都可以有成本合理的认证程序。”  
——Maurice Geraets，恩智浦（NXP）半导体董事会成员

### 半导体行业在各地区的增长

在预测期内，我们预计在所有全球市场中，半导体市场将持续快速增长（见图3）。

• **亚太地区**。该市场将继续是半导体行业收入的主要贡献者，到2022年，复合年均增长率可达到4.8%。电子系统生产将继续以中国为中心。由于中国的制造能力，尤其是消费类电子产品的制造能力出众，因此对半导体的需求日益增长，并成为全球最大的芯片购买国和进口国。发展和增强该行业实力是中国政府的头等大事，因为许多新兴公司正处于初创阶段。



• **欧洲、中东和非洲**。在预测期内，该地区的复合年均增长率将达到3.5%。数据处理一直是欧洲最大的终端用途类别，但我们预计未来两年内将被汽车行业超越。半导体对于许多行业和应用至关重要。欧洲企业在汽车、出行（铁路、航空）和工程等多个领域都处在领先地位。为确保在这些行业的领先局面，并促进人工智能等新应用的发展，欧盟应推广并保护其半导体行业。这包括研发设计、制造以及欧盟创业生态系统。

• **美洲**。在预测期内，该地区的复合年均增长率位居第二，达到4.3%，主要由NAND闪存芯片市场的预期收益推动。在该地区，美国是许多领先半导体公司的所在地，拥有强大的创业生态系统。风险投资是该行业的有力支持因素。然而，美国政府最近否决了若干来自非美国公司的收购计划。



# 学习机器的崛起

半导体行业的需求通常来自颠覆性的新技术推动。在1997年至2007年间，个人电脑的迅速普及推动了对CPU和存储芯片的需求，而互联网的广泛渗透推动了对以太网设备、网络处理器和专用集成电路的需求。智能手机时代始于2007年苹果手机的推出，这增加了对移动处理器的需求，而云计算的采用则推动了服务器CPU和存储的增长。

现在，人工智能很可能成为半导体行业又一个十年增长周期的催化剂。尽管人工智能许多引人注目的新用例将依赖于通过软件而非芯片实现的算法，但对即时计算、连接和传感的需求将会推动未来十年对人工智能定制半导体的巨大需求。

“人工智能提升了学习速度，有助于改善与客户的互动，并能推进流程及程序。”

—— Reinhard Ploss 英飞凌科技（Infineon Technologies）  
首席执行官

## 人工智能与半导体带来的机遇

人工智能是计算机基于对数据集和预定义规则集的复杂分析来模拟智慧人类行为并作出决策或建议的能力。半导体有助于开发和加速人工智能的机会，从而成为推动该领域创新和人工智能增长潜力的关键因素。

人工智能的使用通常取决于三种算法：

- 机器学习（ML）：使用算法分析数据，从中学习，然后对特定情况作出决定或预测。
- 深度学习（DL）：一种基于分析和从特定数据集中学习的机器学习，与特定任务的算法不同。
- 自然语言处理（NLP）：一种分析人机交互的方法，侧重于给计算机编订程序处理和分析大量自然语言数据的方式。

人工智能用例也可按照两种主要的实施类型进行分类：

- 培训系统：利用大量的数据集来学习如何开展特定活动，并不断优化学习算法本身。
- 推理系统：使用预定义的模型进行实时决策。

人工智能适用于几乎所有行业的垂直领域，对云和边缘计算所需芯片数量具有超强预测能力，并且对加速新算法的专业计算需求在不断增长，因而为半导体厂商创造了前所未有的机遇。

“对于人工智能，我们需要在价值链的不同阶段和不同行业之间采用新的合作形式。在新的人机交互方式不断涌现的数字化世界中，安全的身份和新的‘信任授权’均是必需的。为此，可作为道德委员会的国际机构可以发挥重要作用。在欧洲和联合国层面，可将促进关于人工智能道德使用的对话平台作为该机构的基石，进而从人工智能中获益，同时保护个体的利益。”

—— Reinhard Ploss 英飞凌科技首席执行官

## 人工智能驱动的行业增长预测

我们预计到2022年，人工智能相关的半导体市场收入将从目前的60亿美元增至300亿美元以上，复合年均增长率接近50.0%。虽然人工智能驱动的用例会随着时间的推移逐步渗透到每个行业领域，但人工智能的使用将取决于技术投资的规模、技术开发的速度以及实现其效益的速度。

为推理系统提供动力的半导体市场可能仍然是分散的，因为每一个广泛变化的潜在用例，例如面部识别、机器人、工厂自动化、自动驾驶和监控等均需要定制解决方案。相比之下，培训系统将主要基于传统CPU、GPU和现场可编程门阵列（FPGA）基础设施及ASIC。

- **汽车**。仍是市场潜力最大的一个细分市场。我们预计在2022年，ADAS和自动驾驶辅助用例将会带来40亿至47亿美元的收入（见图4）。这其中包括基于推理的系统，用于汽车和边缘计算的自动驾驶和安全辅助；以及基于训练的系统，用于交通规避导航。两者的相对规模将决定需求增长最快的半导体类型——用于边缘计算的GPU和ASIC，以及用于云计算的GPU和FPGA。
- **金融服务**。我们相信，此细分市场将会带来40亿至45亿美元的收入，主要来自交易身份认证和智能投资组合管理的用例。与汽车行业一样，金融服务可能会根据用例实施推理和培训系统。基于认证的用例将在很大程度上依赖于边缘计算的基于推理的人工智能，主要用于智能手机上的面部识别和通过移动CPU或专用人工智能半导体的指纹检测。基于培训的人工智能将主要用于分析海量数据集，以识别智能投资和投资组合管理的趋势；这些活动通常驻留于云端，因为需要基于CPU或GPU基础设施的大量计算。
- **工业**。可能是所有行业中机遇最小的，介于15亿至20亿美元之间，主要来自制造业优化和主动式故障检测。这是因为这些应用非常重视能否利用现有基础架构的培训系统，因此不太可能需要业内最佳的计算能力和更低的延迟。此外，由于工业部署和客户更新周期更长，因此该领域从人工智能获得的收益可能需要比其他行业更长的时间。



汽车行业中的人工智能

汽车行业在电子元件上的开支可谓庞大，在如何利用人工智能加速创新方面进展迅速。到2022年，全球679亿美元的汽车电子元件市场中，人工智能在信息娱乐领域的影响将最为明显，达到85亿美元，ADAS为129亿美元，安全应用57亿美元。

