

国家制造强国建设战略咨询委员会译著系列丛书

人工智能

改变世界，重建未来

[美] 卢克·多梅尔 (Luke Dormehl) ◎著
赛迪研究院专家组◎译

THINKING MACHINES

The Inside Story of Artificial
Intelligence and Our Race to Build the Future



扉页

人工智能

[美] 卢克·多梅尔 著
赛迪研究院专家组 译

中信出版社

目录

推荐序

前言

第一章 表现不俗的传统人工智能怎么了

未来已来

人工智能的三位卓越先驱

达特茅斯会议

传统人工智能的黄金时期

太空机器人沙基

“中文房间”实验

莫拉维克悖论

转变目标

专家系统

谷歌诞生

第二章 以自主学习的方式创建新的人工智能

自主学习的重要性

神经科学之父

以神经元模型为基础的感知器

关于感知器的争论

霍普菲尔德网的兴起

神经网络的守护神

联结主义者

欢迎来到深度学习领域

人工智能新主流

第三章 万物互联的智能时代已经来临

智能设备成为现代生活的必需品

会思考的事物

“控制论”简史

早期的自动机器人

计量生物学上的记录者

未来家庭

机器学习的普及

智慧城市

做好互联网消失的准备

智能设备存在的问题

第四章 人工智能助手如何为我们效劳

打败图灵测试

人工智能助手的兴起

多功能的人工智能助手

从被动人工智能助手到主动人工智能助手

数字民主

与人工智能助手相爱

人工智能的人性化

合格的心理治疗师

出门请带上它们

第五章 人工智能正在改变就业市场

亲爱的沃森

一个技术性失业的世界

技术更新换代的积极意义

新工作机会的产生

土耳其机器人的复仇

人工智能中的人类元素

第六章 人工智能真的具有创造力吗

机器人的创造力

什么叫创造力

“重生”的甲壳虫乐队

天才的灵光一现

人工智能能否成为发明家

洛夫莱斯测试

祝贺沃森大厨

第七章 意识上传实现人类永生

利用机器实现死后永生

个性捕捉

延长人类寿命

模拟神经元

绘制大脑

下一个大事件

意识上传

第八章 人工智能关乎未来的一切

末日的开端

奇点来临

狭义与广义的区别

人工智能带来的风险

“黑箱”风险

我们无法起诉机器人

机器人学的三大法则

人工智能的权利

结论 当机器变得更加智能

致谢

译后记

谨以此书献给我的朋友
亚历克斯·米林顿 (Alex Millington)

推荐序

人工智能是研究人类智能行为规律（如学习、计算、推理、思考、规划等），构造具有一定智慧能力的人工系统，以完成往常需要人的智慧才能胜任的工作。中国、印度和希腊哲学家在公元前近千年已经提出关于形式推理方法的思考，中国、埃及和希腊的能工巧匠也曾尝试制作精巧的歌舞人偶。20世纪30年代末到50年代初，香农、图灵和冯·诺伊曼等奠定了现代信息论和计算机科学的理论基础，图灵还提出了著名的图灵测试：“如果一台机器能够与人类展开对话而不能被辨别出其机器身份，那么称这台机器具有智能。”预言了创造出智能机器的可能。随着计算机科学与技术的发展，1956年夏，美国达特茅斯会议上，人工智能概念和研究领域被正式提出和确立。人工智能涉及计算机技术、控制论、信息论、语言学、神经生理学、心理学、数学、哲学等多学科领域的交叉与融合，其概念与内涵也在随着相关学科和应用领域的发展而持续变化。当前，人工智能主要是指那些通过软硬件结合可以达到人类事件处理能力的机器智能。它已经成为被高度关注、广泛使用、快速发展的科技领域。

