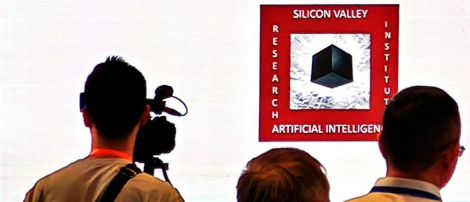
Drivers of Progress in AI

and Silicon Valley's AI Innovation

Piero Scaruffi

Chengdu, September 2024



人工智能发展的驱动力与硅谷人工智能创新

皮埃罗·斯加鲁菲

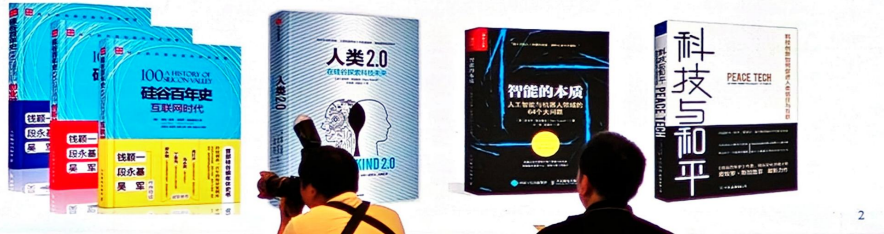
2024年9月 成都

Piero Scaruffi

- 40+ years in Silicon Valley
- Written 20+ books
- Pioneered A.I. and Internet applications in the 1980s
- Founded international inter-university programs (LASERs, LAST Festival)
- Interdisciplinary research at Stanford, UC Berkeley, etc

皮埃罗·斯加鲁菲

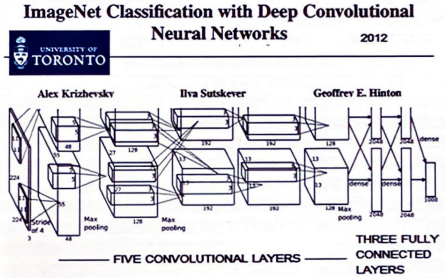
- 40+年硅谷工作和生活经验
- 撰写了20余本著作
- 20世纪80年代开始致力于人工智能和互联网应用的开创性工作
- 创办了国际性跨高校交流活动（LASER，即莱昂纳多艺术科学之夜，LAST节，即融汇生活、艺术、科学、技术）
- 在斯坦福大学、加州大学伯克利分校等高校从事跨学科研究



How did we go from this…

AlexNet (2012)

ImageNet Classification with Deep Convolutional Neural Networks 2012

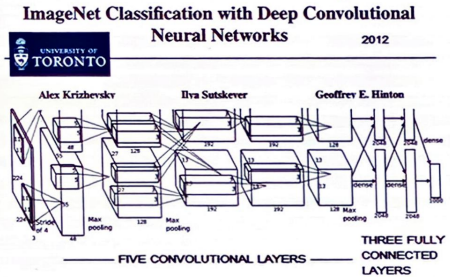


我们是如何从这里开始...

AlexNet (2012)

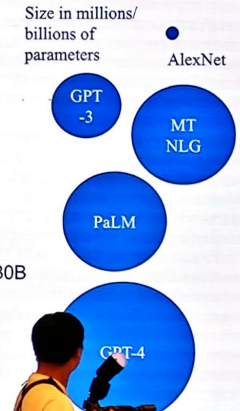
使用深度卷积神经网络进行ImageNet分类 *20*

ImageNet Classification with Deep Convolutional Neural Networks 2012



… to this?

Google BERT (2018) 110M parameters
Facebook's RoBERTa (2019) 123M
Microsoft's XLNet (2019) 110M
OpenAI's GTP2 (2019) 1.5B
Nvidia's Megatron (2019) 8.3B
Google T5 (2019) 11B
Microsoft Turing-NLG (2020) 17.2B
Microsoft DeBERTa (2020) 304M
OpenAI's GTP-3 (2020) 175B
Google DeepMind's Gopher (2021) 280B
Microsoft/Nvidia's Megatron-Turing NLG (2021) 530B
Google PaLM (2022) 540B
Aleph ous (2022) in Germany 200B
Meta and LLaMA (2023)
Op (2024)



…发展到这样的?

谷歌BERT (2018年)：1.1亿参数
脸书RoBERTa (2019年)：1.23亿参数
微软XLNet (2019年)：1.1亿参数
OpenAI的GTP-2 (2019年)：15亿参数
英伟达Megatron (2019年)：83亿参数
谷歌T5 (2019年)：110亿参数
微软Turing-NLG (2020年)：172亿参数
微软DeBERTa (2020年)：3.04亿参数
OpenAI的 GTP-3 (2020年)：1750亿参数
谷歌DeepMind的Gopher (2021年)：2800亿参数
微软/英伟达Megatron-Turing NLG (2021年)：5300亿参数
谷歌PaLM (2022年)：5400亿参数
Aleph Alpha的Luminous (2022年，德国)：2000亿参数
Meta的OPT (2022年)和LLaMA (2023年)
OpenAI的GTP-4 (2023年)

模型参数量大小以百万(M)或十亿(B)为单位

The real drivers of the AI revolution…

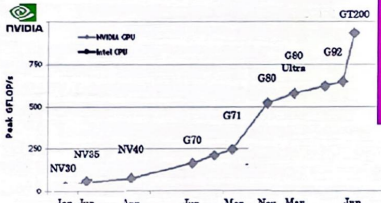
Annus mirabilis of Deep Learning: 2012
Why did it take until 2012?