信息娱乐系统将会用于个人辅助、导航和娱乐。苹果的Car-Play和谷歌的安卓汽车平台已在市场上脱颖而出。

ADAS和安全应用将聚焦驾驶员辅助和自动驾驶，主要通过汽车制造商的专有解决方案（如通用汽车巡航自动化解决方案）或广泛可用的平台（如英特尔的Mobileye和辉达驱动）。到2022年，这些人工智能应用的组件将会集中于传感（光电子学和非光学传感器）、计算（ASIC、ASSP、通用逻辑和微组件）及存储（存储器）领域，可用市场达到208亿美元。模拟和分离组件将是整体解决方案的一部分，但并非人工智能应用的主导因素。

在这些组件中，人工智能注入的逻辑元件将设计用于ASSP、ASIC和微组件，而存储器、光电子和非光学传感器将作为辅助组件来支持整个子系统的设计。到2022年，ADAS、安全和信息娱乐中人工智能芯片的市场预计达到40亿至47亿美元，约占这些应用领域总市场的19.2%-22.6%。

人工智能在自动驾驶汽车中的使用将取决于汽车的自动驾驶能力，根据正常操作所需的人工干预量，通常分为五个级别。

0级不涉及自动化。在1级和2级，ADAS提供自动刹车、稳定性控制和巡航控制。3级在某些情况下包括自动驾驶，而在4级和5级，驾驶是完全自动的。

在4级和5级，自动驾驶子系统必须利用其所有组件在通常情况和特殊情况下提供帮助，完全消除对驾驶员甚至方向盘的需求。摄像、雷达和激光雷达传感器必须能够探测并避开物体。信息娱乐模块充当导航、传感器控制和语音命令的主要数据传输源。最后，核心自主平台发挥人工智能推理系统的作用，用于实时计算和作出关键的安全和导航决策。

对于依赖于学习系统的任务，包括实时路线导航、个性化信息娱乐推荐和数字语音辅助，车载连接功能会将请求发送至云端。人工智能定制化的云基础设施将运用人工智能算法优化这些用例，通常由公共云供应商、或由汽车制造商或服务提供商管理的数据中心掌管。



## 人工智能解决方案堆栈

“谁能最大程度实现启发式编程与算法编程融合，谁将成为人工智能的最大受益者。该融合对欧洲公司而言可能是一个很好的机遇，它们在处理启发式和复杂系统方面有着丰富的经验。”  
—— Reinhard Ploss 英飞凌科技首席执行官

新型创新型人工智能产品或服务将会改变已知世界。虚拟助理能发起拟人电话呼叫并在餐馆订座，人们已经对此醉心不已。但为了更好地理解人工智能机遇能够一展拳脚的领域，我们必须更深入地研究人工智能技术堆栈的底层组件，即构建应用程序的脚手架。

在普华永道看来，人工智能技术堆栈由五个元素或层组成：硬件、库、框架和工具、平台及应用与服务（见图5）。由于对人工智能的大部分注意力集中于人工智能带来的客户体验上，所以从应用程序和服务开始讲起是合乎逻辑的。这是解决方案堆栈的最顶层。此处，最可感知的人工智能功能，在应用级别集合在一起，例如亚马逊的Alexa虚拟助理和苹果的人脸识别。其中部分功能也作为服务提供，例如嵌入软件的推荐引擎。

图5 人工智能技术堆栈的元素

| 堆栈元素         | 描述                                                    |
|--------------|-------------------------------------------------------|
| 应用和服务        | 利用人工智能实现“智能”的软件应用，包括视觉处理、聊天机器人客服、智能助手和算法交易。           |
| 人工智能平台       | 现成的架构模块和服务，可提供机器学习、数据分析、NLP、代理、数据解决方案等可用于构建人工智能应用的功能。 |
| 人工智能框架、工具和接口 | 利用底层机器学习算法为特定应用设计、构建和训练深度学习模型的技术。许多是开源技术，并得到广泛支持。     |
| 人工智能库        | 低级软件功能，帮助优化人工智能框架在特定目标硬件上的部署。                         |
| 人工智能硬件       | 处理器单元和半导体逻辑电路，其设计和优化旨在加速人工智能工作负载和计算的执行。               |

资料来源：普华永道研究

但是，没有深层可重用组件来提供核心功能的应用是什么？这其实是平台层的任务。已有若干公司正在生产人工智能平台，承诺无需处理复杂的算法和深层动态神经网络（DNN）便能构建具备人工智能功能的应用。

这些公司已经建立平台，旨在提供“随时能用”的构建模块和软件服务，即基础人工智能功能，如NLP、代理和决策引擎，这有助于加快人工智能应用程序和服务的开发。示例包括：

- **雨鸟技术（Rainbird Technologies）**。雨鸟技术推出以软件即服务为基础的人工智能平台，旨在提高业务运营的智能化。它提供一个以规则为基础的自动化决策引擎，能够支持执行复杂的任务（如作出预测、建议和业务决策）。它还捕捉平台作出某些决策的依据，这对审计十分有价值，特别是在受监管的行业。
- **语义机器（Semantic Machines）**。这家总部位于加州伯克利，最近被微软收购的初创企业开发了一个基于机器学习的基础技术平台，使用户能与信息系统毫不费力地互动。此类称为“对话式人工智能”的方法有望对我们在电子商务网站上的交易方式、与社交媒体的互动方式，甚至日常使用生产力软件和设备的方式产生深远的影响。

堆栈的中间部分（人工智能框架、工具和接口）允许开发者设计、构建和部署实际的模型和算法。独立软件供应商（ISV）正在为开发者提供人工智能框架、工具和接口，以便使用深层人工智能算法为特定用例构建深度学习模型。其中一些框架也是开源的，有利于其得到广泛采用，并得到人工智能生态系统中大多数参与者的大力支持。

堆栈的底部两层由硬件（处理器、逻辑电路和运行人工智能软件的其他组件）及人工智能库组成，这些基本属于低级软件功能，有助于优化底层硅芯片集的人工智能模型和算法。我们期待传统半导体供应商（如英特尔、辉达、高通和赛灵思）提供用于加速此部分堆栈人工智能用例的优化硅芯片。这些公司还可能提供需要的人工智能库，促进其专属架构的开发和逐级采用，进而帮助在其硅产品上部署人工智能框架。人工智能库示例包括：英特尔 DL SDK/Vision SDK、辉达 cuDNN TensorRT和安谋（ARM）NN。

