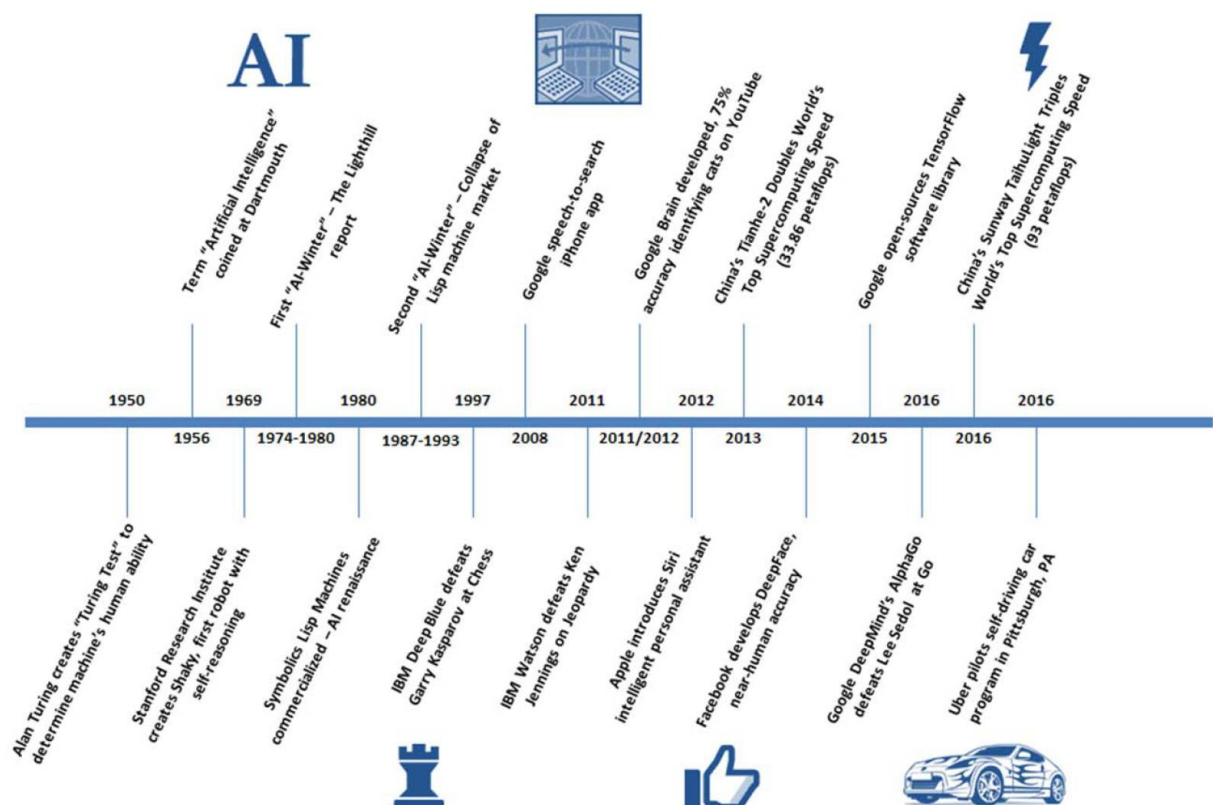


概要

人工智能是信息时代的尖端技术。从人类建立起需要指导控制才干运营的计算机，到计算机拥有可以自己去学习的能力，这一奔腾对各行各业都产生了巨大的影响。虽然此时此刻也许是下一个 AI 冬季（图 8）到来之前的「给予承诺又让人失望」的周期，但这些投资和新技术至少会给我们带来有形的机器学习生产力的经济利益。

Exhibit 8: Evolution of AI: 1950-Present



与此同时，人工智能、机器人和无人驾驶汽车已经成为了流行文化甚至是政治话语的前沿。并且我们在过去一年的研究使我们相信这不是一个错误的开始，而是一个拐点。正如我们将在本报告中探讨的那样，这个变化的因素有显而易见的（更快更强的计算资源和爆炸式增长的数据库），也有细致入微（深度学习，专有硬件和开源的崛起）的。

这个 **AI 拐点 (AI inflection)** 中更令人兴奋的一个方面是「现实世界」的使用案例比比皆是。虽然深度学习使计算机视觉和自然语言解决等技术有了显著的提高，比如苹果公司的 **Siri**，亚马逊的 **Alexa** 和 **Google** 的图像辨认，但是 **AI** 不仅仅是「科技技术」(**tech for tech**)，也就是大数据集与足够强大的技术相结合的情况下，价值正在被慢慢创建，竞争优势也变得越来越明显。