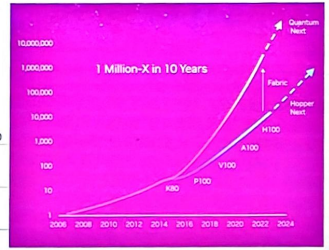
人工智能发展的真正驱动力是…

2012 年被认为是深度学习的“奇迹之年”
为什么直到2012年才取得突破?

1. The GPU

Nvidia's GPUs



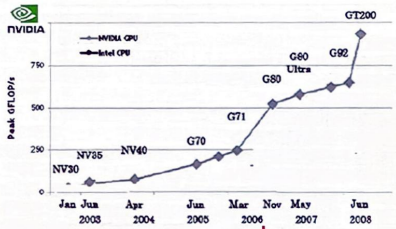


2009: **Andrew Ng (Stanford)** demonstrates a 100M deep belief network trained on 30 Nvidia GeForce GTX 280 GPUs

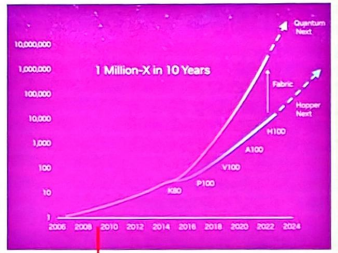
2006: Deep Learning is born

1. 图形处理器（GPU）

英伟达的GPU



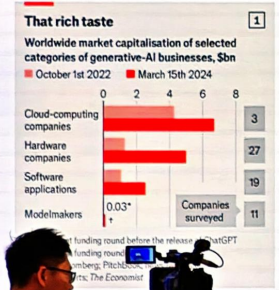
2006: 深度学习诞生

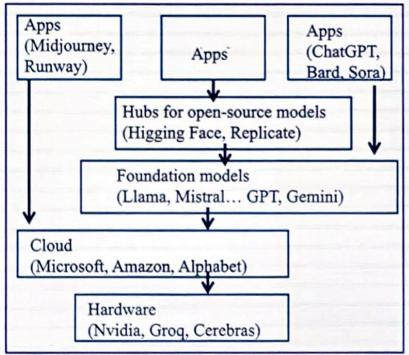


2009: **斯坦福大学**的**吴恩达**展示了一个由1亿参数组成的深度信念网络（DBN）模型，该模型使用了30个英伟达的GeForce GTX 280 GPU进行训练。

A.I. depends on hardware

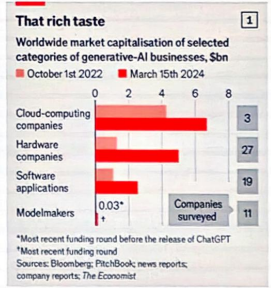
Hardware companies benefits the most

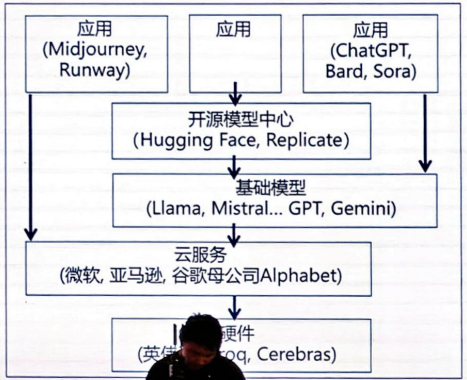


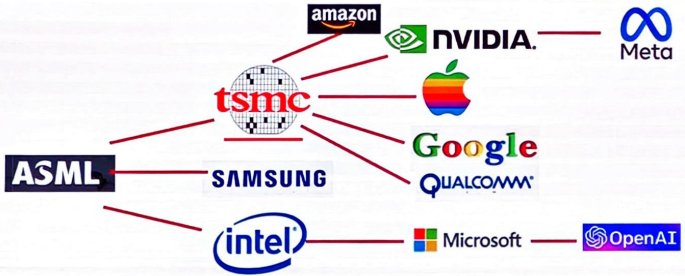


人工智能的发展取决于硬件

硬件公司受益最大

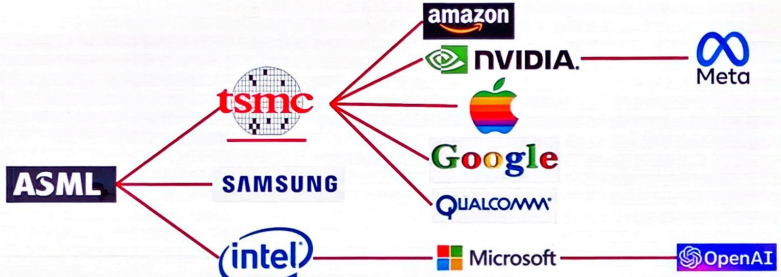






ASML's High NA second-generation EUV will allow foundries to produce chips using a process node under 3nm