有一点正变得日益清晰：硬件层可以说是这种人工智能解决方案堆栈中最有趣的部分。关键原因有两个：首先，人们日益认识到人工智能要求其深层硬件具备独特的处理能力，这导致了选择最佳处理架构的新一轮竞赛——哪种架构将会胜出，是GPU、数字信号处理器（DSP）、FPGA还是定制ASIC，仍有待观察；其次，开发人工智能硬件的参与者数量日益增加，超过了传统芯片制造商以往的名单，这可能会威胁到老牌供应商，并显著改变其市场地位。



图6 对人工智能机遇的高度期望反映在一系列产品中

| 半导体公司            |      |                                                           |                                                |                                 |                                   |                                        |                     | 其他产品供应商                      |                                           |
|------------------|------|-----------------------------------------------------------|------------------------------------------------|---------------------------------|-----------------------------------|----------------------------------------|---------------------|------------------------------|-------------------------------------------|
|                  |      | 英特尔                                                       | 辉达                                             | 高通                              | 赛灵思                               | 恩智浦                                    | 意法半导体               | 安谋                           | 益华                                        |
| 益华               |      | • 数据中心<br>• 自动驾驶汽车<br>• 增强现实/虚拟现实<br>• 无人机<br>• 监控        | • 数据中心<br>• 自动驾驶汽车<br>• 零售分析<br>• 智慧城市<br>• 监控 | • 计算机视觉                         | • ADAS<br>• 无人机<br>• 数据中心<br>• 医疗 | • ADAS视觉处理                             | • 计算机视觉<br>• 智能驾驶   | • 语音助手<br>• 消费型机器人<br>• 智能手机 | • 汽车<br>• 监控<br>• 无人机<br>• 移动/可穿戴式设备      |
| 人工智能库            |      | MKL<br>DL SDK<br>Vision SDK<br>Movidius<br>Myriad Dev Kit | CUDA DNN<br>TensorRT                           | Snapdragon<br>NPE SDK           | reVISION<br>SDAaccel<br>工具包       | S32 Design<br>Studio IDE<br>Vision SDK | STM32<br>STM32 Cube | ARM NN                       | Tensilica NN<br>mapper toolkit<br>DSP SDK |
| 人工智能处理器、<br>硬件/硅 | CPU  | Xeon PHI                                                  |                                                |                                 |                                   |                                        |                     | Cortex-A75<br>Cortex-A55     |                                           |
|                  | DSP  |                                                           |                                                | Hexagon DSP<br>(Snapdragon NPE) |                                   |                                        |                     |                              | Tensilica<br>Vision C5 DSP                |
|                  | GPU  |                                                           | Pascal (Volta)<br>Maxwell<br>Tesla             |                                 |                                   |                                        |                     |                              |                                           |
|                  | FPGA | Arria 10                                                  |                                                |                                 | Zynq MPSoC                        |                                        |                     |                              |                                           |
|                  | 定制   | Nervana NNP<br>Myriad X<br>Loihi NMP                      | ASIC - TPU                                     |                                 |                                   | S32V<br>视觉处理器                          | DCNN的AL SoC         | ARM ML                       |                                           |

资料来源：普华永道研究

培训

推理

两者皆是

人工智能芯片的新兴战场

人们对人工智能领域半导体市场机遇抱有很高期望的一个明确标志是，每个主要供应商均在提供人工智能硅。最常见的目标应用是ADAS、无人机、监测和计算机视觉。

这些应用架构在选择上差异巨大，包括一般用途CPU、DSP、GPU、FPGA和定制的ASIC等（见图6）。不出所料，大多数供应商对人工智能硅架构的选择与其核心能力或优势领域密切相关。例如，赛灵思的Zynq MPSoC是其FPGA产品的可定制变体，辉达的大多数产品也均基于其核心GP-GPU架构。

另一方面，供应商（如恩智浦和意法半导体）大多提供人工智能特定加速和扩展，以增强其现有产品组合而非人工智能特定芯片的能力。

另一个差异是，IP许可供应商安谋和益华（Cadence）提供软CPU和DSPIP核心，前提是未来人工智能处理器将嵌入至ASIC中，而非由专门用于人工智能工作负载的独立运行芯片处理。软CPU和DSP IP核心的模型使硅供应商获得人工智能软核心的许可，得以开发针对其人工智能应用的芯片。与此同时，财力雄厚的供应商（如英特尔）正在对各种不同的架构（CPU、FPGA和定制ASIC）进行广泛投资，旨在满足不同的处理需求。

我们观察到的另一个区别是公司是否生产专门为培训或推理系统设计的芯片。英特尔和辉达面向培训或推理市场提供的芯片集最为多元。英特尔的Arria 10 FPGA和Myriad X ASIC专为推理工作负载而设计，而其Nervana NNP则最适用于培训。同样，辉达生产的Pascal和Volta芯片适用于培训工作负载，Maxwell则用于推理。

两家公司均制造芯片，英特尔的Loihi NMP和辉达的Tesla，都是为了在各自的应用中实现良好的运转而设计的。我们认为并无哪种方法适用于所有情况；根据待分析的数据源类型、数据重力考虑因素和实时处理需求，每个用例的最佳方案可能各有不同。

定制方案

在这一系列的创新中，一些公司可能会尝试开发定制芯片来传递人工智能“圣杯”，即性能卓越，功耗和成本低于第一代引入的任何标准架构。这场架构之战可能会延续到可以预见的未来，我们认为现在宣布赢家还为时过早。定制芯片的整体性能可能是最佳的，但如果只能处理非常有限的应用程序集和用例，则可能无法实现其经济价值，因为由此导致的较低产量可能无法证明前期开发所付出的成本是必要的。

数家非传统芯片制造商已加入这场“军备竞赛”，争夺人工智能半导体的优势地位，并为其特定人工智能需求而设计的定制芯片试水（见图7）。我们在顶级公有云供应商（特别是亚马逊、谷歌和微软）中看到一种显而易见的趋势——所有这些供应商都在探索定制人工智能芯片作为GPU和FPGA的替代产品，以便在云端产品的性能和成本方面获得竞争优势。亚马逊最近宣布为其边缘计算家用设备Alexa开发人工智能芯片；微软正在为其全息透视镜智能眼镜开发人工智能芯片；2017年，谷歌推出了用于神经网络的张量处理单元（TPU），声称在类似的工作负载下，TPU的性能比CPU/GPU芯片高15至30倍，性能功耗比高30至80倍。