例如，在医疗保健中，图像辨认技术可以提高癌症诊断的准确性。在农业中，农民和种子生产商可以运用深度学习技术来提高作物产量。在制药业中，深度学习可以用于改善药物的研发。在能源方面，勘探效率正在提高，设备可用性正在不断增强。在金融服务方面，通过开辟新的数据集，实现更快的分析，从而减少成本，提高回报。**AI** 现在还处在发现其可被运用场景的初期阶段，这些必要的技术会通过基于云的服务实现大众化、平等化，我们相信随之而来的创新浪潮将在每个行业中发明新的赢家和输家。

AI 的广泛应用让我们得出了一个结论：它是一种可以变革全球经济的技术，是提高生产力并结束美国生产率停滞增长的驱动力。结合 GS 首席经济学家 Jan Hatzius 的研究，我们明确了资本深化目前的停滞及其对美国生产率的相关影响。我们相信，AI 技术将会驱动生产力的提高，就像 20 世纪 90 年代那样，驱动公司投资更多的资本和劳动密集型项目，加快发展的脚步，提高赚钱能力以及提高股票的估值。

启示

虽然我们看到了人工智能可以及时地影响到每个公司、行业的一部分经济，但对投资者而言，我们认为这其中四个影响最为显著。

生产率。AI 和机器学习具有激发生产率增长周期的潜力，这会有助于经济的增长，提高公司的赚钱能力，资本回报率和资产估值。根据 GS 首席经济学家 Jan Hatzius 所说：

「大体上而言，AI 看起来似乎比上一次创新浪潮更有也许在记录数据中捕获到更有价值的东西，人工智能可以减少成本，减少对高附加值生产类型的劳动投入。举个例子，这些在商业部门成本节约上的创新也许比在 iPhone 中增长应用程序的可用性和多用性更利于记录学家去捕获有价值的东西。考虑人工智能对商业部门的成本结构的广泛影响，我有理由相信它会被记录学家接受，并且会出现在整体生产力数据中。

尖端技术。**AI** 和机器学习在速度上的价值有助于构建一种在建设数据中心和网络服务时让硬件更便宜的趋势。我们认为这也许推动硬件，软件和服务支出的市场份额的大幅度改变。例如，在「标准」数据中心计算资源上运营的 **AWS** 工作负载的成本低至 **\$ 0.0065 /小时**，而在使用 **AI** 优化过的 **GPU** 上运营的成本为 **0.900 美元一小时**。

竞争优势。我们看到了 **AI** 和机器学习具有重新调整每个行业的竞争秩序的潜力。未能投资和运用这些技术的管理团队在和受益于战略智能的公司竞争时，有很大也许会被淘汰掉，由于这些技术可以让公司的生产力提高，并为它们发明资本效益。在第 **41** 页开始的短文中，我们将研究这些竞争优势是如何在医疗保健、能源、零售、金融和农业等领域发展起来的。

创办新公司。我们发现了 **150** 多家在过去十年中创建的人工智能和机器学习公司（附录 **69-75**）。虽然我们相信人工智能的大部分价值都掌握在具有资源、数据和投资能力的大公司手中，但我们也盼望风险投资家、公司家和技术专家可以继续推动新公司的创建，从而促进实质性的创新和价值发明，即使最后创业公司会被收购。当然我们也不能忽视人工智能巨头（人工智能领域的谷歌或 **Facebook**）的出现。

在接下来的篇幅中，我们将进一步探讨 AI 的技术，历史，机器学习的生态系统以及这些技术在行业和领头公司中的应用。

什么是人工智能？

人工智能是做出可以以人类智能的方式学习并解决问题的智能机器和计算机程序的理工科。传统而言，该领域涉及自然语言解决与翻译、视觉感知与模式辨认，以及决策制定。但该领域以及应用的复杂度都在急剧扩展。

在此报告中，我们的大部分分析集中在机器学习（人工智能的一个分支）与深度学习（机器学习的分支）上。我们强调两点：

简言之，机器学习是从样本和经验（即数据集）中进行学习的算法，而不是依靠硬编码和预先定义的规则。换言之，也就是开发者不再告诉程序如何区分苹果和橘子，而是向算法输入数据（训练），然后自己学习如何区分苹果和橘子。

深度学习的重大发展是人工智能拐点背后的重要驱动。深度学习是机器学习的一个子集。在大部分传统的机器学习方法中，特性（即有预测性的输入或属性）由人来设计。特性工程是一大瓶颈，需要大量的专业知识。在无监督学习中，重要特性并非由人预定义，而是由算法学习并发明。

为了更加明了，我们不注重真人工智能、强人工智能或通用人工智能这样的概念，它们意味着复制人类智能，也经常出现在流行文化中。虽然已有了一些有潜力的突破，比如谷歌 DeepMind 的 AlphaGo 系统，我们还是更注重立即有实在经济的人工智能发展。