The $150 Million Machine Keeping Moore's Law Alive



ASML's High NA second-generation EUV will allow foundries to produce chips using a process node under 3nm

The $150 Million Machine Keeping Moore's Law Alive

Chiplets

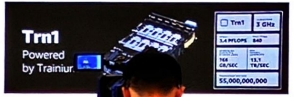
- Build a microprocessor by packaging together many small chips
- AMD (2019): Zen 2 architecture (with chiplets made by TSMC with 7nm technology)
  - the third generation of Ryzen microprocessors for personal computers
  - the first generation of EPYC processors
- 2021 Apple M1 Max chip
- 2022 Intel announces Meteor Lake chip

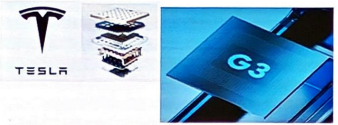
芯粒

- 通过将多个小芯片组合封装在一起来构建一个微处理器
- AMD (2019): Zen 2 架构 (采用台积电7nm工艺制造的芯粒)
  - 用于个人电脑的第三代Ryzen微处理器
  - 第一代EPYC “霄龙” 处理器
- 2021年苹果发布M1 Max芯片
- 2022年英特尔发布Meteor Lake芯片

A.I. depends on hardware

Apple, Tesla, Google, Microsoft and Amazon are now making their own ASIC chips
Amazon (2024): second generation of Trainium
Microsoft (2023): Maia
Google (2023): Tensor G3
Meta (2023): MTIA
Tesla (2024?): next-gen Dojo





人工智能的发展取决于硬件

苹果、特斯拉、谷歌、微软和亚马逊现在都在自主开发ASIC芯片
亚马逊 (2024年)：第二代Trainium
微软(2023年): Maia
谷歌(2023年): Tensor G3
Meta (2023年): MTIA
特斯拉(2024年?): 下一代 Dojo

2. The datasets

ImageNet (2009, Princeton Univ)
Atari games (2013, Univ of Alberta)
COCO (2014, Microsoft)
LSUN (2015, Princeton Univ)
UM (2014, Macau Univ)
SQuAD (2016, Stanford)
MARCO (2016, Microsoft)
YouTube-8M (2016, Google)
Kinetics-600 (2017, Oxford)

IMAGENET
Large Scale Visual
Recognition
Challenge

2. 数据集（Dataset）

ImageNet (2009年, 普林斯顿大学)
Atari 游戏 (2013年, 阿尔伯塔大学)
COCO (2014年, 微软)
LSUN (2015年, 普林斯顿大学)
UM (2014年, 澳门大学)
SQuAD (2016年, 斯坦福大学)
MARCO (2016年, 微软)
YouTube-8M (2016年, 谷歌)
Kinetics-600 (2017年, 牛津大学)

IMAGENET
Large Scale Visual
Recognition
Challenge

李飞飞

The Datasets drive AI

1991: IBM creates a dataset of 700,000 chess games played by chess masters
1997: Deep Blue beats the world champion of chess
2009: Feifei Li's ImageNet large dataset of tagged images
2012: Spectacular improvement in image recognition
2013: Michael Bowling's dataset of Atari games
2015: DeepMind's videogame-playing program
2016: DeepMind's dataset of 150,000 weiqi games
2017: AlphaGo beats the world master of weiqi
2018: OpenAI's WebText
2019: OpenAI's GPT2

OpenAI

数据集驱动着人工智能的发展

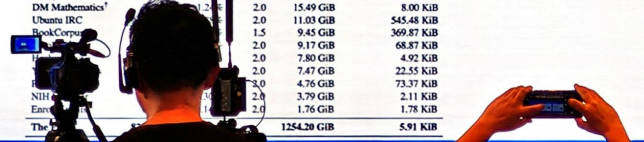
1991年：IBM创建了一个包含70万场国际象棋大师比赛记录的数据集。
1997年：深蓝（Deep Blue）击败了国际象棋世界冠军。
2009年：李飞飞发布了 ImageNet 大规模标记图像数据集
2012年：图像识别能力显著提升
2013年：Michael Bowling的 Atari 游戏数据集发布
2015年：DeepMind推出了视频游戏玩家程序
2016年：DeepMind发布了15万盘围棋对局的数据集
2017年：AlphaGo击败围棋世界冠军
2018年：OpenAI发布WebText数据集
2019年：OpenAI发布GPT-2模型

OpenAI

Datasets of 2024

The Pile (hosted on The Eye): a 825 GiB open-source data set that consists of 22 smaller datasets combined together