包括苹果、三星和特斯拉在内的几家公司正在开发自己的人工智能硅，并根据其产品要求量身定制。苹果为iPhone XR和iPhone XS智能手机引入A12仿生芯片。该产品包括用于面部识别和动画表情符号应用程序的神经引擎，以及计算摄影和像素处理功能的影像处理器。

鉴于公有云供应商和产品公司纷纷开发自己的定制硅，用于优化自身应用程序和用例，争夺人工智能优势地位竞赛中出现的重大转变必然会威胁到传统芯片制造商（如英特尔、辉达和赛灵思）的市场地位，并有可能颠覆其传统商业模式。尽管 有定制活动，我们预期GPU和FPGA将继续在云端共存，以加速人工智能工作负载。近期，鉴于开发新型硅设计需要大量投资和资源，且这种方式要实现盈利需要达到高销量，大多数应用程序将通过商用硅产品提供服务。

人工智能初创企业的前景

半导体制造商必须抗衡的另一个趋势是庞大的初创企业队伍，这些初创企业正在开发针对人工智能优化的革命性新型芯片架构并将其商业化。关键问题是，这些年轻的公司是否会对现有企业构成威胁，抑或甚至能够把握机遇超越其他竞争对手并在人工智能领域胜出？

利用人工智能前景的热潮正在引发解决方案堆栈领域的大量创新。近年来，对人工智能初创企业的风投基金大幅增加，2017年对人工智能和机器学习公司的投资达到创纪录的110亿美元。不出所料，这种行动有很大一部分发生在堆栈上层的软件和算法领域。初创企业在该领域构建专注于特定人工智能用例的可扩展平台，并寻求开发能够集成到现有应用程序中的人工智能软件，以使其更智能化。

堆栈底层同样能引起人们巨大的兴趣和兴奋之情，越来越多的初创企业在瞄准新型硅架构，将其优化来满足人工智能工作负载带来的独特处理需求。我们对风投基金的分析显示，人们对半导体初创企业的兴趣有所回升，2017年，半导体初创企业吸引了近7.5亿美元的风投资金，是前两年所获资金总和的三倍多，也是2015年前所有投入人工智能芯片初创企业的资金的12倍。

如图8所示，前19家人工智能半导体初创企业中有11家位于美国，且大多数均在探索针对各种人工智能和深度学习工作负载专门定制的处理 器架构。其中9家正在构建深度学习处理器，三家正在开发所谓的神经形态处理器，该处理器基于超前的全新架构，试图模仿人脑的运行方式。