为什么人工智能发展加速？

深度学习能力的极大发展是如今人工智能拐点背后的催化剂之一。深度学习的底层技术框架——神经网络，已经存在了数十年，但过去 5 到 10 年的 3 种东西改变了深度学习：

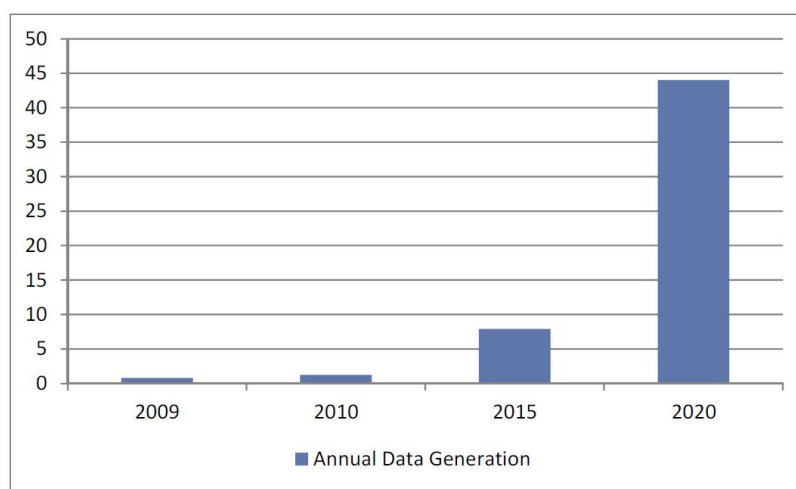
1. 数据。随着全球设备、机器和系统的连接，大量的无结构数据被发明出来。神经网络有了更多的数据，就变得更为有效，也就是说随着数据量增长，机器学习可以解决的问题也增长。手机、IoT、低成本数据存储和解决（云）技术的成熟使得可用数据集的大小、结构都有了极大增长。

例如，特斯拉收集了 780mn 英里的驾驶数据，并且通过他们的互连汽车，每 10 小时就能增长 100 万英里的数据。此外，Jasper 有一个平台，能让多家汽车制造商和电信公司进行机器间的交流，这家公司于今年 2 月份被 Cisco 收购。Verizon 在 8 月份做了类似

的投资，宣布收购 Fleetmatics，Fleetmatics 做的是将汽车上的远程传感器通过无线网络连接到云软件。

未来，5G 网络的上线将会加速数据生成与传输的速率。据 IDC 的 Digital Universe Report 显示，年度数据生成预期到 2023 年达成 44zettabytes，表白我们正在见证应用这些技术的使用案例。

Exhibit 1: Annual data generation is expected to reach 44 zettabytes (44 trillion GB) by 2020, according to EMC/IDC
annual data generation globally (in ZB)



Source: EMC, IDC

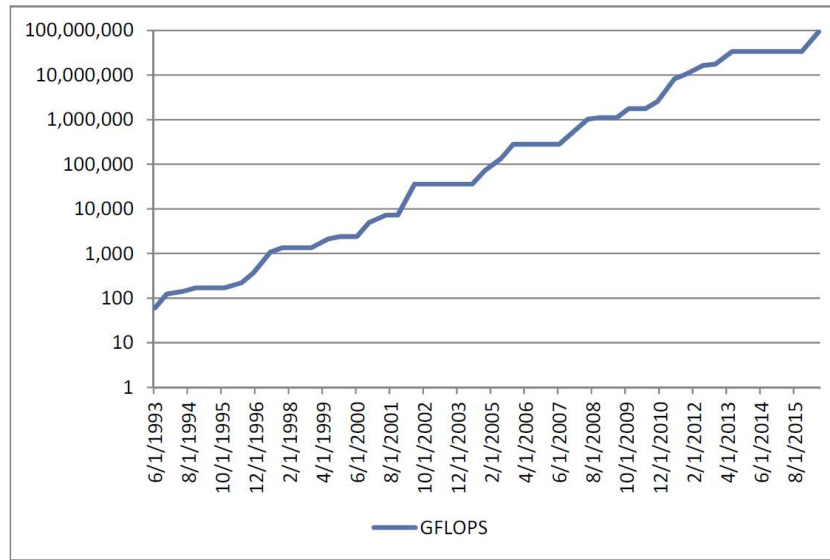
图 1：年度数据生成预期到 2023 年达成 44zettabytes

2. 更快的硬件。GPU 的再次使用、低成本计算能力的普遍化，特别是通过云服务，以及建立新的神经网络模型，已经极大的增长了神经网络产生结果的速度与准确率。GPU 和并行架构要比传统的基于数据中心架构的 CPU 能更快的训练机器学习系统。通过使用图像芯片，网络能更快的迭代，能在短期内进行更准确的训练。同时，特制硅的发展，比如微软和百度使用的 FPGA，可以用训练出的深度学习系统做更快的推断。此外，从 1993