Component	Raw Size	Weight	Epochs	Effective Size	Mean Document Size
Pile-CC	227.12 GiB	18.11%	1.0	227.12 GiB	4.33 KiB
PubMed Central	90.27 GiB	14.40%	2.0	180.55 GiB	30.55 KiB
Books3[†]	100.96 GiB	12.07%	1.5	151.44 GiB	538.36 KiB
OpenWebText2	62.77 GiB	10.01%	2.0	125.54 GiB	3.85 KiB
ArXiv	56.21 GiB	8.96%	2.0	112.42 GiB	46.61 KiB
Github	95.16 GiB	7.59%	1.0	95.16 GiB	5.25 KiB
FreeLaw	51.15 GiB	6.12%	1.5	76.73 GiB	15.06 KiB
Stack Exchange	32.20 GiB	5.13%	2.0	64.39 GiB	2.16 KiB
USPTO Backgrounds	22.90 GiB	3.65%	2.0	45.81 GiB	4.08 KiB
PubMed Abstracts	19.26 GiB	3.07%	2.0	38.53 GiB	1.30 KiB
Gutenberg (PG-19)[†]	10.88 GiB	2.17%	2.5	27.19 GiB	398.73 KiB
OpenSubtitles[†]	12.98 GiB	1.55%	1.5	19.47 GiB	30.48 KiB
Wikipedia (en)[†]	6.38 GiB	1.53%	3.0	19.13 GiB	1.11 KiB
DM Mathematics[†]	7.75 GiB	1.24%	2.0	15.49 GiB	8.00 KiB
Ubuntu IRC	5.52 GiB	0.88%	2.0	11.03 GiB	545.48 KiB
BookCorpus2	6.30 GiB	0.75%	1.5	9.45 GiB	369.87 KiB
EuroParl[†]	4.59 GiB	0.73%	2.0	9.17 GiB	68.87 KiB
HackerNews	3.90 GiB	0.62%	2.0	7.80 GiB	4.92 KiB
YoutubeSubtitles	3.73 GiB	0.60%	2.0	7.47 GiB	22.55 KiB
PhilPapers	2.38 GiB	0.38%	2.0	4.76 GiB	73.37 KiB
NIH ExPorter	1.89 GiB	0.30%	2.0	3.79 GiB	2.11 KiB
Enron Emails[†]	0.88 GiB	0.14%	2.0	1.76 GiB	1.78 KiB
The Pile	**825.18 GiB**			**1254.20 GiB**	**5.91 KiB**



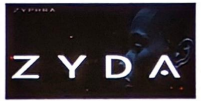
2024年的数据集

The Pile（托管于The Eye网站平台）：一个825 GiB的**开源**数据集，由22个子数据集组合而成。

Component	Raw Size	Weight	Epochs	Effective Size	Mean Document Size
Pile-CC	227.12 GiB	18.11%	1.0	227.12 GiB	4.33 KiB
PubMed Central	90.27 GiB	14.40%	2.0	180.55 GiB	30.55 KiB
Books3[†]	100.96 GiB	12.07%	1.5	151.44 GiB	538.36 KiB
OpenWebText2	62.77 GiB	10.01%	2.0	125.54 GiB	3.85 KiB
ArXiv	56.21 GiB	8.96%	2.0	112.42 GiB	46.61 KiB
Github	95.16 GiB	7.59%	1.0	95.16 GiB	5.25 KiB
FreeLaw	51.15 GiB	6.12%	1.5	76.73 GiB	15.06 KiB
Stack Exchange	32.20 GiB	5.13%	2.0	64.39 GiB	2.16 KiB
USPTO Backgrounds	22.90 GiB	3.65%	2.0	45.81 GiB	4.08 KiB
PubMed Abstracts	19.26 GiB	3.07%	2.0	38.53 GiB	1.30 KiB
Gutenberg (PG-19)[†]	10.88 GiB	2.17%	2.5	27.19 GiB	398.73 KiB
OpenSubtitles[†]	12.98 GiB	1.55%	1.5	19.47 GiB	30.48 KiB
Wikipedia (en)[†]	6.38 GiB	1.53%	3.0	19.13 GiB	1.11 KiB
DM Mathematics[†]	7.75 GiB	1.24%	2.0	15.49 GiB	8.00 KiB
Ubuntu IRC	5.52 GiB	0.88%	2.0	11.03 GiB	545.48 KiB
BookCorpus2	6.30 GiB	0.75%	1.5	9.45 GiB	369.87 KiB
EuroParl[†]	4.59 GiB	0.73%	2.0	9.17 GiB	68.87 KiB
HackerNews	3.90 GiB	0.62%	2.0	7.80 GiB	4.92 KiB
YoutubeSubtitles	3.73 GiB	0.60%	2.0	7.47 GiB	22.55 KiB
PhilPapers	2.38 GiB	0.38%	2.0	4.76 GiB	73.37 KiB
NIH ExPorter	1.89 GiB	0.30%	2.0	3.79 GiB	2.11 KiB
Enron Emails[†]	0.88 GiB	0.14%	2.0	1.76 GiB	1.78 KiB
The Pile	**825.18 GiB**			**1254.20 GiB**	**5.91 KiB**