从发展阶段看，20世纪中叶至今，人工智能虽快速发展，但仍处于单台或局域成组机器处理简单任务的弱人工智能阶段。依靠全球网络、大数据和云计算能力，终端云端协同实现具有人类思考方式和多样复杂任务处理能力的强人工智能还处于发展初期。近年来，人工智能技术呈现加速发展态势，其发展水平也达到了前所未有的高度。从几年前打败国际象棋大师的IBM“深蓝”，到现在打败李世石的谷歌AlphaGo、无人机与无人驾驶汽车等；从依靠海量计算能力的强计算弱智能，到依靠深度学习大数据、云计算的网络协同强智能，人工智能技术和发展模式正在发生深刻而本质性的转变。

人工智能已经在深刻而广泛地改变着人们的生活。从手机智能系统、机器视觉到图像识别，从嵌入软件到智能控制，从大数据采集到分析理解等，都渗透着人工智能的创新应用。人们依靠智能导航出行，通过语音与机器互动，应用智能工具搜索智识信息……我们已自觉或不自觉地处于人工智能的环境中。许多过去被视为只有依靠人的智力才能胜任的复杂工作，也开始出现被机器完美替代的可能。今天，我们要选择的已经不是是否接受人工智能，而只能是选择以科学理性的方式应用人工智能造福于人类，并共同应对人工智能应用可能带来的法律伦理、公共和国家安全等新的挑战。

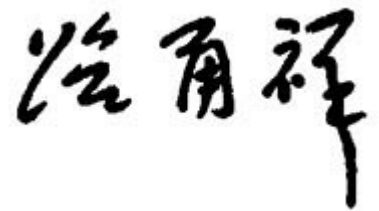
人工智能在生产领域中的应用，为新时代的产业结构升级变革提供了新的机遇。在这万物互联的时代，人工智能以增强机器智力的方式替代人或与人协同工作，促进网络智能设计制造与服务，将大幅提升劳动生产率，重塑产业链与价值创造和分配方式。灵活多样的众创、众筹能力与人工智能技术相结合将为节能减排、绿色发展，提质增效、产业升级创造新的机会。

本书从人工智能的萌芽到人工智能未来发展方向给予了系统解读。围绕人工智能在生产、生活中存在的方式和可能发挥的作用，展示了人工智能诸多应用场景和技术创新发展的方向。虽然当前的人工智能还不具备自主创造能力，但它对生产生活引发的深刻变革，它所展示的勃勃生机将激励政产学研用金各方去协同创新发掘。伴随着众多人工智能相关书籍、刊物和网络新媒体的广泛传播，相信更多的人将对人工智能产生更

大的兴趣、给予更大的关注，将会对它有更深刻的认知，必将会有更多的人才、资源投入人工智能领域。

我国从20世纪70年代末开始启动人工智能研究，并将其列入国家高技术研究计划。进入新世纪后，国家更加重视人工智能的发展，不断加大对相关技术研究和应用发展的支持力度，计算机视觉、语音识别和语言技术处理等人工智能技术快速发展，已进入国际先进行列。《中国制造2025》提出以加快新一代信息技术与制造业的深度融合为主线，以推进智能制造为主攻方向，通过智能技术创新和智能制造的应用发展，适应个性化、定制式设计制造服务，促进中国制造向全球制造产业链、价值链的中高端拓展，实现由大转强。

国家制造强国建设战略咨询委员会是实施制造强国战略的咨询机构，始终把研究推动智能制造发展作为一项重要的职责。人工智能是智能制造的基础核心技术。为跟踪国际人工智能的发展趋势，为读者提供人工智能的最新发展动态，我们组织有关专家翻译出版本书。值得指出的是，知识网络时代，在全球多样化、个性化、定制式市场需求的推动下，在万物互联、大数据、云计算等信息·物理·计算环境的支持下，人工智能技术与制造服务深度融合创新发展与产业化速度将愈加快捷迅猛。国家制造强国建设战略咨询委员会将继续发挥自身的专业优势，通过组织编译国外相关优秀著作，努力为社会各界特别是制造业领域提供更多的关于人工智能及智能制造的前沿研究成果和创新发展的信息。

A large, bold, black handwritten signature in Chinese characters, reading '郭君强' (Guo Junqiang).