年开始超级计算机的原计算能力有了极大发展（图 2）。在 2023 年，单张英伟达游戏显卡就有了类似于 2023 年之前最强大的超级计算机拥有的计算能力。

Exhibit 2: Raw compute performance of global supercomputers, measured in GFLOPs, has increased exponentially since 1993

Peak GFLOPS of #1 ranked global supercomputers on the Top500 list

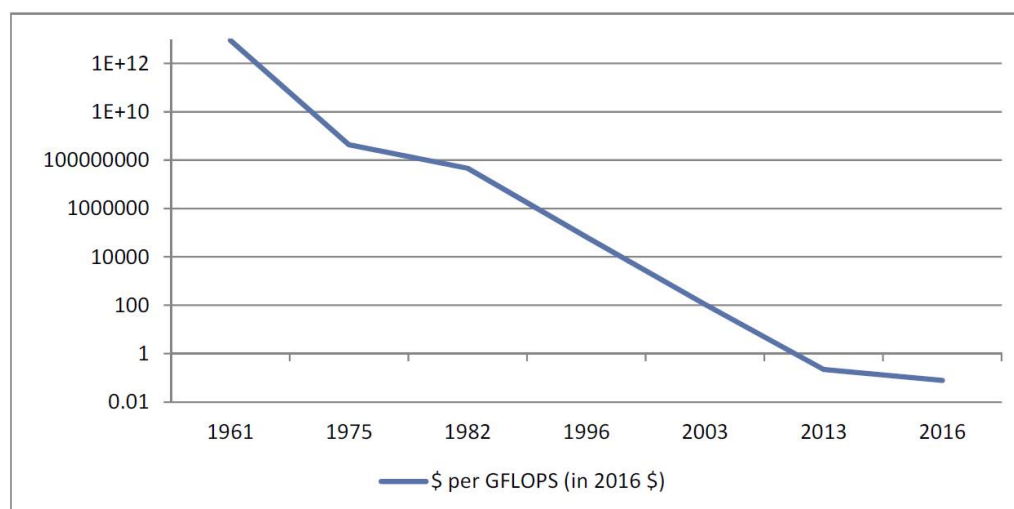


Source: top500.org, compiled by Goldman Sachs Global Investment Research

图 2：全球超级计算机的原计算性能，以 GFLOPs 测试

成本也有了极大的减少。英伟达 GPU（GTX 1080）有 9 TFLOPS 的性能，只要 700 美元，意味着每 GFLOPS 只要 8 美分。在 1961 年，串够 IBM 1620s 每提供 1 GFLOPS 需要的钱超过 9 万亿。

Exhibit 3: Price per unit of compute has decreased drastically over time
\$ per GFLOPS in representative computer systems



Source: IBM, Cray, Sony, Nvidia, Press reports, Goldman Sachs Global Investment Research

图 3：每单位计算的价格有了极大下降

3. 更好、更普遍可用的算法。更好的输入（计算和数据）使得更多的研发是面向算法，从而支持深度学习的使用。例如伯克利的 **Caffe**、谷歌的 **TensorFlow** 和 **Torch** 这样的开源框架。比如，刚开源一周年的 **TensorFlow**，成为了 **GitHub** 上有最多 **forked repositories** 的框架。虽然不是所有的人工智能发生于普遍可用的开源框架中，但开源的确在加速发展，并且也有更多先进的工具正在开源。

方向

虽然本报告的重点是人工智能的发展方向以及公司如何把握这个方向，但是了解人工智能对我们生活的影响限度也是很重要的。

在线搜索。就在一年多以前，谷歌透露，它们已经开始将大量的搜索工作移植到了 **RankBrain**

（一个人工智能系统），使其和链接（**links**）以及内容（**content**）成为了谷歌搜索算法的三个最重要的标志。

推荐引擎。**Netflix**，亚马逊 和 **Pandora** 都在使用人工智能来拟定推荐什么样的电影和歌曲，突出哪些产品。5 月，亚马逊开源了它们的深度可扩展稀疏传感网络引擎（**the Deep Scalable Sparse Tensor Network Engine**（**DSSTNE**），简称「**Destiny**」），它被用于产品推荐，同时可以被扩展以实现超越语言和语言理解以及异议辨认的目的。

人脸辨认。**Google**（**FaceNet**）和 **Facebook**（**DeepFace**）都投入了大量的技术来拟定您的照片中的人脸和真实的人脸是不是几乎完全吻合。1 月，苹果采用了进一步措施，购买了 **Emotient**（一个致力于通过读取人的面部表情来拟定其情绪状态的 **AI** 创业公司）显然，这些技术远远不止于对照片进行标记。