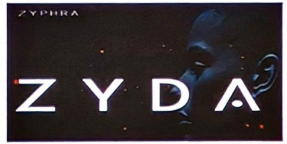
Datasets of 2024

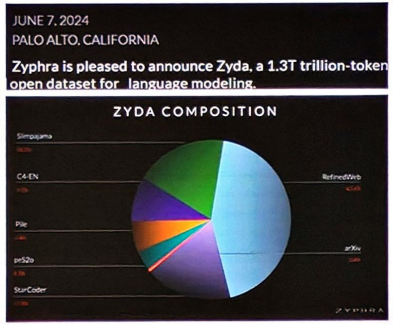
Zyphra's Zyda: 1.3 trillion tokens, a mashup of RefinedWeb, Starcoder, C4, Pile, Slimpajama, pe2so, and arxiv



2024年的数据集

Zyphra的**Zyda**：包含1.3万亿个token，由RefinedWeb、Starcoder、C4、Pile、Slimpajama、pe2so和arxiv多个知名数据集混合而成。





3. Open Source

Open-source platforms for deep learning

- Torch (Ronan Collobert @ IDIAP, Switzerland, 2002)
- Theano (Bengio's group @ Univ of Montreal, Canada, 2010)
- Caffe (Yangqing Jia @ UC Berkeley, 2013)
- TensorFlow (Rajat Monga @ Google, 2015)

3. 开源

开源深度学习平台

Torch（2002年，Ronan Collobert 在瑞士IDIAP研究所开发）

Theano（2010年，Bengio团队在加拿大蒙特利尔大学开发）

Caffe（2013年，贾扬清在加州大学伯克利分校开发）

TensorFlow（2015年，Rajat Monga在谷歌公司开发）

Open-source models

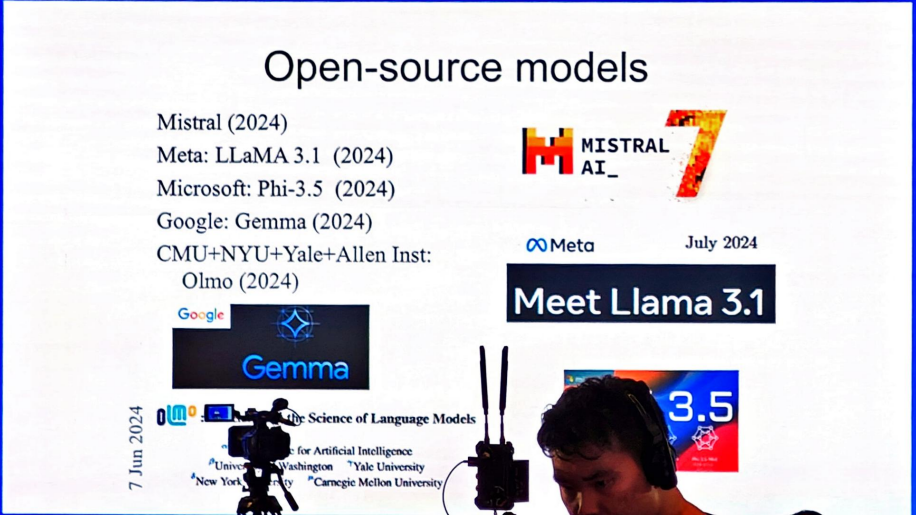
EleutherAI: open-source versions of GPT-3 (GPT-Neo, GPT-J, GPT-NeoX, Polyglot, Pythia)
Carnegie Mellon University and Google: XLNet (2019)
Facebook AI and Univ of Washington : RoBERTa (2019)
Microsoft: DeBERTa (2020)
Meta: XLM-RoBERTa (2019), LLaMA (2023), XGLM (2022)
Google: FLAN (2021) and Flamingo (2022)
BigScience: BLOOM (2022)
Cerebras' Cerebras-GPT (2023)
Zhejiang Univ's HuggingGPT (2023)
Databricks: Dolly (2023)
Stanford: Alpaca and Vicuna (2023)
Salesforce: Xgen (2023)
Abu Dhabi: Falcon (2023)
Mistral (2024)
Meta: LLaMA 3.1 (202
HyperW
d on LLaMA 3.1

开源模型

EleutherAI：GPT-3的开源版本 (GPT-Neo, GPT-J, GPT-NeoX, Polyglot, Pythia)
卡内基梅隆大学和谷歌：XLNet (2019)
Facebook AI 和华盛顿大学：RoBERTa (2019)
微软：DeBERTa (2020)
Meta：XLM-RoBERTa (2019), LLaMA (2023), XGLM (2022)
谷歌：FLAN (2021) 和 Flamingo (2022)
BigScience：BLOOM (2022)
Cerebras：Cerebras-GPT (2023)
浙江大学：HuggingGPT (2023)
Databricks：Dolly (2023)
斯坦福大学：Alpaca 和 Vicuna (2023)
Salesforce：Xgen (2023)
阿布扎比：Falcon (2023)
Mistral (2024)
Meta：LLaMA 3.1 (2024)
HyperWrite：Reflection 70B (2024), 基于LLaMA 3.1
微软：Phi-3.5 (2024)

