

# 2018 AI 大事件



# 01

# NLP 突破



# 自然语言处理（NLP）的突破

- BERT 模型发布：[https://www.youtube.com/watch?v=t45S\\_MwAcOw&pp=ygUEQkVSVA%3D%3D](https://www.youtube.com/watch?v=t45S_MwAcOw&pp=ygUEQkVSVA%3D%3D)
  - 谷歌推出的 BERT（Bidirectional Encoder Representations from Transformers）通过双向 Transformer 架构和预训练-微调模式，显著提升了语言理解能力，在 11 项 NLP 任务中刷新了记录，成为 NLP 领域的新里程碑。
- 预训练模型的普及：[https://www.youtube.com/watch?v=t45S\\_MwAcOw&pp=ygUEQkVSVA%3D%3D](https://www.youtube.com/watch?v=t45S_MwAcOw&pp=ygUEQkVSVA%3D%3D)
  - ELMo 和 ULMFiT 等预训练模型进一步推动了迁移学习在 NLP 中的应用，标志着 NLP 的 ImageNet 时代的到来。
- 意义：
  - BERT 发布不仅是 NLP 发展史上的里程碑，更是 2018 年 AI 领域最重要的事件。它彻底改变了语言模型的设计思路，推动了 NLP 技术的快速进步和实际应用，为后续的 AI 研究和商业化奠定了坚实基础。



# 算法研究的重大进展

- BigGAN 的惊艳表现: <https://www.youtube.com/watch?v=Unf-7sDUdZ0&pp=ygUGQmInR0FO>
  - DeepMind 的 BigGAN 在图像生成任务中表现出色, 生成的图像逼真度远超以往模型, 展示了生成对抗网络 (GAN) 的潜力。
- 强化学习的突破: <https://www.youtube.com/watch?v=q1o8fNOiWFY&pp=ygUdRG9wYW1pbmUgb3BlbmFpIHJlaW5mb3JjZW1uZXQ%3D>
  - OpenAI 发布了强化学习入门教程 Spinning Up, 并推出了多巴胺 (Dopamine) 框架, 简化了强化学习的研究和应用。



02

# AI 框架发展



# 开源框架的崛起

- PyTorch 1.0 发布: <https://www.youtube.com/watch?v=WMITGIZCFfw&pp=ygUMUHIUb3JjaCAxLjAg>
  - 2018年, Facebook 的 PyTorch 与 Caffe2 整合, 推出了 PyTorch 1.0, 其灵活性和动态计算图机制使其在研究领域迅速普及, 进一步巩固了其在深度学习框架中的地位, 尤其是在研究领域和开发社区中, 成为 AI 领域的重要框架。
- TensorFlow 生态系统扩展: <https://www.youtube.com/watch?v=yjprpOoH5c8&t=81s&pp=ygUKVGVuc29yRmxvdw%3D%3D>
  - TensorFlow 发布了多个重要版本, 增加了即时执行和发行策略等功能, 并推出了 TensorFlow.js、TensorFlow Lite 等工具, 进一步扩展了其应用场景。
- ONNX 的推动: [https://www.youtube.com/watch?v=ThGiPWrsxU0&list=PLhBhgortqAcih4oJ7G\\_kjmUKefNQi2gJY](https://www.youtube.com/watch?v=ThGiPWrsxU0&list=PLhBhgortqAcih4oJ7G_kjmUKefNQi2gJY)
  - Open Neural Network Exchange (ONNX) 标准逐渐成熟, 促进了不同框架之间的模型互操作性。

# 03

## AI 芯片与硬件



# 硬件与芯片的突破

- TPU 广泛应用: [https://www.youtube.com/watch?v=kPMpmcl\\_Pyw&pp=ygUOQ2xvdWQgVFBVIFBvZHM%3D](https://www.youtube.com/watch?v=kPMpmcl_Pyw&pp=ygUOQ2xvdWQgVFBVIFBvZHM%3D)  
<https://www.youtube.com/watch?v=kBjYK3K3P6M&t=14s&pp=ygUHVFBVIHYzIA%3D%3D>
  - 谷歌 TPU（张量处理单元）在2018年进一步普及，支持大规模训练和推理，推动了BERT等模型的突破性进展。谷歌还推出了Cloud TPU Pods，使大规模机器学习训练更加便捷。
  - TPU v3 于 2018 年 5 月在谷歌 I/O 开发者大会上发布，性能较 TPU v2 提升了 8 倍，每个 TPU v3 Pod 的计算能力达到每秒 100 千万亿次浮点运算。
  - TPU v3 Pod 由 1024 个 TPU v3 芯片组成，计算能力达到每秒 100 千万亿次浮点运算，是 TPU v2 Pod 的 8 倍。  
<https://www.youtube.com/watch?v=kNch2v1bh40&pp=ygUIRWRnZSBUUFU%3D>
  - 2018 年，谷歌发布了 Edge TPU，专为 IoT 设备和边缘计算设计。Edge TPU 能够以每秒 30 帧以上的速度处理高分辨率视频，并运行多个先进的计算机视觉模型。



