

从 Siri 闭门造车到开放机器学习, 详解苹果 AI 实践的三个阶段

人工智能 2019-04-25

在这篇长文里, 我将梳理过去 9 年苹果在人工智能领域的探索, 包括以下几个方面:

乔布斯为何要买下 Siri? Siri 是如何被苹果「毁掉」的? 为什么 2016 年 WWDC 大会是苹果 AI 实践的转折点? Core ML 带来那三重意义? 2018 年苹果 AI 产品和团队调整的意义如何? 2019 年苹果会如何将 iPhone 与 AI 结合起来?

1. AI 大拿加盟

本月初, CNBC 率先报道了 AI 领域知名学者、GANs(对抗式生成网络)提出者 Ian Goodfellow 离开 Google 并加入苹果的消息。

CNBC 发现, Ian Goodfellow 在周四更新了自己的 LinkedIn 个人资料, 显示这位 AI 领域炙手可热的研究者已经在今年 3 月加入苹果, 并成为苹果「特殊项目小组」的机器学习负责人, 该小组直接向苹果 CEO 库克汇报。

媒体对于 Ian Goodfellow 的关注较多, 此君在 2013 年以实习生身份加入 Google, 一年后, 他完成了一篇题为《Generative Adversarial Nets》的论文, 该论文提供了一个计算机如何生成图片的思路——将两个神经网络相互对抗。这个方法

引发行业震动，包括 Facebook AI 首席科学家 Yann LeCun 以及参与 Google Brain 和百度大脑建设的吴恩达，都肯定了该方法对于机器学习发展的的重要意义。

值得一提的是，此前「AI 换脸」技术 DeepFake，其核心技术也是来自 Ian Goodfellow 的发明。

目前尚不清楚 Ian Goodfellow 会给苹果带来哪些变化，考虑其 3 月份刚刚入职，而且 Ian Goodfellow 主要关注在研究领域，因此可以断定，短期内，Ian Goodfellow 的研究不会直接落地到苹果的产品里。

但 AI 之于苹果，已然变得越来越重要了。

2. 乔布斯为何收购 Siri

如果以当下的视角去看，苹果的 AI 探索始于收购 Siri。

作为乔布斯生前收购的最后一家公司，Siri 肩负着乔布斯对于未来交互模式的远见思考。

Human-Computer Interaction (1830s – 2015)



正如上图所示，「互联网女皇」 Mary Meeker 在人机交互历史回顾中指出，几百万年前，人类从声音的交互开始，而文字的发明，则让交互从声音转入文本阶段，于是出现了各种与文本打交道的交互工具。即便如此，人类交流中的 90% 依然还是通过声音来完成，这种更自然也更原始的交互模式在过去半个世纪里并未得到突破性的进展。

自动语音识别 (Automatic Speech Recognition ，简称 ASR) 是一种机器可识别人类说话单词的能力。很长一段时间里，ASR 在准确度上都无法与人类识别相媲美，而在 2010 年，ASR 在识别准确度上迎来一个拐点。

从上图的变化可以看出，2010—2015 的五年时间，ASR 准确度得到大幅提升，这种变化超过了过去 30 年到 40 年的变化。事实上，我们现在已经接近一个时刻：机器对于语音的识别能力即将超过人类。

乔布斯看到了基于声音的交互才是人类天性的体现，而技术进步则帮助人类重新回到了最原始的声音交互时代。

Siri 在这种情况下成为苹果布局声音交互的重要产品，当时，乔布斯曾被问到收购 Siri 是否想与 Google 竞争搜索业务，乔布斯给出的答案很明确：Siri 不是搜索产品，他们要做人工智能。（他原话是「 [Siri is] not a search company. They're an AI company. We have no plans to go into the search business. We don't care about it -- other people do it well.」）

只可惜，乔布斯并未看到搭载 Siri 的第一部 iPhone—— iPhone 4S 的发布仪式。彼时，Siri 给世界带来的惊喜随着一次次的恶作剧般的调侃趋于平淡，人们突然发现，Siri 几乎没有任何用处。她（他/它）常常答非所问，又经常无故跳出（不小心长按 home 键）.....