2016年9月

前言

计算机的那只红眼睛能洞悉一切，它目不转睛，也不眨眼。说话的语气从容平静。

“嗨，”它说，“我们还要继续比赛吗？”

它指的是那天早些时候开始的棋局。但是你真的没有心情继续这盘棋，这不是因为计算机总是赢你（尽管这是事实），事实上，你之所以恼怒，是因为它犯下了一个令人费解的错误，这个错误与重要机器的关键部件可能出现的故障息息相关，甚至关系到你能否继续生存下去。尽管计算机坚持称这些部件已经出了故障，但你亲自检查后，却发现它们依然处于良好的工作状态。现在你想知道真正的答案。

“是的，我知道你发现它们还在正常运转，但我可以向你保证，它们将要出故障。”计算机一边说，一边试图用那种惯常的、毫无情绪的声音安抚你。

你感觉自己已经怒火中烧了。

“够了，根本不是这么回事，”你气急败坏地嚷道，“这些零件完全没有问题，我们在百分之百的过载条件下进行过测试。”

“我并不是质疑你的话，但是这的确是有问题的。”计算机继续说道。

随后它又说了6个字，你知道这些话是对的，但这绝对不是你现在想听到的话：“我不可能出错。”

影迷们可以很快发现，这是斯坦利·库布里克（Stanley Kubrick）的经典电影《2001：太空漫游》中的一幕，讲的是有意识的计算机流露杀机并且企图杀死太空船中全部船员的故事。

多年以来，这曾是我们看待人工智能的方式：它是一种若隐若现的威胁，但只安全地存在科幻小说中。

但现在已经不再是这样了。

今天，人工智能的梦想早已走出了影院与科幻小说，进入了人们的现实生活。人工智能可以驾驶汽车，交易股票，仅通过观看视频网站YouTube就能学会施展复杂的技能，在不同语种间任意切换，它比我们更加精确地识别人脸，以及创立原始的假设从而发现治愈疾病的新药。而这仅仅是个开端。

《人工智能》讲述的是人工智能的进化之旅及其对人类的意义。在阅读这本书的过程中，我们将遇到假装变童癖的计算机、会跳舞的真空吸尘器、会下棋的人工智能以及旨在与死后的人进行对话而被上传了个人意识的计算机。

这是关于我们如何设想未来的故事，以及在一个充满技术的世界，我们在面临计算机智能不断加速发展的情况下如何为人类塑造自己的角色的故事。这是一个关于创造性的本质、未来就业以及当所有知识成为数据并且可以电子存储时，即将发生什么的故事。当我们制造的机器比我们更加聪明，我们将要做什么，人类

现在还具备什么优势，以及你和我是否也是一种会思考的机器。

1950年，英国数学家和计算机科学家先驱人物艾伦·图灵（Alan Turing）曾预测，到20世纪末，“词语的使用与一般的教育主张将发生巨大改变，届时人们谈论机器思维时，不会有人对此进行反驳”。

就像许多未来学家对技术的预测一样，图灵对未来非常乐观，但并不过度乐观。在21世纪初期，与“智能”相关的技术以及“机器学习”已经是很常见的话题，而这些概念在图灵时代曾使许多人感到匪夷所思。

如今，人工智能作为一门学科已经诞生60年，它依然是人类最重大、最具雄心的项目：努力建造真正的会思考的机器。随着时光的流逝，技术人员离实现这个目标越来越近，未来的一抹曙光也正在天际迅速明朗。

《人工智能》讲述的正是这种熠熠生辉且近在咫尺的未来，隐藏在角落里的变化，以及它们将如何永久性地改变我们的生活。

第一章 表现不俗的传统人工智能怎么了

第一章

表现不俗的传统人工智能怎么了

[2]的“土星5号”运载火箭船尾的复制品。该运载火箭不久后在阿波罗太空计划中协助人类成功登上了月球。在港务局大楼，人们排队观看壮观的纽约世界贸易中心“双子塔”的建筑模型。两年后的1966年，世界贸易中心破土动工。