Open-source models

Mistral (2024)
Meta: LLaMA 3.1 (2024)
Microsoft: Phi-3.5 (2024)
Google: Gemma (2024)
CMU+NYU+Yale+Allen Inst: Olmo (2024)

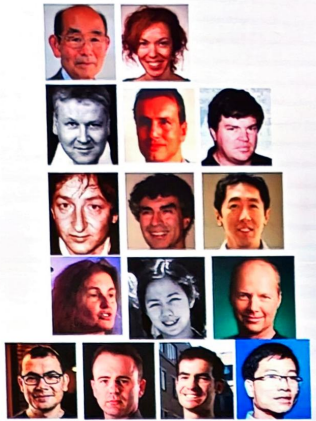


开源模型

Mistral (2024年)
Meta: **LLaMA 3.1** (2024年)
微软: **Phi-3.5** (2024年)
谷歌: **Gemma** (2024年)
卡内基梅隆大学+纽约大学 +耶鲁大学+艾伦研究所: **Olmo** (2024年)

4. International Collaboration

Kunihiko Fukushima: Japan
Hava Siegelmann: Israel
Sepp Hochreiter: Germany
Juergen Schmidhuber: Switzerland
Yann LeCun: France
Geoffrey Hinton: Britain/ Canada
Yoshua Bengio: France/ Canada
Andrew Ng: China
Daniela Rus: Romania
Fei-fei Li: China
Sebastian Thrun: Germany
DeepMind: Britain/ New Zealand
Ilya Sutskever: Russia
Quoc Le: Vietnam



4. 国际合作

福岛邦彦（Kunihiko Fukushima）：日本
哈娃·西格尔曼 （Hava Siegelmann）：以色列
塞普·霍赫赖特 （Sepp Hochreiter）：德国
于尔根·施米德胡伯 （Juergen Schmidhuber）：瑞士
杨立昆（Yann LeCun）：法国
杰弗里·辛顿（Geoffrey Hinton）：英国/ 加拿大
约书亚·本吉奥（Yoshua Bengio）：法国/ 加拿大
吴恩达（Andrew Ng）：中国
丹妮拉·鲁斯（Daniela Rus）：罗马尼亚
李飞飞（Fei-fei Li）：中国
塞巴斯蒂安·特伦（Sebastian Thrun）：德国
DeepMind：英国/ 新西兰
伊利亚·苏茨克维（Ilya Sutskever）：俄罗斯
Quoc Le：越南

E.g.: the Stars of Deep Learning

Kunihiko Fukushima: Japan
Hava Siegelmann: Israel
Sepp Hochreiter: Germany
Juergen Schmidhuber: Switzerland
Yann LeCun: France
Geoffrey Hinton: Britain/ Canada
Yoshua Bengio: France/ Canada
Andrew Ng: China
Daniela Rus: Romania
Fei-fei Li: China
Sebastian Thrun: Germany
DeepMind: Britain/ New Zealand
Ilya Sutskever: Russia
Quoc Le: Vietnam
Jitendra Malik: India

Dong Yu: China
Oriol Vinyals: Spain
Ian Goodfellow: Canada
Karen Simonyan: Armenia
Andrew Zisserman: Britain
Christian Szegedy: Germany
Aja Huang: China
Kaiming He: China
Jian Sun: China
Andrej Karpathy: Slovakia
Pieter Abbeel: Belgium
Ronan Collobert:
Yangqing Jia:
Rajat Monga:
Richard Socher:

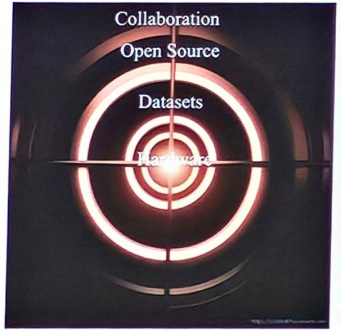
如：深度学习的领军人物

福岛邦彦(Kunihiko Fukushima)：日本
哈娃·西格尔曼 (Hava Siegelmann)：以色列
塞普·霍赫赖特 (Sepp Hochreiter)：德国
于尔根·施米德胡伯 (Juergen Schmidhuber)：瑞士
杨立昆(Yann LeCun)：法国
杰弗里·辛顿(Geoffrey Hinton)：英国/ 加拿大
约书亚·本吉奥(Yoshua Bengio)：法国/ 加拿大
吴恩达(Andrew Ng)：中国
丹妮拉·鲁斯(Daniela Rus)：罗马尼亚
李飞飞(Fei-fei Li)：中国
塞巴斯蒂安·特伦(Sebastian Thrun)：德国
DeepMind：英国/ 新西兰
伊利亚·苏茨克维(Ilya Sutskever)：俄罗斯
Quoc Le：越南
吉滕德拉·马利克 (Jitendra Malik)：印度