图7 加入人工智能竞赛的其他参与者



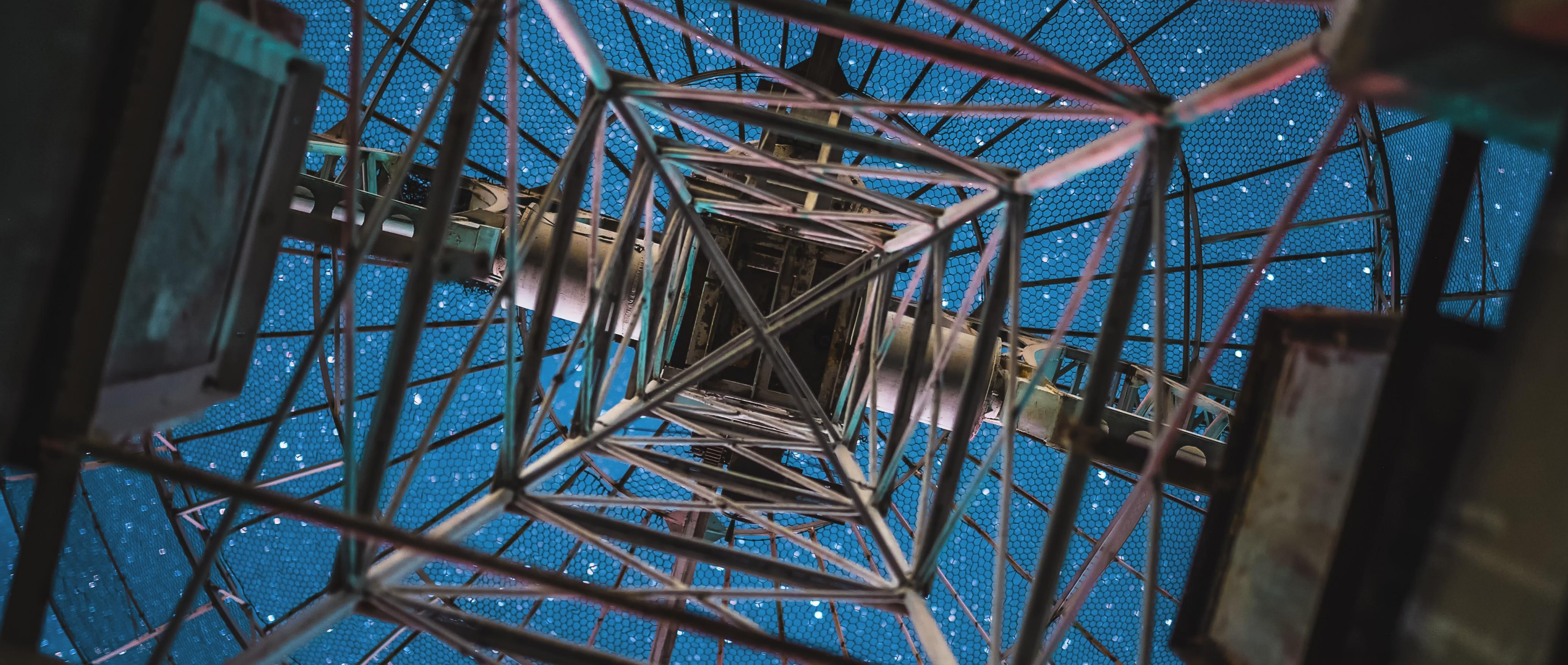


图8 人工智能初创企业获得的风投资金正在增加

| 初创企业                         | 成立年份 | 总部          | 阶段  | 迄今为止获得的资金<br>(百万美元)          | 战略投资者                    | 技术                       |
|------------------------------|------|-------------|-----|------------------------------|--------------------------|--------------------------|
| 地平线机器人 (Horizon Robotics)    | 2015 | 中国北京        | B轮  | <div><div></div></div> 700   | 海力士半导体 (SK Hynix), SK 中国 | 成像DSP                    |
| Graphcore                    | 2016 | 英国布里斯托尔     | D轮  | <div><div></div></div> 310   | 微软、宝马                    | 深度学习处理器                  |
| 寒武纪科技 (Cambricon Technology) | 2016 | 中国北京        | B轮  | <div><div></div></div> 210   | 国投 (SDIC)                | 深度学习处理器                  |
| Wave Computing               | 2010 | 加州坎贝尔       | E轮  | <div><div></div></div> 203.3 | 三星                       | 深度学习处理器                  |
| Vicarious                    | 2010 | 加州旧金山       | C轮  | <div><div></div></div> 120   | 三星                       | 神经形态处理器                  |
| Rigetti Computing            | 2013 | 加州伯克利       | B轮  | <div><div></div></div> 119   |                          | 光学/量子人工智能计算              |
| Cerebras                     | 2016 | 加州洛斯阿尔托斯    | C轮  | <div><div></div></div> 112   |                          | 深度学习处理器                  |
| Vayyar                       | 2011 | 以色列耶胡德      | C轮  | <div><div></div></div> 80    |                          | 成像DSP                    |
| 熠知电子 (ThinkForce)            | 2017 | 中国上海        | A轮  | <div><div></div></div> 68    |                          | 人工智能加速引擎                 |
| Movidius                     | 2006 | 加州圣马特奥      | E轮  | <div><div></div></div> 56    | 英特尔 (收购)                 | 神经计算机引擎加速器 (应用程序: 成像DSP) |
| Mythic                       | 2012 | 加州红木城、德州奥斯汀 | B轮  | <div><div></div></div> 55.2  |                          | 神经形态处理器                  |
| KnuEdge                      | 2005 | 加州圣地亚哥      | 种子轮 | <div><div></div></div> 47    |                          | 神经形态处理器                  |
| Nervana                      | 2014 | 加州圣地亚哥      | A轮  | <div><div></div></div> 40    | 英特尔 (收购)                 | 深度学习处理器                  |
| Xanadu                       | 2016 | 加拿大多伦多      | 种子轮 | <div><div></div></div> 25    |                          | 光学/量子人工智能计算              |
| Reduced Energy Microsystems  | 2014 | 加州旧金山       | 种子轮 | <div><div></div></div> 6.9   |                          | 深度学习处理器                  |
| LightOn                      | 2016 | 法国巴黎        | 种子轮 | <div><div></div></div> 3.3   |                          | 光学/量子人工智能计算              |
| CyberSwarm                   | 2017 | 加州圣马特奥      | 种子轮 | <div><div></div></div> 2     |                          | 人工智能辅助网络安全CPU            |
| Tenstorrent                  | 2016 | 加拿大多伦多      | 种子轮 | <div><div></div></div> 1     |                          | 深度学习处理器                  |
| Vathys                       | 2015 | 俄勒冈州波特兰     | 种子轮 |                              |                          | 深度学习处理器                  |

A轮 B轮 C轮 D轮 E轮 种子期

资料来源：普华永道研究

值得注意的是，其中只有少数公司拥有战略投资者，最活跃的战略投资者为英特尔和三星。事实上，英特尔已收购Movidius和Nervana，并已开始将这两家公司的产品集成到自己的人工智能布局图中。值得注意的是，辉达缺席战略投资者名单，该公司反而在堆栈上层利用风险部门向正在构建平台和应用程序的公司进行进一步投资。

迄今为止，人工智能初创企业获得的风投总额在早期阶段的A轮和B轮投资以及后期阶段的C轮和D轮交易平均分布。尽管大多数后期阶段的初创企业均位于美国和欧洲，但大部分早期阶段的企业位于亚太地区。

中国最大的初创企业包括寒武纪科技公司、地平线机器人公司、熠知电子和深鉴科技公司（DeePhi Technologies），迄今为止它们共筹集3亿美元的风险投资，前两家公司约占投资总额的三分之二。

据迄今为止的分析，我们相信制造最好的人工智能硅材料的竞赛才刚刚开始，竞争将十分激烈，难以预料未来会对现有企业造成什么样的影响。每一家希望在这一竞争激烈的领域大展身手的半导体公司都必须从现在开始着手准备，然而只有时间才能告诉我们谁是最后的赢家。



抓住人工智能机遇

历史表明，尽管半导体公司从颠覆性增长周期中获利颇丰，但无论是将超出芯片本身的新技术货币化，还是对这些技术支持的新商业模式进行扩张，它们仍有很多次未能获得每个周期全部价值的应得份额。

几乎可以确定的是，人工智能的崛起会成为未来十年半导体行业最强大的驱动力。正如我们的分析显示，现有企业和初创企业都在努力开发驱动人工智能的硬件。但是它们能否充分把握住这个机遇？它们能否跳出从开发和销售人工智能芯片中获得利润的模式，转而全身心参与到人工智能革命中？

我们相信他们可以，但若要做到这一点，他们必须深思熟虑以重新评估其人工智能战略和商业模式，精心设计其技术和产品战略，并深入了解如何在整个人工智能生态系统中发挥其作用。以下是公司在准备应对人工智能带来的变化并充分利用这一机遇时需要考虑的若干关键因素和建议。

**战略和商业模式。**制定公司的人工智能愿景，然后利用这一愿景更好地理解需要关注的核心领域，至少在竞赛的早期阶段应当如此。这主要包括探索新的方法将公司资产和专业领域货币化。能否利用数据？能否提供相关服务？该方法还将告诉我们如何更好地发展当前的投资组合，使其变得与人工智能更相关、更适用，同时与总体战略保持紧密一致。以下是公司可以选择聚焦的三个领域：

- 增长细分市场：在利用公司现有市场并提供巨大增长潜力的细分市场中，识别并瞄准新型人工智能用例，例如ADAS和物联网。
- 硅之外的货币化：探索人工智能特定知识产权许可机遇或提供可货币化的人工智能相关服务的机遇，包括托管人工智能服务化和匿名人工智能用例数据，用于改进培训系统和算法。
- 产品组合：仔细评估在何处下大赌注来构建新型人工智能功能，而非进行增量投资来增强当前投资组合中人工智能的适用性。

**技术和产品供应。**鉴于芯片的具体用例，为芯片选择正确的技术和架构至关重要，但公司也必须确保所选择的技术路径在人工智能堆栈的其他地方能够得到支持。一种行之有效的方法是，定义产品供应，以广泛地包含人工智能库、工具包和堆栈中的其他软件元素，但是公司必须确定究竟是自己构建所有元素，还是将产品与技术合作伙伴的元素整合。公司可以通过以下几种方式侧重于产品：