如今，大部分上述进展会让人不禁缅怀过去的科技进步。在“更大、更高、更重”的雄伟壮阔中，这些进步完成了与过去时代的对话。那个时代对当年的世博会参观者还是个未知数，而如今却即将走向终结。工业时代逐渐淡出，被个人电脑所引领的信息时代所取代。对于那些1964年以后出生的孩子来说，数字将代替他们工程师梦想中的铆钉。苹果公司的史蒂夫·乔布斯在纽约世博会举办那年只有9岁，谷歌联合创始人拉里·佩奇和谢尔盖·布林在纽约世博会举办近10年后才出生，而Facebook（脸谱网）的马克·扎克伯格则在纽约世博会举办20年后才来到这个世界。

最后，法拉盛草原可乐娜公园中最前卫的区域是IBM（国际商用机器公司）的展区。IBM在1964年世博会上的任务是在航天火箭和核反应堆等更负盛名的奇观中，强化公众对于计算机（更具体而言是人工智能）的认知。为此，IBM选择世博会作为全新System/360系列大型计算机的推介平台——据说，该系列计算机的性能非常强大，足以作为智能电脑首台原型机的基础。

IBM在世博会上最引人瞩目的展示当属美国夫妻档设计大师查尔斯（Charles）和蕾·伊姆斯（Ray Eames）设计的巨蛋形展馆。这颗飞艇般大小的“巨蛋”耸立在一片由45棵32英尺高的几何形金属片树木所构成的“森林”中；14 000片灰色和绿色树脂玻璃叶片展开后形成一个一英亩的硕大穹顶。兴奋不已的世博会参观者需要乘坐一台专用的液压电梯进入“巨蛋”，然后可以坐在一个高科技放映厅内观看一段关于人工智能未来的视频。一位激动不已的评论者借用IBM从19世纪20年代就开始使用的品牌口号“Think”，这样写道：“观看着，思考着，并且惊叹着人类及其所创造的机器的奇妙智慧。”

IBM在这场盛会上展示了多项令人印象深刻的技术。开创性的手写识别计算机便是其中之一，并被世博会官方手册称为“光学扫描和信息检索”系统。在该系统的演示中，参观者可以在一张小卡片上亲笔写下任意一个1851年后的历史日期。然后，卡片被插入“光学字符读取器”，并在该设备中被转换成数字格式，然后再转发给当时最先进的IBM 1460计算机系统。随后，计算机磁盘上一个大型数据库中存储的当日重大新闻事件被打印在一张穿孔卡片上，颇有纪念意义，也让使用者惊叹不已。一张保存下来的穿孔卡片上如是写道：

《纽约时报》在你所写下的日期报道了以下新闻事件：

1963年4月14日：30 000名朝圣者在复活节造访耶路撒冷；教皇约翰·保罗为人类的真理和爱祷告。

[3]之外的纽约金士顿一台强大的IBM大型计算机之间的专用数据连接，所有繁重的计算工作都由这台计算机来完成。

机器翻译对计算机的清晰思维如何引领我们走向乌托邦进行了简明且出色的概括。政治家们可能无法结束冷战，但他们也只是普通人——有着人类可能具有的所有缺点。议员、将军，乃至总统都严重缺乏当时学术界刚刚开始提出的“机器智能”。智能机器能否比人类做得更好？在1964年的世博会上，激动万分的公众一睹科研人员最乐观的愿景。人工智能给出的建议是，只要能够在机器中尽量利用和复制人脑的最深层奥秘，那么全球和谐终会得以保障。

世博会的官方标语“理解构筑和平”是这一切的最佳总结。

未来已来

1964年纽约世博会上凸显了人工智能愿景的两种表现。第一种表现是人们对于等待着他们的未来有多么乐观。尽管冷战威胁时隐时现，但在19世纪60年代，人们却在许多方面都表现出惊人的乐观态度，毕竟在60年代初，约翰·肯尼迪总统宣布：在10年内，人类将登上月球。而在这10年的结尾，人类确实做到了。如果连这都能做到，那么解析和重建意识看似更容易实现。政治学家、人工智能创始人之一赫伯特·西蒙（Herbert Simon）于1960年宣称：“我们离复制人类大脑解决问题的能力与信息处理的能力已经不远了。”可能是借鉴了肯尼迪向对手下战书的方式，他随意地加上了自己估计的时间：“如果我们在未来10年还未实现这一愿景，那才令人惊讶。”