3. Siri 是如何被玩坏的

如乔布斯当年所言，语音交互正在成为人工智能最有可能爆发的领域。看看市面上的科技公司们，亚马逊、微软以及 Google、Facebook，无一不在这个领域发力。

但苹果显然缺乏应对之策。自乔布斯逝世后，原来的 Siri 团队核心成员几乎都离开了苹果，包括 Siri 早期开发者的 Dag Kittlaus 和 Adam Cheyer 等人又创立了一家 Viv 的公司，其产品可以理解为新一代的 Siri。

2018 年 3 月，付费科技媒体 The Information 发表了一篇长文，详细展示了过去几年 Siri 是如何被苹果玩坏的。

文章提到了两个对立的原因，一方面是负责 Siri 的苹果老员工 Richard Williamson 被指责没有创新意识；另一方面 Richard Williamson 则不断强调 Siri 是一个不成熟的产品，自己作为维护者无法承担其如此大的责任，或者换句话说，这个锅应该由 Siri 创始团队来背。

文章同时还指出，苹果在 2013、2015 年做了两笔与 Siri 相关的收购，希望将两家创业公司 Topsy 的搜索能力以及 VocallQ 的自然语言理解能力整合到 Siri 里，但结果并不尽如人意。Information 获取的消息源称，Topsy 团队的成员表示不愿与 Siri 团队合作，原因是 Siri 的技术陈旧且无法满足当下的用户需求。

随后苹果内部做了一个「艰难」的决定：放弃对 Siri 的修修补补，重写 Siri。这也是一个艰巨的任务，此时的 Siri，已经成为数以亿计的苹果设备里的重要产品，颇具

讽刺的是，或许是因为 Siri 太傻，或许 Siri 的打开率太低，重写 Siri 的过程并没有在如此庞大的用户群体掀起多大波澜。

撇开这些无法确定的事实来看，Siri 无法变得聪明也是苹果内部的必然结果。

首先，缺乏人才。有媒体盘点过 2015 年苹果的挖人名单：

2015 年年初，前苏黎世联邦理工学院自动系统实验室主管 Paul Furgale 加入苹果；7 月，前意大利经典汽车厂商菲亚特克莱斯勒集团质检总监 Doug Betts 加入苹果。8 月，大众前工程师 Megan McClain、特拉斯前工程经理 Hal Ockerse、前高级工程师 Jamie Carlson 加入苹果。9 月，芯片制造商 Nvidia 公司人工智能专家、深度学习软件主管 Jonathan Cohen 加入苹果；10 月，电动摩托厂商 Mission Motors 申请破产，旗下一部分优秀工程师投奔到了苹果旗下；上述几位除了 Jonathan Cohen 之外都与苹果汽车项目相关（当然这个项目如今已经被大量裁员）。直到 2015 年 9 月，苹果才罕见发布招聘广告，大规模招募人工智能和机器学习方面的人才。

其次，苹果公司文化强调的保密原则，根本不适用于机器学习领域。这也可以解释，为何当 Google、Facebook 等公司的机器学习工程师发布论文，而难见苹果公司的原因。

这不禁让人发问：苹果在 AI 领域到底是毫无作为还是隐藏实力？

4. iOS 10 开启了苹果的 AI 新征程

如果将时钟拨回到三年前的 WWDC，苹果的 AI 探险之路也是从这个时候开始的。

与历年 WWDC 类似，这一次同样发布了全新的 iOS——iOS 10。

iOS 10 里，苹果官方的照片应用可以自动识别人脸、物体，并能够相册里的人物自动分类，新加入的「回忆」，则可以通过对照片的位置信息、人脸照片信息、场景信息等聚集在一起，还支持插入音乐。

苹果将上述功能加入到一台运行 iOS 10 的 iPhone 中实属不易，要知道类似的功能在 Google 与微软那里，都是通过云端数据处理后的返回结果。

苹果之所以采用这种方法既有历史原因，也有诸多现实考量。

历史上看，乔布斯时代的苹果对人机交互有着异常的坚持，比如 iOS 曾经异常的简洁性——这也意味着毫无选择性的交互方式。乔布斯也曾表示，如果他来做一款 DVD，那么唯一的按钮就是「烧录」，其他都可以舍弃。正是这样对于简洁性的追求，使得苹果的产品线长期以来在功能上的变化都极其克制。

乔帮主已仙逝多年，iOS 自 iOS 7 开始在交互上开始复杂，但将人工智能纳入到苹果产品的过程依然十分缓慢，摆在苹果面前的现实问题有两个，其一，苹果的云服务不如 Google 和微软，至今苹果 iCloud 的服务还有一部分跑在亚马逊 AWS 上，