虽然个人助理应用产品有无数的用户，比如苹果的 **Siri**，信用贷，保险风险评估，甚至天气预测。在接下来的篇幅中，我们探讨公司该如何使用这些技术来加速增长，减少成本和控制风险。从这些技术及其使用这些技术的应用的发展速度来看，它们充其量但是可认为公司和投资者提供一些方向，以保持他们的竞争力。

加强未来的生产率

美国的劳动生产率在 90 年代中期的快速增长和过去十年的缓慢增长和之后，近年来已经停止增长了。我们认为，就像 20 世纪 90 年代互联网技术被广泛采用那样，消费类机器学习和人工智能的扩散有也许大幅度地改变全球产业的生产范式。

在整个行业中，我们发现在自动化的促使下，劳动时间减少了约 0.5%-1.5%，同时，由于 AI / ML 技术带来的效率增益，到 2025 年，这些技术将对生产力增长产生高达 51-1154 个基点（bps）的影响。虽然我们盼望 AI / ML 可以随着时间同时提高生产率的分母和分子，但是我们认为最重要的是，初期的影响将是低工资任务的自动化，即以更少的劳动时间推动类似的产出增长水平。我们的基本案例 AI / ML 驱动提高了 97 个基点，这意味着 2025 年的增长生产率中的 1.61% 将由 IT 奉献，比 1995 - 2023 年高出 11 个基点（图 9,10）。

Exhibit 9: Productivity analysis

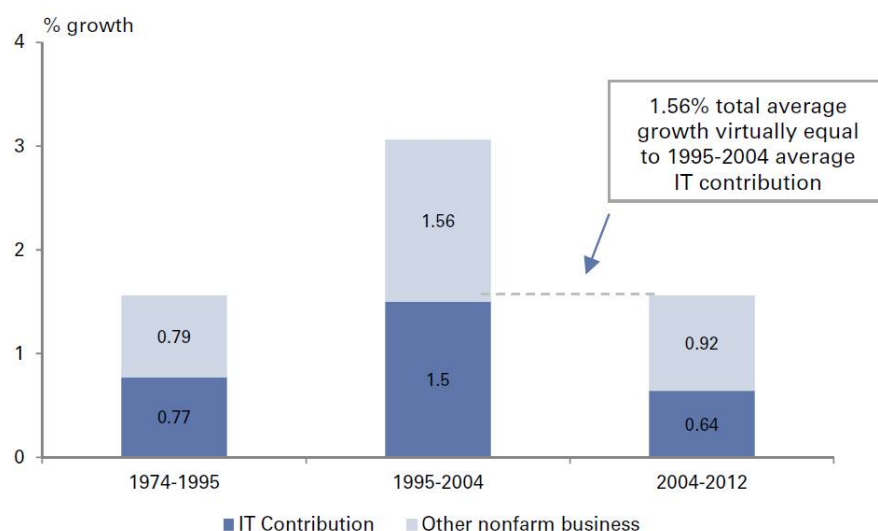
\$ millions, assumes linear nominal GDP growth beyond 2019

US	2016E	2017E	2018E	2019E	2020E	2021E	2022E	2023E	2024E	2025E
Output										
US Nominal GDP* (\$bn)	18,552	19,300	20,045	20,757	21,470	22,183	22,895	23,608	24,321	25,034
	2.9%	4.0%	3.9%	3.6%	3.4%	3.3%	3.2%	3.1%	3.0%	2.9%
Productivity										
Labor productivity	69.0	70.4	71.8	73.1	74.3	75.4	76.5	77.6	78.6	79.7
yoy growth (%)	0.9%	2.1%	2.0%	1.7%	1.6%	1.6%	1.5%	1.4%	1.3%	1.3%
Labor hours (mn)	268,958	273,992	279,026	284,060	289,094	294,128	299,162	304,196	309,230	314,264
ML/AI impact										
	Low			Base			High			
Labor hours reduction (mn)	(1,571)			(2,969)			(4,714)			
Reduction	-0.5%			-1%			-1.5%			
2025 Labor hours (mn)	312,693			311,295			309,550			
2025 GDP (\$bn)	25,034			25,034			25,034			
Labor productivity	80.1			80.4			80.9			
yoy growth (%)	1.8%			2.2%			2.8%			
Improvement (bps)	51			97			154			

Source: OECD, US Bureau of Labor Statistics, Goldman Sachs Global Investment Research

图 9：生产力分析；单位百万美元，假设 2023 年之后
GDP 线性增长

Exhibit 10: Late 90s: IT-producing sectors contribute nearly half of productivity growth
But lose value and share in growth post-tech boom