# 硬件与芯片的突破

- NVIDIA RAPIDS 平台: <https://www.youtube.com/watch?v=OTOGw0BRqK0&pp=ygUMbnZpZGIhIERHWC0y>
- DGX-2: 2018 年 3 月, 英伟达发布了 DGX-2, 这是全球首款性能达到 2 千万亿次浮点运算的 AI 超级计算机。它由 16 块 Tesla V100 GPU 组成, 支持大规模深度学习训练和推理任务。
- Tesla V100 32GB: 英伟达推出了内存翻倍的 Tesla V100 32GB GPU, 进一步提升了 AI 训练和推理的性能。
- Jetson AGX Xavier: 英伟达发布了 Jetson AGX Xavier, <https://www.youtube.com/watch?v=XoWW5HiGHsg&pp=ygURSmV0c29uIEF1WCByYXZpZXI%3D> 这是一款面向机器人和边缘计算的 AI 模块, 具备高性能和低功耗的特点。



# 硬件与芯片的突破

- 华为昇腾 Ascend:
  - 2018年，华为发布了基于Ascend（昇腾）系列芯片的Atlas智能计算平台，旨在打造面向端、边、云的全场景AI基础设施方案。
- AI芯片创业热潮：
  - AI芯片创业公司吸引了超过15亿美元的投资，多家初创公司致力于开发专为AI优化的硬件，如神经形态芯片和边缘计算设备。
- 其他：<https://www.youtube.com/watch?v=Efjwy5PEXZw&pp=ygUQQ2FtYnJpY29uIE1MVTEwMA%3D%3D>
  - 寒武纪推出 Cambricon MLU100，首款面向云端智能推理的AI芯片。
  - 中国AI芯片初创公司深鉴科技被美国FPGA巨头赛灵思收购。



# 04

## 产品与解决方案



# AI 产品解决方案

- 智慧城市解决方案: <https://www.youtube.com/watch?v=P3CmPrDvVWQ&pp=ygUNc21hcnQgY2l0eSBBSQ%3D%3D>
  - 华为发布了面向智慧城市的数字平台，整合物联网、大数据、视频云等资源，推动智慧城市建设。此外，地平线发布了基于AI芯片的未来城市解决方案，覆盖多个城市场景。
- 智慧医疗: <https://www.youtube.com/watch?v=qetKUFD4A&pp=ygUTU21hcnQgaGVhbHRoY2FyZSBhaQ%3D%3D>
  - 智慧医疗领域进一步崛起，飞利浦发布了智能互联睡眠呼吸管理解决方案Dream Family，通过可穿戴设备和数字化互联技术实现健康管理。
- AI在传统行业的应用:
  - AI 技术开始渗透到更多传统行业，例如腾讯与罗氏制药合作，探索AI在医疗行业的应用。
  - 此外，英特尔与腾讯优图实验室合作开发了AI摄像机和AI盒子，用于安防领域。



# AI 产品落地应用

- Google Duplex 惊艳亮相: <https://www.youtube.com/watch?v=D5VN56jQMWM&pp=ygUOR29vZ2xIIER1cGxleCA%3D>
  - 在 2018 年谷歌I/O开发者大会上, Google展示了Duplex, 一个能够进行自然语言对话的AI系统, 可以代替用户打电话预约餐厅或理发店。其流畅的对话能力和自然语音合成技术引发了广泛关注。Duplex 展示了 NLP 技术的实际应用潜力, 但也引发了关于AI伦理和透明性的讨论。
- 自动驾驶技术的商业化: [https://www.youtube.com/watch?v=hA\\_-MkU0Nfw&pp=ygUYd2F5bW8gYXV0b25vbW91cyBkcml2aW5n](https://www.youtube.com/watch?v=hA_-MkU0Nfw&pp=ygUYd2F5bW8gYXV0b25vbW91cyBkcml2aW5n)
  - Waymo在2018年推出了商业化自动驾驶服务, 标志着自动驾驶技术从测试阶段迈向实际应用。此外, 百度Apollo无人车在央视春晚上亮相, 展示了中国在自动驾驶领域的进展。自动驾驶技术的商业化落地为未来交通革命奠定了基础。



# 思考



# 总结

- 2018年，AI技术在开源框架、硬件芯片、算法研究和产品应用等方面取得了显著进展。从BERT的突破到 Google Duplex 的惊艳亮相，再到 TPU 和 RAPIDS 的硬件创新，这些事件共同推动了AI技术的快速发展和实际落地。
- 2018 年是英伟达技术全面爆发的一年，从显卡产品线的更新到 AI、自动驾驶和数据中心领域的突破，英伟达展示了其在图形计算和 AI 领域的领先地位。这些发布不仅推动了游戏和内容创作的进步，也为未来的技术发展奠定了基础。





# Thank you

把AI系统带入每个开发者、每个家庭、  
每个组织，构建万物互联的智能世界

Bring AI System to every person, home and  
organization for a fully connected,  
intelligent world.

Copyright © 2024 XXX Technologies Co., Ltd.  
All Rights Reserved.

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. XXX may change the information at any time without notice.



GitHub <https://github.com/chenzomi12/AIFoundation>