这也使得苹果无法像 Google 微软那样通过云端强大的处理性能来处理照片分类、语音识别等。

其二，苹果长期以来的价值观就是对用户隐私的保护。以虚拟助理来说，Siri 早在 2011 年就已经发布，但比起 Google Now、微软的 Cortana 来说，Siri 真的很不智能，原因就在于 Siri 对于用户数据的收集非常少，而另外两家公司的产品则有点「贪得无厌」，尤其是 Google，默认的用户协议里，Google 可以收集用户几乎所有的数据为其所用。

但现实则是，时下流行的深度学习，迫切需要海量数据的「喂养」。这一年的 WWDC 上，苹果也重新定义了所谓「大数据」，正式启动苹果的「Differential Privacy」，这个概念由微软的 C. Dwork 提出，其基本含义就是基于统计学原理，在一个群体层面收集数据，而非像 FB、Google 那样收集个体的数据。

由于苹果的 iMessage、FaceTime 早已部署了端到端的加密技术，因此能够让用户数据以加密的方式上传到苹果的服务器，然后苹果在对于这些群体层的数据进行分析和优化，从而改善用户的体验。

2016 年秋天，随着新 iPhone 设备的发布，iOS、macOS 等系统里也开始集成「Differential Privacy」。而到了这年冬天，在一个行业 AI 大会的非开放会场，新晋上任的苹果公司机器学习负责人 Russ Salakhutdinov 介绍了苹果在自动驾驶、

神经网络、AI 芯片等领域的思考，这是几乎是苹果首次向外界（至少是学术界）透露自己的 AI 布局。

也是在 12 月，Russ Salakhutdinov 还代表苹果宣布了另一项重大决定：苹果将会允许自己的 AI 研究人员公布自己的论文研究成果！

Machine Learning at Apple
Engagement with Research Community

Can we publish?

Yes.

Do we engage with academia?

Yes.



紧接着，苹果在 2017 年 7 月正式发布了机器学习博客，利用这个平台发布一系列围绕 AI 的研究进展，其第一篇博客谈如何基于小规模数据训练，实现算法的提升，其中也提及使用 GANs 的思路。

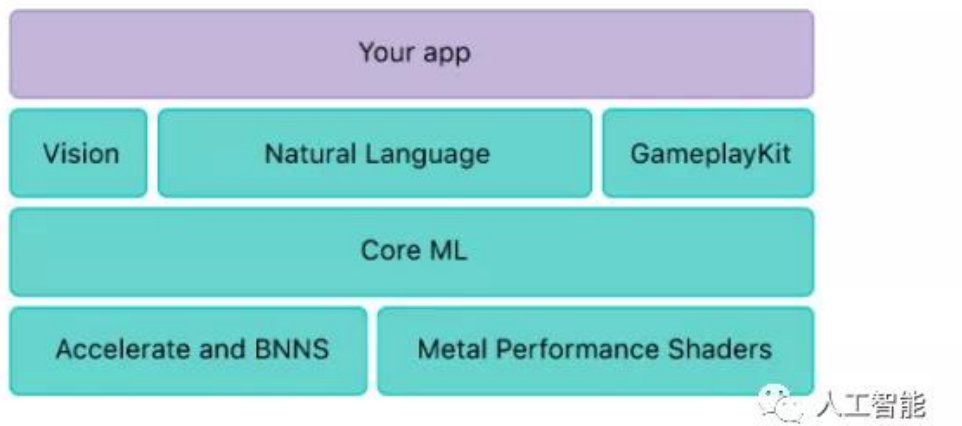
5. 如何让开发者加入苹果 AI 生态？

2017 年可以说是苹果正式向外界展示 AI 能力的开始。这一年的 WWDC 大会上，苹果向世界展示了自己对于 AI 之于手机的新思考：帮助开发者将 AI 能力带入到智能手机里，或者说，帮助开发者将 AI 带入 iOS。