- 不断进化的硬件架构：探索具有定制架构的设计，例如神经形态处理和子系统设计，这些都是为了通过独特的自学功能来加速深度学习算法而定制的，例如通过集成逻辑和内存功能。
- 人工智能库和工具包：尽可能为产品组合中的现有产品开发软件开发工具包（SDK）和编译器，以优化和加速人工智能算法。
- 全栈产品：与合作伙伴合作，在硅、平台、工具和人工智能库方面提供全栈解决方案，使应用程序的开发和差异化轻松易行。

**合作伙伴关系和人工智能生态系统。**对上述人工智能解决方案堆栈的深入研究清晰地表明，人工智能的成功在很大程度上取决于建立一个完整的合作生态系统，无论是通过技术联盟，还是作为企业将产品推向市场。为推动长期增长，聪明的参与者应识别并有效利用生态系统来缩短上市时间，制定高效的销售策略以进军特定的垂直行业或应用，并评估战略合作伙伴关系和投资，如合资还是收购。以下是一些可供公司探索的选项：

- 许可：通过授予来自第三方供应商，如安谋、思华（CEVA）和铿腾（Cadence）的知识产权核心的许可，探索既可降低风险又能加快人工智能专用产品开发的选项。
- 合作伙伴：与整个堆栈中的竞赛者建立新的战略伙伴关系，推动自身人工智能芯片和解决方案的采用。
- 战略投资：考虑投资于人工智能初创企业，增强与核心业务相一致的用例功能。



# 行业数字化

半导体行业自问世以来一直是数字化的先驱，提供数字化服务并追求新的数字商业模式。例如，在上世纪70年代，英特尔出售芯片测试设备的收入即超过其出售芯片本身的收入。上世纪80年代，随着数字设计和仿真工具以及通信技术的普及，无晶圆厂和铸造模型涌现出来并颠覆了当时盛行的整合元件制造商（IDM）模式。后来出现了纯粹的知识产权竞争，高通、安谋等公司纷纷效仿，进一步侵蚀了半导体价值链。

如今，其他行业，尤其是汽车行业在数字化方面明显超过了半导体行业。这不免令人啼笑皆非，因为汽车制造商自身在数字化方面的成功很大程度上来自于半导体行业产品的支持。对半导体公司而言，现在比以往任何时候都更需要考虑如何最好地利用数字化，以及找到最有利于其组织发展的机遇。

在考虑数字化的最佳推进方式时，芯片制造商可以衡量三大战略（数字化产品和服务、数字商业模式和数字半导体价值链）。请注意，所有这些策略都必须得到可靠数据和分析技术的支持。

**数字化产品和服务。**我们看到半导体公司在实现产品和服务数字化时有三个选择：数据货币化、增强和定制（见图9）。数据货币化允许半导体公司利用自己或他人设备产生的大量数据；增强后产品和服务通过人工智能或整合其他产品或服务，来丰富公司的现有技术；定制化通过更高的精度和效率为客户带来额外的价值。

图9 半导体公司的数字战略

|             | 数据货币化                                                                                    |                                                                         | 增强                                                                            |                                                                                                            |                                                                                               | 定制                                                                                                      |
|-------------|------------------------------------------------------------------------------------------|-------------------------------------------------------------------------|-------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------|
| 应用          | <ul style="list-style-type: none"><li>提供原始数据</li><li>“清洗”和组织数据</li><li>以数据生成洞见</li></ul> | <ul style="list-style-type: none"><li>人工智能辅助的集成电路（IC）设计服务</li></ul>     | <ul style="list-style-type: none"><li>人工智能辅助IC集成服务</li></ul>                  | <ul style="list-style-type: none"><li>物联网平台集成</li></ul>                                                    | <ul style="list-style-type: none"><li>完整的设备</li></ul>                                         | <ul style="list-style-type: none"><li>按需供应的芯片</li></ul>                                                 |
| 描述          | <ul style="list-style-type: none"><li>为客户提供的IC设计数据</li><li>为客户提供数据清洗、生成洞见的算法</li></ul>   | <ul style="list-style-type: none"><li>为客户提供基于人工智能的设计服务和错误修复服务</li></ul> | <ul style="list-style-type: none"><li>基于人工智能的电路设计服务主要围绕供应商的IC产品基础工作</li></ul> | <ul style="list-style-type: none"><li>芯片供应商与超大规模云端平台供应商合作，确保IC产品编程工具在云端平台的集成开发环境中可用，并能通过应用编程接口启用</li></ul> | <ul style="list-style-type: none"><li>通过掌握数据，更好地捉住物联网机遇</li><li>越接近客户越能实现更成功的特定应用设计</li></ul> | <ul style="list-style-type: none"><li>展望未来，用于芯片设计的人工智能应用的激增，将允许客户通过人工智能引擎创建的设计选项来设置IC的性能和功能参数</li></ul> |
| 示例/<br>潜在用例 | <ul style="list-style-type: none"><li>三星Artik云物联网平台</li></ul>                            | <ul style="list-style-type: none"><li>英特尔EDA设计流程（芯片）*</li></ul>         | <ul style="list-style-type: none"><li>英特尔EDA设计流程（板级设计）*</li></ul>             | <ul style="list-style-type: none"><li>高通/Linaro*</li><li>云端服务</li></ul>                                    | <ul style="list-style-type: none"><li>英特尔Mobileye</li></ul>                                   | <ul style="list-style-type: none"><li>英特尔（实验阶段）</li></ul>                                               |

\*潜在用例来源：普华永道调查



**数字商业模式。**我们看到半导体公司可以选择五种创新的数字商业模式（见图10）：

1. **剃刀与刀片**：企业可以利用该模式，以较低的利润率提供核心产品（如计算架构），然后开发一种附加产品（如云平台服务）。此类产品依赖于该架构，但利润更高。未来，半导体供应商可以向云端服务供应商出租设备，从而实现硬件利用的货币化。
2. **平台**：公司可以通过促进芯片制造商和客户之间基于数据或硬件的交换创造价值。这有利于平台参与者，并允许创办人影响标准，增加竞争对手的转换成本。

3. **开源**：利用这一机遇，半导体公司可以创建一个平台，允许客户构建定制化的开源芯片。这有助于与第三方共享软件源代码和集成电路设计，从而分摊研发成本并缩短上市时间。
4. **XaaS（一切即服务）**：半导体制造商可以使用该方法，围绕基础设施、硬件和软件开发来创新服务，例如提供计算即服务等。由此，硬件和功能更新可以作为服务计费。
5. **市场**：发展双边市场，使用人数的增多可以产生积极的网络效应，从而增加价值。市场可以包括基于云的算法即服务，例如共享人工智能培训数据等。客户可以通过通用接口访问算法，开发者则可以上传新的算法和模型来丰富市场。