Source: Federal Reserve Board, Goldman Sachs Global Investment Research

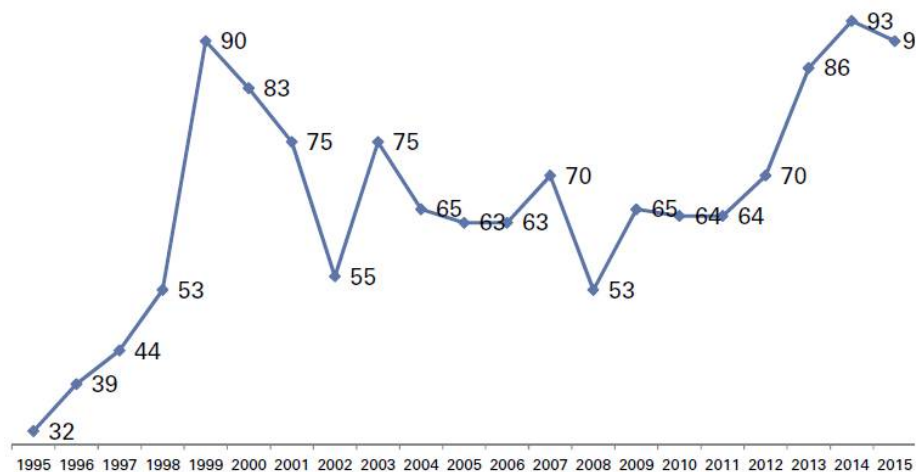
生态系统：云服务，开源在未来的 AI 投资周期中的关键受益人

我们相信，在未来的几年中，一个公司运用人工智能技术的能力将成为体现公司在所有重要行业竞争力的一个属性。虽然战略会因公司规模和行业而有所不同，但假如管理团队不会把重心放在领导人工智能和在此基础上的利益上，那么未来产品创新、劳动效率和资本杠杆都会存在落后的风险。因此，我们认为公司需要投资这些新技术以保持竞争力，同时这将导致对人工智能所以依赖的人才、服务和硬件的空前的需求。

作为比较，20 世纪 90 年代技术驱动的生产力繁荣推动了相应的激增。增长对技术的资本支出导致了新的公司和业务的增长来捕获这些资本支出。在不可避免的行业整合发生之前，成立软件、硬件和网络公司开始发生转变。下图 13 突出了软件行业内的这种模式。在 1995 - 1999 年期间，在通货膨胀调整后，市值在 20 亿美元到 50 亿美元之间的公共软件公司的数量几乎增长了两倍，在 2023 年代中期才得到巩固。

Exhibit 13: Rapid growth in the enabler ecosystem accompanied the 1990s productivity boom

of software companies with inflation adjusted market cap \$200mn-\$5bn (2015 dollars)

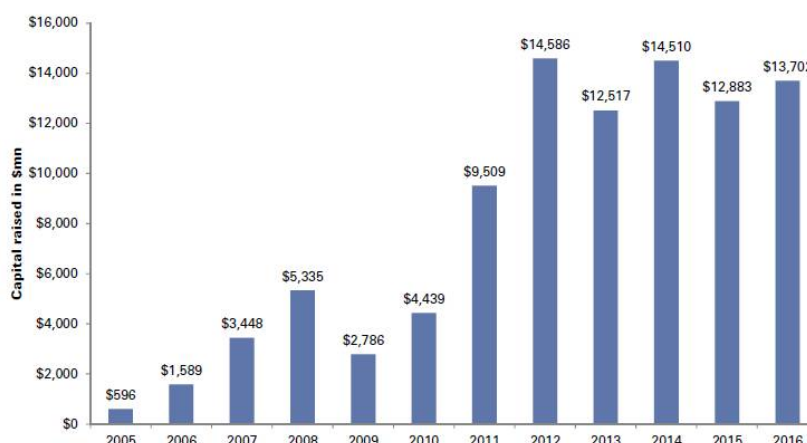


Source: Factset, Goldman Sachs Investment Research

图 13：随着 20 世纪 90 年代生产力激增的驱动者生态系统（enabler ecosystem）

Exhibit 14: Venture capital investing in AI has inflected this decade

Total VC funding in AI/ML by year



Source: PitchBook

图 14：这十年来，投资人工智能的风险资本出现了暴增

我们看到了由 AI 驱动的生产率具有产生下一个相似的繁荣周期的潜力，可以通过运用这些潜力，把软件、硬件、数据和服务提供商作为商业投资来发明价值。如上图 14 所反映的那样，与 AI 相关的初创公司的风险投资在这十年中急剧增长。AI 公司投资的繁荣现象的巨大潜力也开始推动整合。特别是云平台对 AI 相关人才进行了大量投入，自 2023 年以来，谷歌、亚马逊、微软和 Salesforce 共进行了 17 项与 AI 相关的收购（下图）。