俞栋(Dong Yu)：中国
奥里奥尔·维尼亚尔斯 (Oriol Vinyals)：西班牙
伊恩·古德费洛(Ian Goodfellow)：加拿大
卡伦·西莫尼扬(Karen Simonyan)：亚美尼亚
安德鲁·齐瑟曼 (Andrew Zisserman)：英国
克里斯蒂安·塞杰迪 (Christian Szegedy)：德国
黄士杰 (Aja Huang)：中国
何恺明(Kaiming He)：中国
孙剑 (Jian Sun)：中国
安德烈·卡帕斯 (Andrej Karpathy)：斯洛伐克
皮特·阿比尔 (Pieter Abbeel)：比利时
罗南·科洛贝尔 (Ronan Collobert)：法国
贾扬清 (Yangqing Jia)：中国
拉贾特·蒙加 (Rajat Monga)：印度
理查德·索查 (Richard Socher)：德国

The real drivers of the AI revolution…

1. Hardware
2. Datasets
3. Open Source
4. International Collaboration

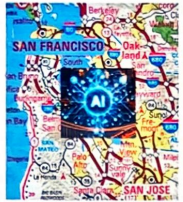


人工智能发展的真正驱动力是…

1. 硬件
2. 数据集
3. 开源
4. 国际合作

Why Silicon Valley… again?

- 2024: Silicon Valley = AI Valley
- 2022: The Bay Area accounted for about 25% of AI conference papers, patents and startups
- Today: 1425 Artificial Intelligence jobs available in San Francisco Bay Area out of 28656 Artificial Intelligence jobs available on Indeed.com (5% of all jobs out of a population that is 2% of the USA)

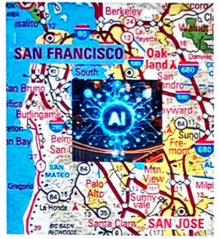


为何又是硅谷...?

2024年：硅谷 = AI谷

2022年：湾区在人工智能会议论文、专利和初创企业方面占比约25%

如今：Indeed.com网站上有28,656个人工智能相关的职位，其中1,425个在旧金山湾区（占职位总数的5%，而该地区人口仅占美国的2%）



Silicon Valley in 1950

1950年的硅谷

Today


1950


Berkeley
Mt Diablo
zymergen
Oakland
Alameda
San Leandro
Hayward
Pleasanton
Daly City
Genentech
Pacifica
Fremont
Sunol
GILEAD
ORACLE
23andMe
PayPal
Stanford University
Palo Alto
NVIDIA
zoom
Milpitas
vmware
Google
intel
SILICON VALLEY
cisco
San Jose
Tassajara
Diablo
Don Edwards San Francisco Bay National
El Granada
Half Moon Bay
Lobitos

如今的硅谷


1950


Berkeley
Mt Diablo
zymergen
Oakland
Alameda
San Leandro
Hayward
Pleasanton
Daly City
Genentech
Pacifica
Fremont
Sunol
GILEAD
ORACLE
23andMe
PayPal
Stanford University
Palo Alto
NVIDIA
zoom
Milpitas
vmware
Google
intel
SILICON VALLEY
cisco
San Jose
Tassajara
Diablo
Don Edwards San Francisco Bay National
El Granada
Half Moon Bay
Lobitos

Why Silicon Valley?

- What Silicon Valley does best
  - Not invented here: computer, transistor, integrated circuit, robots, Artificial Intelligence, programming languages, databases, videogames, Internet, personal computers, World-wide web, search engines, social media, smartphones, wearable computing, space exploration, electrical cars, driverless cars, generative AI…

为何是硅谷?

- **硅谷最擅长做什么?**
  - **以下这些并非都是在硅谷发明的:**

    计算机、晶体管、集成电路、机器人、人工智能、编程语言、数据库、电子游戏、互联网、个人电脑、万维网、搜索引擎、社交媒体、智能手机、可穿戴设备、太空探索、电动车、自动驾驶汽车、生成式AI...

Why Silicon Valley?

- What Silicon Valley does best
  - Invented here: disrupting products

UBER airbnb

You
vc investment by state of the USa in 2023

ChatGPT
I don't have real-time access to current dat
investments by state in 2023. However, you
reports published by research firms specia

Silicon Valley takes inventions and [illegible] into disruptive [illegible]ogies

为何是硅谷？

- 硅谷最擅长做什么？
  - 这里真正擅长的：发明颠覆性产品

UBER airbnb

You
vc investment by state of the USa in 2023

ChatGPT
I don't have real-time access to current dat
investments by state in 2023. However, you
reports published by research firms specia

硅谷将发明转变为颠覆性技术。

What is really special?

Innovation that came from small companies

- Integrated circuit (Fairchild) — 1960s
- Microprocessor (Intel)
- Unix server (SUN)
- GUI (Apple) — 1970s
- SQL Database (Oracle)
- Networking switches (Cisco)
- Biotech (Genentech) — 1980s
- Relational DB (Oracle)
- Ecommerce (eBay) — 1990s
- Search engine (Google)
- Online payment (Paypal) — 2000s
- Social network (Facebook, Twitter)
- Sharing economy (Airbnb, Uber) — 2010s
- Video [illegible]acing (Zoom)
- Elect[illegible]driving car (Tesla) — 2020s
- Genera[illegible]AI (OpenAI)

硅谷真正的特别之处?