西蒙的预测后来无望地落空了，但事实证明，有关纽约世博会的第二种表现就是：IBM当时并没有错。尽管我们坚持认为人工智能尚未普及，但1964年令人惊掉下巴的三项技术如今已司空见惯。光学扫描和信息检索已演变成互联网，让我们在一瞬间获得的信息比原本一生可能获得的信息还要多。虽然我们仍然无法预见未来，但是基于用户不断预测我们可能感兴趣的新闻故事、书籍或歌曲所生成的巨大数据集，我们正在人工智能方面取得巨大的进展。这种预测的连通性不仅限于传统意义上的计算机，而且还通过大量智能传感器和设备嵌入到了周围的设备、车辆和建筑中。

福尔摩斯木偶剧旨在展示通过计算机逻辑可以完成多种任务的方式。我们的计算机逻辑方法已经在某种程度上做出了改变，但是在查看两张照片并且判断照片中是否为同一人时，福尔摩斯可能还是会被精确度高于人类的现代人脸识别算法所折服。福尔摩斯的创造者、作家柯南·道尔（Conan Doyle）是一位毕业于爱丁堡大学（如今是英国顶尖人工智能院校之一）的专业医生。如果他看到疾病诊断效率高过许多人类医生的人工智能Modernizing Medicine（现代化医疗），也很可能会对此赞叹不已。

最后，世博会上神奇的机器译员已发展演变成为今天我们最为熟悉的谷歌翻译：一项拥有惊人的高准确率、涵盖58种不同语言（共计3 306项独立翻译服务）的免费机器翻译服务。如果说世博会设想的是俄语与英语之间的即时翻译，那么谷歌翻译则更进一步，它能够提供冰岛语和越南语或者波斯语和意第绪语等此前受到局限的语种之间的翻译。凭借云计算，我们甚至不需要固定的主机来执行计算，只需要比一副纸牌还小、被称为智能手机的便携式计算机即可。

在某种程度上，由于所有这些技术现已存在，而且不仅存在于研究实验室中，而是几乎每个想要使用它们的人都能唾手可得，因此很难说我们还没有进入人工智能时代。与我们在生活中为自己制定的许多虚无缥缈的目标一样，人工智能成为计算机科学的“梦幻岛”：一种总是潜伏在下一个转角的虚空“假设”。

但是，如果因此就认为人工智能自60年前诞生到今日一直在直线发展就大错特错了。在开始普及目前推动人类重大进步的大规模“深度学习神经网络”前，我们必须对人工智能的历史有更多的了解。

我们也必须了解，长期以来，在走上错误的道路前，人工智能是看上去发展得很顺利的。

人工智能的三位卓越先驱

数千年来，人类一直梦想着为无生命的物体注入生命。但当谈到人工智能在普通人中的普及时，就得从世界上第一台可编程计算机——重达30吨的巨型计算机ENIAC开始说起。在1945年“二战”结束后仅6个月，宾夕法尼亚大学就制造出了兼具电子数字集成器和计算器功能的ENIAC。这台计算机耗费了500 000美元的美国军事基金，其运算速度比作为竞争对手的其他机电机器快了大约1 000倍。这台机器及其所代表的理念令媒体极为感兴趣。他们将这台计算机称为“巨脑”。