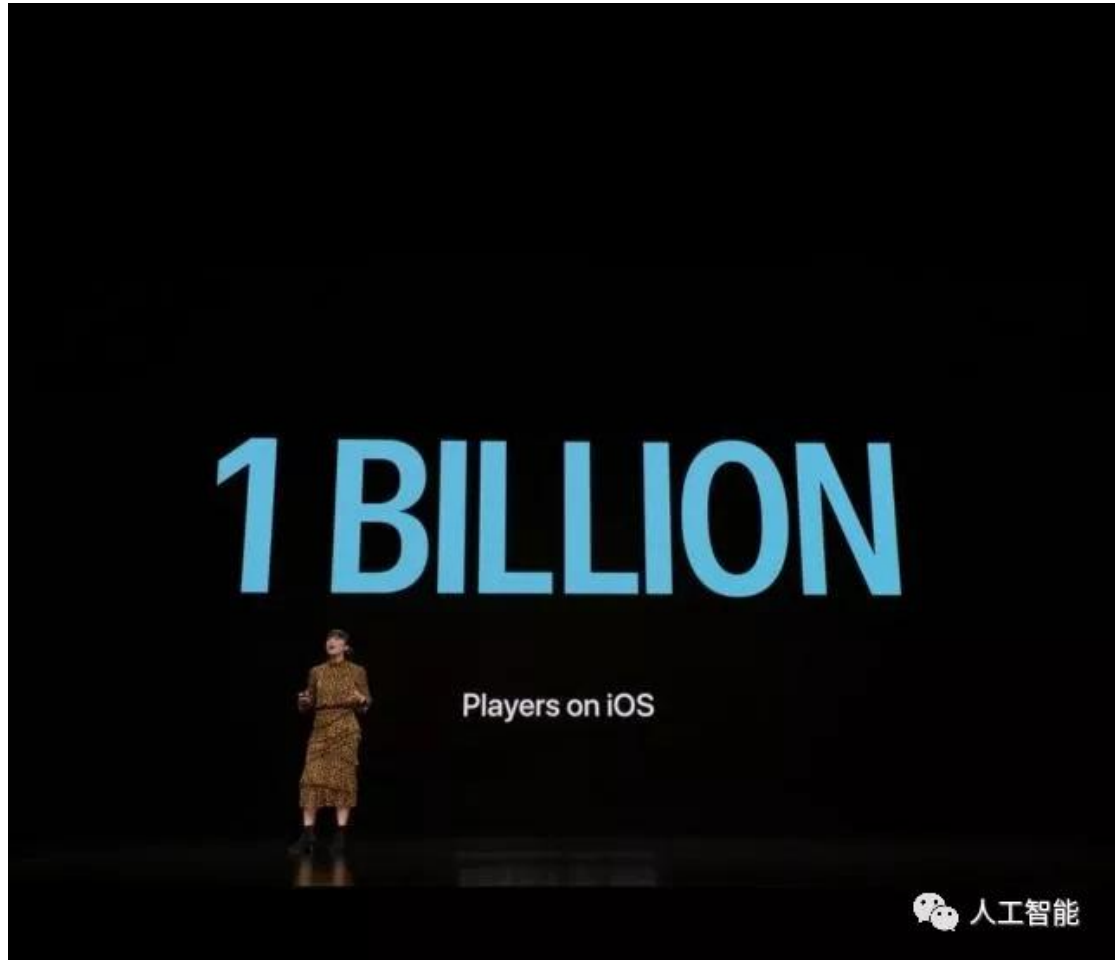
这年的 WWDC 发布的 Core ML 野心巨大。坦率来说，Core ML 并不是第一个将机器学习引入智能手机的开发平台。2017 年的 5 月的 Google I/O 大会上，Google 发布了面向移动设备的机器学习框架 TensorFlow Lite，使得开发者可以方便地在移动端部署 AI 应用。

但苹果 CoreML 的出现，对于行业以及苹果而言有三重意义。

其一，相比于其他的机器学习框架，CoreML 坚持了苹果一贯以来的「设备智能」的原则，从训练到部署，都在设备端展开，这充分展现了苹果的「AI 价值观」，也因为苹果的巨大影响力，能够引发整个机器学习行业对于数据、隐私的关注。



其二，对于开发者而言，苹果的硬件生态依然拥有巨大的吸引力，基于 iOS 体系下的用户基础依然是一个金矿。上月的苹果软件服务发布会上，有一组数字足够震撼，如下图所示，iOS 上的游戏玩家已经突破 10 亿。



这样的硬件生态和用户基础，也将吸引更多开发者加入到苹果对于未来的应用体验定义之中，从 AR 到 AI，App Store 上的应用生态也在变得越来越多元化。

其三，对于苹果来说，正是 App Store 以及上面的开发者成就了苹果当下无可争议的「第一生态」，在当下 iPhone 销量下滑、iPad 无法接力的现实语境，Core ML 也成为提升 iOS 应用体验（讨好消费者）、推动 iOS 应用开发创新（讨好开发者）的重要产品。

Core ML 2

Smaller. Faster. Customizable.