图10 半导体公司的创新商业模式

| 类型    | 应用          | 描述                                                                                                 | 示例                                                |
|-------|-------------|----------------------------------------------------------------------------------------------------|---------------------------------------------------|
| 剃刀与刀片 | • 云平台与新型计算  | • 提供领先的计算架构支持的云平台服务<br>• 支持应用程序特有的计算架构（如深度学习）<br>• 未来，芯片供应商可以向云服务供应商出租设备，从而实现硬件使用的货币化              | 谷歌云平台<br><br>辉达GPU云                               |
|       | • 数据平台      | • 为合作伙伴创建数据交换和货币化平台，使平台参与者受益，并允许创办人设立和改变标准，增加竞争对手的转换成本                                             | TSMC开放创新平台                                        |
| 平台    | • 平台软件基础设施  | • 作为人工智能硅芯片知识产权的一部分，开发标准的应用程序API，允许第三方开发者在芯片上创建和部署人工智能模型<br>• 创建标准工具包、编译器和/或库，为特定应用程序中的芯片建立“开发者粘性” | 思华和辉达深度学习SDK<br><br>安谋NN SDK<br>TensorFlow库       |
|       | • 物联网平台     | • 支持集成电路产品与外部供应商增值服务的互操作性，使得硬件即服务等商业模式成为可能                                                         | 安谋和Linaro开源ARM平台<br>英特尔物联网平台                      |
| 开源    | • 开源集成电路设计  | • 创建一个平台，允许客户构建定制化开源硅芯片<br>• 与传统设计相比，定制服务可以快速整合并以更快、更低成本向客户交付                                      | SiFive开源设计                                        |
| XaaS  | • 计算即服务     | • 硬件和功能更新作为服务计费<br>• 服务以动态方式重新编程现场可编程门阵列，新型和改进的推理模型被下载至芯片上<br>• 芯片功能即服务（收费）                        | IBM Q（量子计算）                                       |
| 市场    | • 人工智能算法/模型 | • 为基于云端的算法即服务创建市场<br>• 客户可以通过通用接口访问算法<br>• 开发者可以上传新的算法/模型来丰富市场                                     | Algorithmia人工智能算法市场                               |
|       | • 终端用户应用数据  | • 提供服务：帮助整合不同客户和用例的人工智能相关数据，将其匿名化并在人工智能数据市场上销售<br>• 这一数据的潜在购买者可以获得反馈以改进训练/算法，并了解终端客户如何使用人工智能       | Bottos分布式人工智能数据市场<br><br>OpenMind和Ocean人工智能数据共享市场 |

资料来源：普华永道研究



**数字半导体价值链。**半导体企业可以通过对端到端的纵向和横向价值链进行数字化来获得巨大利润，不仅可以利用新的人工智能驱动的能力，还可以充分利用它们提供的其他数字机会。图11显示公司应该考虑采取的若干措施。

总而言之，半导体公司应考虑通过所有三种数字战略所能获得的各种机会（见图12）。

**运营。**除了可以用来提高收入的各种数字战略外，芯片运营商还可以考虑通过应用人工智能和机器学习改善运营来提高收益的机会。选项包括：

- **中央控制塔：**可以让公司始终实时详细了解所有供应链环节和运营（包括供应商和客户的运营）。公司可以设计用于虚拟控制室操作的仪表盘，并为关键利益相关者提供可视化。
- **生态系统感知：**使用人工智能从供应商和客户生态系统收集数据和见解，识别最相关的信号，从而指明下一步的机会，并为采取的应对措施提供建议。
- **人工智能辅助的长期/短期需求预测：**收集来自生态系统的需求信号，如大型企业采取的行动、来自供应链的信号和相关新闻等，并分析其对改善实时生产组合和供应调度的影响。

- **人工智能辅助的设计和调试：**使用机器学习系统为集成电路设计解决方案提供建议，识别潜在的错误设计元素，并实施更高效的产品设计分支。
- **制造工艺优化：**将工厂控制中心设在远程设施或办公空间旁边，以提高学习和响应能力，从而通过分析生产设备的传感器日志和相关事件来提高设备效率。此外，还可以使用计算机视觉工具来发现故障群。人工智能也可以支持释放在前端和后端的晶圆批次。

半导体公司如果希望受益于进一步的数字化和人工智能发展机遇，就应通过以下经过深思熟虑的步骤有系统地前进：第一步，探索和学习。公司必须选择合适的机会，通过试点项目进一步理解和开发新的用例和产品。第二步，建立能力。公司应识别现有和所需要的能力，然后建立缺失的能力。随后需设计并推出数字转型计划，以实现新的能力。第三步，扩大规模，进一步发现数字机会和相关用例。这之后，公司必须确定将这些机会融入到新数字组织的路线图。企业也必须利用新的功能集和组织来确保新产品开发和引进的顺利实施。