建造这样一个“巨脑”的想法令人们为之神往。“二战”结束以前，“电脑”专指在记账等领域进行计算的人。忽然之间，电脑所代表的不再是人了，而是装有电子管和晶体管的机器，而且这些机器的计算速度甚至比最有计算天赋的人还要快得多。“二战”及其余波激发了人们对认知心理学的兴趣。仅在战时，美国心理学会的会员人数就从2 600人增至4 000人。到1960年，仅15年间，会员人数就激增到了12 000人。认知心理学研究人员将人类大脑想象成一台机器，复杂行为可以通过这台机器表现为多重简单响应的综合结果。认知心理学家关注的是人类行为中切实可观测到的事物，而非在无法证实的“心理实体”上浪费时间。“行为主义”等领域因此应运而生，因老鼠实验而闻名的著名心理学家伯尔赫斯·弗雷德里克·斯金纳（B. F. Skinner）称行为主义为“行为的技术”。

工程师们之前一直对心理学形而上学的层面避而不谈，但是他们都被一个概念深深吸引，即人脑可能成为电脑。他们同样开始醉心于研究记忆、学习和推理，许多心理学家都将其视为人工智能的基础。激动人心的是，他们同样意识到，与人类相比，计算机具有潜在优势。例如，ENIAC每分钟能够计算多达20 000次乘法，这样的计算能力着实令人惊讶。人类的记忆并不可靠，而相比之下，一台每微秒访问数千条项目的机器更能凸显出优势。

专门介绍现代计算诞生的书籍有很多，其中三名作者脱颖而出，他们为我们现在所熟知的人工智能领域奠定了哲学和技术基础。他们就是：约翰·冯·诺依曼、艾伦·图灵和克劳德·香农。

冯·诺依曼是土生土长的匈牙利人，1903年出生在布达佩斯的一个犹太银行家家庭。1930年，冯·诺依曼来到普林斯顿大学担任数学教师，三年后成为新成立的普林斯顿高等研究院的6位教授之一，并担任这一职位直至去世。无论以什么标准来衡量，冯·诺依曼都堪称才智惊人。相传，他在6岁时就能计算8位数的除法。“二战”期间，冯·诺依曼在洛斯阿拉莫斯（Los Alamos）参与了曼哈顿计划（Manhattan Project），他的一项艰巨任务就是计算氢弹爆炸的精确高度，即在什么位置爆炸能够造成最大的破坏效果。冯·诺依曼对计算领域做出的主要贡献是，提出在计算机存储器中设置计算机程序存储的理念。实际上，冯·诺依曼是第一个将“记忆”这一人类专用术语应用于计算机的人。与同时期的其他人不同，冯·诺依曼并不认为计算机能够像人类一样思考，但是他却创建了与人类面相学相并行的学说。他在一篇文章中指出，计算机零件“相当于人类神经系统中的神经元。感觉神经元（传入神经元）和运动神经元（传出神经元）之间的共同点仍有待探讨”。其他人会很高兴接过这一挑战。

艾伦·图灵是英国的数学家和密码专家。“二战”期间，他在英国布莱切利公园密码破译中心担任政府密码学校的领导，并提出了许多破译德军密码的方法，其中最著名的就是发明了能够破译英格玛（Enigma）密码机设置的机电装置。这使他在破译截获信息方面发挥了重要作用，最终帮助同盟国战胜了纳粹。图灵对“思考的机器”这一想法十分着迷，并设计了著名的图灵测试，我们在后面的章节会进行详细探讨。图灵小时候非常喜欢一本书——《每个孩子都应该知道的自然奇观》（*Natural Wonders Every Child Should Know*），作者艾德文·坦尼·布鲁斯特（Edwin Tenney Brewster）称：“这本书试图引导8—10岁的孩子对如下问题自问自答：‘我与其他生物有什么相同之处？我们之间又有哪些不同？’”布鲁斯特在书中写道：

当然，这就是一台机器。虽然它比以往任何机器都复杂得多，但归根结底也只是一台机器。人们曾把它比作一台蒸汽机，但那时我们对它的了解远没有现在多。它实际上是一台燃气机：就像汽车、摩托车或飞行器的引擎。

图灵的一个重要理念就是通用图灵机（Universal Turing Machine）。与其他拥有单一功能的电脑相比，图灵诠释了如何通过阅读磁带上的分步说明来使用电脑去完成多个任务。他写道，电脑“实际上能够作为其他机器的模型”。这就意味着没必要在完成每项任务时使用不同的机器。图灵指出：“我们无须再为不同的任务制造不同的机器，只需对通用机进行‘编程’即可完成这些任务。”