2018 年的 WWDC 上，Core ML 进入 2.0 时代，提供了更小、更快也给更容易定制的机器学习模型，同时也大幅提升了训练、部署速度。可以预见，两个月之后的 WWDC 还会有更多的升级。

6. 苹果开始认真做 AI 了吗？

事实上，2018 年苹果在 AI 领域最大突破则是将 John Giannandrea 招致麾下。这位在 Google 工作 8 年，先后负责机器智能、搜索团队的大拿的到来，标志着苹果开始认真做起了 AI。

根据 TechCrunch 的报道，2018 年 7 月，苹果内部对于 AI 团队做了调整。将 Core ML 和 Siri 团队合并，由 John Giannandrea 统一领导，并直接向库克汇报。

这次人员调整也让这一年 WWDC 上推出的 Siri Shortcuts 具有更令人看好的前景。基于 Siri Shortcuts，第三方应用开发者可以将快速将自己应用的某些功能「输送」给 Siri，实现最快的语言唤醒。

而到了 2018 年 12 月，苹果官方宣布 John Giannandrea 晋升为管理团队成员，其职位名称为「机器学习和 AI 战略高级副总裁」。

Executive Profiles



Tim Cook
CEO



Katherine Adams
Senior Vice President and
General Counsel



Angela Ahrendts
Senior Vice President
Retail



Eddy Cue
Senior Vice President
Internet Software and
Services



Craig Federighi
Senior Vice President
Software Engineering



John Giannandrea
Senior Vice President
Machine Learning and
AI Strategy



Jonathan Ive
Chief Design Officer



Luca Maestri
Senior Vice President and
Chief Financial Officer

在苹果官方新闻稿里，对 John Giannandrea 的工作范围做了这样的介绍：

Giannandrea oversees the strategy for AI and Machine Learning across all Apple products and services, as well as the development of Core ML and Siri technologies. His team's focus on advancing and tightly integrating machine

learning into Apple products is delivering more personal, intelligent and natural interactions for customers while protecting user privacy.

Giannandrea's team is also helping to enhance the developer experience with tools including Core ML and Create ML, making it easier and faster to integrate machine learning into apps.

通过上面的架构图以及对 John Giannandrea 职位范围的整理，可以清晰地看到苹果对于 AI 的重视程度，至此，AI 也成为苹果体系内与零售、服务、软件、设计并列的核心业务。而 John Giannandrea 在学术界以及工业届的号召力，也将进一步带动苹果 AI 领域的人才积累和研究突破，如今，随着 Ian Goodfellow 的到来，苹果在 AI 领域的话语权也越来越大。

那么问题来了，苹果会如何定义手机 AI？

7.2019 iPhone 的 AI 如何体现？

过去五年，正是人工智能再一次成为主流技术潮流的五年，同时也是智能手机由盛而衰的五年，而智能手机与人工智能的结合，也是近两年来的新潮流。

包括苹果在内的智能手机行业，必须面对一个残酷的事实：尽管所有人都知道 AI 会给手机带来变革，但没有人知道到底从何处入手，以及如何去做。

如果结合苹果以及其他公司的实践，或许我可以提供几个值得关注的线索。

第一，操作系统。Google 与苹果，作为全球两大移动 OS 提供商，肩负着如何将 AI 植入 Android 和 iOS 的重任，好在 5 月、6 月的两场开发者大会将让我们见识这两家公司的具体策略；

第二，语音。去年亮相的 Google Duplex 已经扩大到更多设备，这是语音技术的一次巨大突破，相比于更漂亮的实验室数据，基于实际应用场景的 Duplex 或许并不完美，但却提供了语音技术应用的新场景。

与此同时，苹果的 Siri Shortcuts 的确在一定程度上提高了 iOS 的自动化能力，但还没有完全激发出开发者、消费者的使用热情，随着苹果在越来越多设备里加入可随时唤醒的 Siri，Siri Shortcuts 可以提供更多应用空间。

第三，图像/摄像头。2018 年，Google 和华为分别从软件（计算机视觉）、硬件（面积更大的传感器）定义了两个不同的手机拍照路线（详见这期会员通讯），尽管苹果也在 2018 年用智能 HDR 予以还击，但在夜景上还是完败。

这也让外界好奇 2019 年 iPhone 的拍照能力是否可以实现新的突破，不管是硬件还是软件，苹果的确需要拿出一个好相机，挽回正在丢失的用户。

- END -