Target	Date Announced/ Reported	Description
Acquirer: AMZN		
2lemetry Inc.	3/16/2015	IoT: track and manage connected devices
Orbeus	Fall 2015	Photo-re cognition technology based on AI
Acquirer: AAPL		
Vocal IQ	10/2/2015	Speech recognition based on AI
Perceptio	10/5/2015	Uses AI to classify photos on smartphones
Emotient	1/7/2016	Uses AI to read people's emotions by analyzing facial expressions
Turi	8/5/2016	Machine learning platform for developers and data scientists
Tuplejump Software	9/26/2016	Machine learning/big data technology developer
Acquirer: CRM		
Tempo AI	5/29/2015	AI-based smart calendar app
MinHash	12/16/2015	Developed an AI platform and a virtual personal assistant (AILA) for marketers
PredictionIO	2/19/2016	Developed an open source-based machine learning server
Metamind	4/4/2016	Deep Learning platform (natural language processing, computer vision, database predictions, etc)
Acquirer: MSFT		
Equivio	1/21/2015	Machine-learning based text analytics service for legal and compliance
Revolution Analytics	1/27/2015	Open-source analytics company that specializes in R programming language for statistical computing
Wand Labs	6/16/2016	Messaging app/chat bot developer
Genee	8/23/2016	AI scheduling assistance services
Acquirer: GOOGL		
DeepMind	1/16/2014	AI firm that specializes in machine learning, advanced algorithms, systems neuroscience
Emu	8/6/2014	Mobile messaging app with an AI assistant
Jetpac	8/15/2014	AI-based mobile photo app
Dark Blue Labs	10/23/2014	Deep learning startup specializing in understanding natural language
Vision Factory	10/23/2014	Deep learning startup specializing in visual recognition systems
Timeful	5/4/2015	Machine-learning based scheduling tool
Speaktit (Api.ai)	9/21/2016	Speech recognition and natural language understanding solutions

Source: Bloomberg, company data, FactSet, The Guardian, Techcrunch, VentureBeat

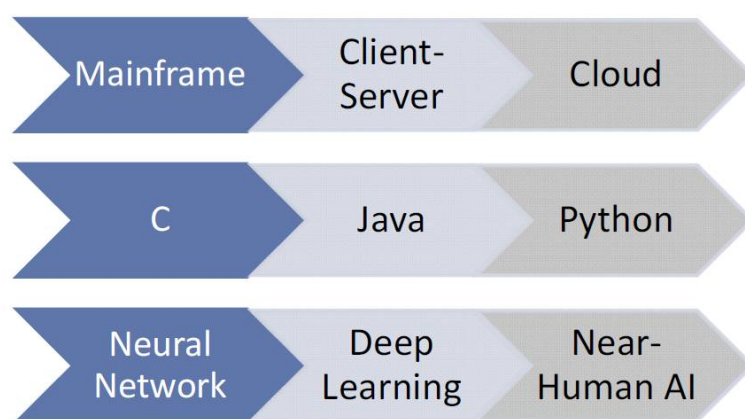
在上下文中对 AI 和 ML 技术的发展和历史技术周期的比较中，我们看到了前者的一些益处。与过去 50 年的其他重要技术的周期同样，计算（和摩尔定律）一直是进步的克制剂和推动者。

例如，在系统架构方面，我们目睹了从大型机系统转变为客户端-服务器模型的整个过程，并且近年来已经开始被云/移动模式所取代。这种进化的驱动因素是计算能力、存储容量和带宽的改善。每个转换都随着着应用开发的转变，涉及各种新编程语言的出现和演变（见图表 15）和各种也许的应用程序各种的类型。如上下文中所提及的同样，AI 这个概念已经存在几十年了，其中神经网络的概念出现在 20 世纪

60 年代，虽然直到最近几年，计算能力才开始让神经网络能在实际环境中使用。

我们相信我们正处在 AI 平台的初期阶段，就如同 20 世纪 50 年代大型机才开始商业化到 21 世纪的手机和云的商业化。随着平台曲线的变化（我们认为它正在发生），应用程序、工具和服务驱动者（enabler）会爆炸式增长，我们将在下面更具体地讨论。

Exhibit 15: AI advances can be compared to historical technological evolutions in systems architecture and programming language adoption, though we believe we are still in very early stages of development and adoption