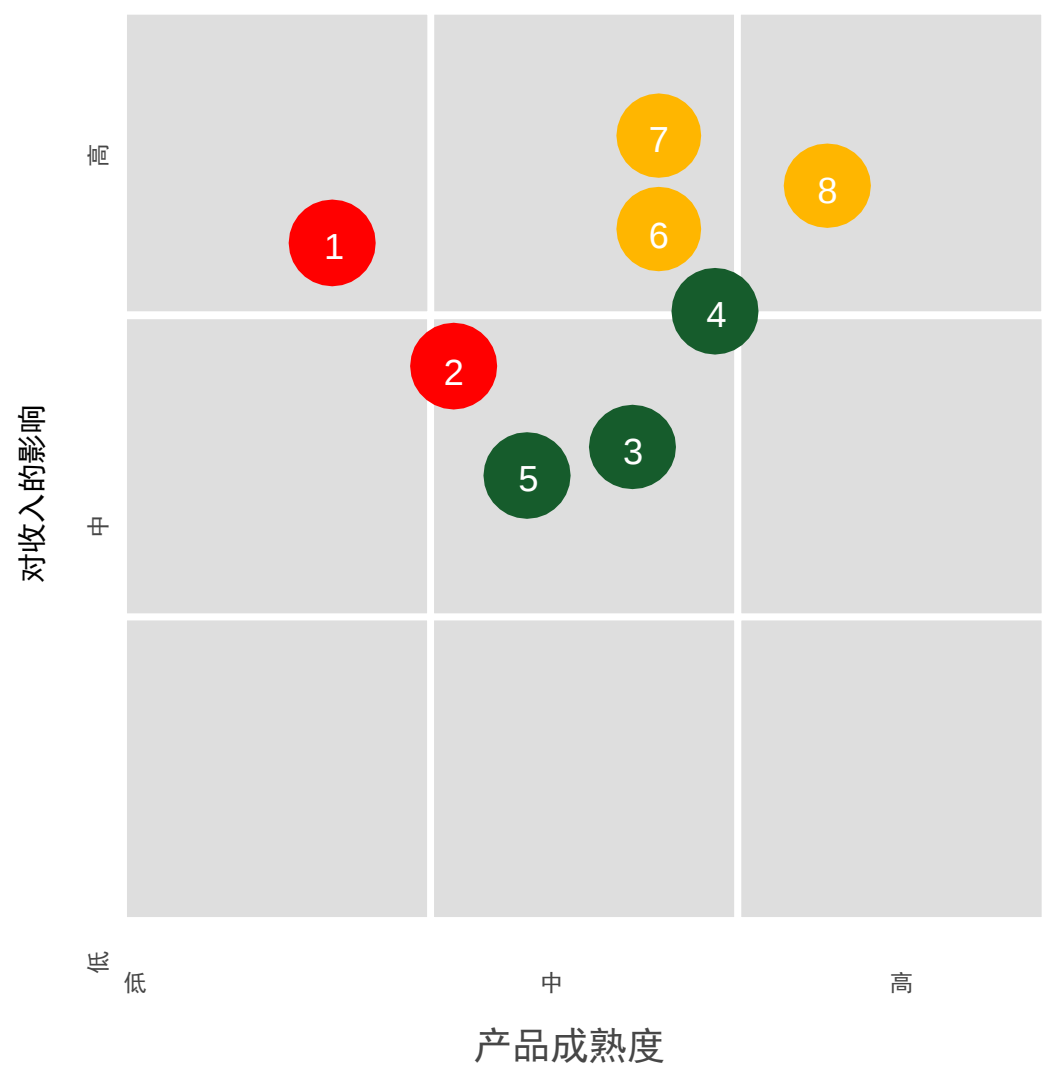
图11 价值链中的各种机会



资料来源：普华永道研究



图12 按成熟度和影响划分的机会



### 产品和服务的数字化

- 1 数据货币化——人工智能辅助的集成电路设计服务
- 2 产品定制——按需芯片
  - 数据货币化——处理服务
  - 增强型产品——人工智能辅助的集成电路集成服务
  - 增强型产品——物联网平台集成
  - 增强型产品——设备运行

### 创新的数字商业模式

- 3 开源——集成电路设计
- 4 一切即服务——计算即服务
- 5 市场——人工智能算法/模型
  - 剃刀与刀片——云平台 and 新型计算
  - 平台——数据
  - 平台——物联网
  - 市场——人工智能终端用户/应用数据

### 价值链数字化

- 6 生态系统感知
- 7 人工智能辅助的设计和调试
- 8 制造工艺优化
  - 中央控制塔
  - 人工智能辅助的长期/短期需求预测

# 未来道路

显然，未来几年甚至几十年，半导体公司获利概率仍然较高。在截止到2022年的预测期内，我们预计在全球所有市场中，半导体市场将持续快速增长，达到5,750亿美元。

到2022年，七类元件中的内存芯片将继续占据最大的市场份额，大部分增长由云计算和智能手机等终端设备的虚拟现实所推动。

此外，全球经济的乐观前景表明，到2022年，以汽车和数据处理市场为主导的应用市场或将继续增长。带动这些细分市场的将是人工智能相关芯片的需求。

能够最大限度利用这一增长并充分实现其市场潜力的半导体公司很可能会是那些能够把握人工智能机遇的公司。随着新兴初创公司和科技界其他领域参与者加入竞争，争夺市场的竞争只会日益激烈。除了提供芯片之外，半导体公司还必须找到合适的方法，实现新技术的货币化以超越实际芯片本身，或者拓展由这些技术支持的新商业模式。采取此类行动的公司将会繁荣发展，反之则会被更敏捷的竞争对手超越。



# 联系信息

## 杜塞尔多夫

Werner Ballhaus  
普华永道德国合伙人  
+49-211-981-5848  
werner.ballhaus@pwc.com

Seung-Min Shin  
普华永道德国经理  
+49-211-981-1228  
seung.min.shin@pwc.com

## 慕尼黑

Marcus Gloger  
普华永道思略特德国合伙人  
+49-89-5452-5572  
marcus.gloger@pwc.com

Tanjeff Schadt  
普华永道思略特德国总监  
+49-89-5452-5521  
t.schadt@pwc.com

Sören Tolzmann  
普华永道德国总监  
+49-89-5790-5202  
soeren.tolzmann@pwc.com

Andris Ogrins  
普华永道思略特德国高级经理  
+49-89-5452-5632  
andris.ogrins@pwc.com

## 圣何塞

Conall Dempsey  
普华永道美国合伙人  
+1-215-518-0149  
conall.dempsey@pwc.com

Mark McCaffrey  
普华永道美国合伙人  
+1-408-817-4199  
mark.mccaffrey@pwc.com

Anand Sundaram  
普华永道美国总监  
+1-480-319-1931  
anand.sundaram@pwc.com

Amit Adalti  
普华永道思略特美国总监  
+1-408-624-6714  
amit.adalti@pwc.com

## 深圳

周伟然（Wilson Chow）  
普华永道全球TMT行业主管合伙人  
+[86] (755) 8261 8886  
wilson.wy.chow@cn.pwc.com

## 上海

高建斌（Jianbin Gao）  
普华永道中国TMT行业主管合伙人  
+86-21-2323-3362  
gao.jianbin@cn.pwc.com



普华永道

---

在普华永道，我们的使命是建立社会信任以及解决重要问题。我们的企业网络遍布158个国家，拥有超过250,000名员工，致力于提供高质量的审计、咨询和税务服务。敬请访问[www.pwc.com](http://www.pwc.com)了解更多信息，或告知我们您的需求。

本文内容仅作为一般信息用途，不应取代专业顾问的咨询。

© 2019 普华永道。保留所有权利。普华永道是指普华永道网络和/或其一个或多个成员公司，各成员公司均为独立的法律实体。

更多详情请参见[www.pwc.com/structure](http://www.pwc.com/structure)。PMS-000106-1