图灵猜测，其中一项任务就是模仿人工智能。他在《智能机器》这篇文章中将在机器中复制人工智能称为“一项特殊的挑战”，因为当时的计算机还存在一定的局限性。他认为“人脑的存储容量大约为100亿二进制数字，但其中大部分存储容量被用于记忆视觉印象，或以其他方式浪费掉了。人们有理由希望通过数百万字节的计算机内存而在人工智能方面取得实际的进展”。

人工智能的另一位创始人为克劳德·香农，也就是今天的“信息理论”之父。香农出生于1916年，是三位创始人中最年轻的一位，他为计算领域做出的最大贡献是提出了晶体管的工作原理。晶体管是数十亿个微小的开关，电脑就是由这些开关组成的。算法就是指令序列，通过打开或关闭晶体管向电脑传达指令。香农认为，电脑在接到其他晶体管指令后打开或关闭特定的晶体管，通过这一方式进行基本的逻辑分析。他指出，如果在打开2号和3号晶体管的同时打开1号晶体管，那么这就是一个逻辑运算。如果在打开2号或3号其中一个晶体管时打开1号晶体管，这就构成第二个逻辑运算。如果在2号晶体管关闭的情况下打开1号晶体管，这就是第三个逻辑运算。就像简单的口语词汇一样，所有计算机算法都可以归为以下三种状态之一：与、或、否。香农将这些简单的状态合并成复杂的指令序列，并且建议执行复杂的逻辑推理链。

达特茅斯会议

上述三人之中，只有香农继续探索并积极推动人工智能正式成为一个学科。遗憾的是，图灵和冯·诺依曼分别在41岁和53岁时英年早逝，但他们的理念和影响却延续至今。艾伦·图灵是同性恋，在那个时代的英国，同性恋是违法的。尽管图灵的密码破译工作对英国战胜纳粹德国起到了至关重要的作用，但功不抵过，1952年他被指控严重猥亵。作为惩罚，图灵必须在坐牢或化学阉割中做出选择，而他选择了后者。两年后，图灵吞下涂有氰化物的苹果自杀。2013年，他的罪行得到英国皇家的赦免，并且有人提议制定“图灵法”，以赦免那些在历史上被指控犯有猥亵罪的同性恋。

冯·诺依曼死于癌症。他之所以患上癌症，很可能是因为他曾参与原子弹项目。与冯·诺依曼关系密切的一位同事在《经济学杂志》（*Economic Journal*）上刊登的讣告中称，他的思维“十分特别，有些人（同样是杰出的科学家）不禁扪心自问：他是否将人类心智发展带入了一个新阶段”。

随着两位创始人的离去，其他年轻研究人员接过了建造思考的机器的重任。第二波人工智能研究人员第一次正式为该领域命名，将其确定为一个独立的学科。1956年夏天注定是一个不平凡的夏天：埃尔维斯·普雷斯利的电臀舞（hip gyrations）震惊了观众；玛丽莲·梦露嫁给了剧作家亚瑟·米勒；德怀特·艾森豪威尔总统将“我们信仰上帝”确定为美国法定箴言。关于人工智能的首次正式会议也在这个夏天召开。研讨会持续6周，众多学科中最杰出的学者在新英格兰地区汉诺威市达特茅斯学院的269英亩土地上齐聚一堂。除克劳德·香农以外，大会还有另外两名年轻的组织者，他们分别是约翰·麦卡锡（John McCarthy）和马文·明斯基（Marvin Minsky），他们都在蓬勃发展的人工智能领域发挥了重要的作用。

他们写道：“人们将在一个假设的基础上继续进行有关人工智能的研究，那就是学习的各个方面或智能的各种特性都能够实现精确描述，以便我们能够制造机器来模仿学习的这些方面和特性。人们将尝试使机器读懂语言，创建抽象概念，解决目前人们的各种问题，并且能自我完善。”