Source: Goldman Sachs Global Investment Research

图 15：人工智能的发展可以与历史上的系统架构和编程语言的采用的技术革命相比，尽管我们认为我们目前仍然处在人工智能发展和应用的初期阶段

stack 的演变过程以及和 AI 之间的相应关系

蓝色 = 专有供应商，橙色 = 开源，绿色 = 云服务（注意：一些供应商，如 IBM 和 Microsoft 都是专有服务和云服务）

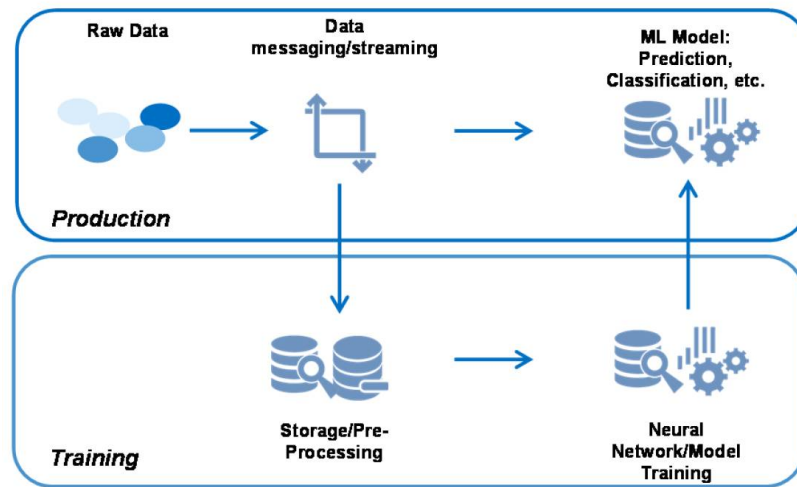
EVOLUTION OF THE STACK				
	Mainframe	Client-Server	Current Era (Hybrid)	AI
 Analytics Tools (Select Vendors)	Decision Support Systems, Reporting, Statistics IBM, SAP, SAS, SPSS	Business Intelligence (BI), Reporting, Statistics Cognos, Business Objects, Crystal Systems, Microstrategy, SAS	BI, Reporting, Data Visualization, Statistics IBM, SAS, SAP, Microsoft, Tableau, Qlik, Domo, Looker, Amazon, Salesforce	Visualization, Machine Learning Services and Frameworks Tableau, IBM, NVIDIA, SAS, Skyminid, Databricks, Amazon, Google, Microsoft, Alibaba, Baidu
 Programming Languages (Select Vendors)	COBOL, Pascal, ALGOL IBM, Borland Software, Burroughs, NCR, DEC	Java, .Net, C, SQL Sun Microsystems, Microsoft	Java, .Net, javascript, Ruby, SQL Oracle, Microsoft	Python, R, SQL Microsoft
 Data Processing (Select Vendors)	Hierarchical Databases IBM, DEC, Unisys	Relational Databases IBM, Oracle, Microsoft, Informix, Sybase, Teradata, MySQL	Relational, NoSQL, Hadoop Oracle, Microsoft, SAP, IBM, Teradata, MongoDB, DataStax, Cloudera, Hortonworks, Amazon, Google	Neural Networks, Stream Processing, NoSQL, Hadoop, Relational IBM, Databricks, Cloudera, Hortonworks, Amazon, Google, Microsoft, Qubole, Alibaba, Baidu
 Data Movement/Integration (Select Vendors)	Limited Data Movement IBM	ETL/Data Integration (DI), Messaging, Application Integration (EAI) Prism, Informatica, Ab Initio, Ascential, TIBCO, BEA, Microsoft	DI, EAI, Streaming, Messaging IBM, Oracle, Informatica, Talend, Confluent, Cloudera, Hortonworks, Mulesoft, Amazon, Google	Streaming, Preparation, Messaging IBM, Talend, Confluent, Cloudera, Hortonworks, Databricks, Skyminid, Amazon, Google, Microsoft, Alibaba, Baidu
 Storage (Select Vendors)	Mainframes, Tape DEC, IBM	SAN, NAS EMC, NetAPP, IBM	SAN, NAS, Hadoop, Cloud Object Storage Dell, NetApp, IBM, Cloudera, Hortonworks, MapR, Amazon, Microsoft, Google, Baidu	Hadoop, Cloud Object Storage IBM, Dell, Cloudera, Hortonworks, MapR, Amazon, Microsoft, Google, Baidu, Alibaba
 Compute (Select Vendors)	Mainframes DEC, IBM, NCR, Burroughs	Unix Servers, x86 IBM, Dell, HP	x86, cloud compute instances HP, Dell, Lenovo, Huawei, Cisco, Amazon, Google, Microsoft	x86, cloud compute instances IBM, HP, Dell, Lenovo, Huawei, Cisco, Amazon, Google, Microsoft, Baidu, Alibaba
 Components (Select Vendors)	HDD, CPU IBM, Fairchild, DEC, CDC, EMC	HDD, CPU IBM, Intel, AMD, Western Digital, Seagate, Sandisk, TMS	HDD, SSD, CPU Intel, ARM Holdings, Western Digital, Seagate	HDD, SSD, CPU, GPU, FPGA Intel, ARM Holdings, Western Digital, Seagate, NVIDIA, Xilinx, Google

proprietary vendors | open source | cloud services

Source: Company data, Goldman Sachs Global Investment Research

Exhibit 17: Machine Learning in Production

How various open source and cloud technologies would be utilized in the machine learning pipeline



Source: Company data, Goldman Sachs Global Investment Research

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/295222041112011214>