源自小公司的创新

- 集成电路（仙童） — 1960s
- 微处理器（英特尔）
- Unix 服务器系统（SUN）
- 图形用户系统（苹果） — 1970s
- SQL 数据库（甲骨文）
- 网络交换机（思科）
- 生物技术（基因泰克） — 1980s
- 关系数据库（甲骨文）
- 电子商务 （eBay） — 1990s
- 搜索引擎（谷歌）
- 在线支付（Paypal） — 2000s
- 社交网络（脸书, 推特）
- 共享经济（Airbnb, Uber） — 2010s
- 视频会议（Zoom）
- 电动自动驾驶汽车（特斯拉） — 2020s
- 生成式人工智能 （OpenAI）

What is really special?

Open Source

- GItHub has 100 million developers

GitHub Trending

硅谷真正的特别之处?

开源

GItHub拥有一亿名开发者。

GitHub Trending

What is really special?

Open Source gave us:

- Bitcoin/ blockchain
- Big data infrastructure (Hadoop, Spark, Cassandra…)
- Deep learning platforms (Caffe, Theano, Keras, TensorFlow)
- Android
- Firefox
- Linux
- Apache
- Java
- LLMs

TensorFlow learn Keras PYTORCH theano gensim Caffe Chainer SM 將軍

硅谷真正的特别之处?

开源带给我们:

- 比特币/ 区块链
- 大数据基础设施（Hadoop, Spark, Cassandra等）
- 深度学习平台（Caffe, Theano, Keras, TensorFlow）
- 安卓系统
- 火狐浏览器
- Linux操作系统
- Apache网络服务器
- Java编程语言
- 大语言模型（LLMs）

TensorFlow learn Keras PYTORCH theano gensim Caffe Chainer SM 將軍

What is really special?

Talents from all over the world

We find that 44.6 percent, or **223** companies, in the Fortune 500 were founded by immigrants or their children. Of those companies, 101 were founded directly by foreign-born individuals while another 122 were founded by the children of immigrants. Mar 14, 2022

https://www.shrm.org › companies-with-immigrant-roots ▾

New American Fortune 500 in 2019 - SHRM

More Than Half of America's Unicorns Have Immigrant Founders Billion-dollar startups by those born outside the U.S. have grown more than 500 percent since 2018.

https://www.inc.com

硅谷真正的特别之处?

来自世界各地的人才

We find that 44.6 percent, or **223** companies, in the Fortune 500 were founded by immigrants or their children. Of those companies, 101 were founded directly by foreign-born individuals while another 122 were founded by the children of immigrants. Mar 14, 2022

我们发现，财富500强中有44.6%，即223家公司，是由移民或其子女创立的。其中，101家公司是由外国出生的人直接创立的，另外122家是由移民的子女创立的。

https://www.shrm.org › companies-with-immigrant-roots ▾

New American Fortune 500 in 2019 - SHRM

2019年“新美国”财富500强概况 –美国人力资源管理协会（SHRM）

More Than Half of America's Unicorns Have Immigrant Founders Billion-dollar startups by those born outside the U.S. have grown more than 500 percent since 2018.

https://www.inc.com

超过半数的美国独角兽公司是由移民创立的

自2018年以来，由外国出生者创立的十亿美元级初创公司增长了超过500%

What is really special?

Interdisciplinary thinking

- Scientists
- Artists
- Musicians
- Politicians
- Historians
- Entrepreneurs
- Investors
- …

The best colleges for startup founders

1. Stanford **University**.
2. **University** of California, Berkeley.
3. Massachusetts Institute of Technology (MIT)
4. Harvard **University**. Harvard
5. **University** of Pennsylvania. .
6. Cornell **University**
7. **University** of Michigan.
8. **University** of Texas.

businessinsider.com/

硅谷真正的特别之处?

跨学科思维

- 科学家
- 艺术家
- 音乐家
- 政治家
- 历史学家
- 企业家
- 投资者
- …

The best colleges for startup founders

1. Stanford **University**.
2. **University** of California, Berkeley.
3. Massachusetts Institute of Technology (MIT)
4. Harvard **University**. Harvard
5. **University** of Pennsylvania. .
6. Cornell **University**
7. **University** of Michigan.
8. **University** of Texas.

businessinsider.com/

What is really special?

Research = “Moonshot” projects

Let’s put a computer in every room!

Let’s connect the whole world!

Let’s index the whole Web!

Let’s put a computer in a telephone

Let’s reinvent the car

硅谷真正的特别之处?

研究 = “登月计划”项目

让我们使每个房间都拥有一台电脑!

让我们连接整个世界!

让我们给整个世界建立索引!

让我们把电脑放进电话里!

让我们重新发明汽车!

The End… for now

SILICON VALLEY RESEARCH INSTITUTE ARTIFICIAL INTELLIGENCE

（暂时）完结

SILICON VALLEY RESEARCH INSTITUTE ARTIFICIAL INTELLIGENCE