然而，他们的雄心壮志和信念都太过绝对，而时间却有些紧迫。他们在达特茅斯会议的提议中指出：“我们认为，如果精心挑选一组科学家，并让他们专心研究人工智能，那么仅需一个夏天我们就能够在这些问题上取得重要进展。”

毋庸置疑，实际花费的时间肯定要比这更长。

传统人工智能的黄金时期

随着感兴趣的研究者越来越多，人工智能开始细分为不同的领域，这意味着人们探索的范围开始变得更加广阔。从某种意义上来说，这一现象是大势所趋。在达特茅斯会议上，人们发现，就连为各自的新领域取一个大家都认同的名字都十分困难。约翰·麦卡锡力荐“闪光探测人工智能”，但却并不能令其他人信服。另一位名为亚瑟·塞缪尔（Arthur Samuel）的研究者认为这个名字听起来很“做作”，而艾伦·纽厄尔（Alan Newell）和赫伯特·西蒙则立即又将其作品重新命名为“复杂信息编程”。

人工智能很快就细分为不同专业，仅从1958年在英国密德萨斯的特丁顿国家物理实验室召开的“思维过程的机械化”会议就能看出这一点。达特茅斯会议召开仅仅两年后，人工智能就已经细分为多个领域，包括人工思维、字符和模式识别、学习、机器翻译、生物学、自动编程、产业规划和办公机械化。

接下来的一段时间成了传统人工智能的黄金时期。新的领域层出不穷，研究者们不断取得重要进展，思考的机器也呼之欲出。由于有美国国防部高级研究计划局（ARPA）等政府机构的资金支持，巨大的投资使研究没有受到什么干扰。1963年6月，国防部高级研究计划局为麻省理工学院提供了220万美元的拨款，用于研究“计算机辅助识别”。据拨款受益人称，国防部高级研究计划局采取一次性付款，之后也未太过关心这笔款项是如何使用的。而类似的事情不止一次发生。

约翰·麦卡锡将这一现象称为人工智能的“看！不用手也行！”时代，这一时代成功唤起了青少年的热情和勇气。在之后的十多年里，研究者们应对了众多挑战，并反驳了“没有机器能够做到X”这样一个假设。X所代表的含义取决于提问者是谁。一位研究者曾经编写了一套国际跳棋程序，该程序能够打败大多数业余选手，包括编程者自己。另一项突破是感知人工智能，它能够在一个人使用机器人手的桌面上重新排列不同形状的彩色模块：这无疑是计算机视觉方面的一项壮举。一个名为SAINT的项目能够解开大学一年级课程水平的微积分中的积分问题。另一个名为ANALOGY的项目还能够解开智力测验中的几何问题，而STUDENT则能够破解复杂的代数难题，例如：“如果汤姆的客户数量是广告数量20%的平方，他一共做了45个广告，那么请问汤姆一共有多少个客户？”

计算推理领域有一个特别引人注目的例子，那就是Sir（语义信息检索）项目。Sir不仅能够理解英语句子，甚至能够分析宾语之间的关系，这和人工智能有些类似。实际上，这一“知识”依赖于预先设定的模板，例如，A是B的一部分，然后用名词来代替变量。然而，正如马文·明斯基所说的那样，人们可以采用类似的方法解决一系列其他问题。有一个有关Sir的典型示例，案例中未加粗的句子为用户输入的文本，加粗的句子为机器给出的答案：

每个男孩都是人。

我知道。

手指是手的一部分。

我知道。

每个人都有两只手。

上述句子表达不清楚。请改述。

每个人都长两只手。

我知道。

约翰有几根手指？

上述句子表达不清楚，但是我猜（有）指的是（长有）。我不知道手指是否是约翰身体的一部分。

约翰是个男孩。

我知道。

约翰有几根手指？

上述句子表达不清楚，但是我猜（有）指的是（长有）。每只手长有几根手指？

每只手都长有5根手指。

我知道。

约翰长有几根手指？

答案是**10**根。

他们构思人工智能和智能行为的方法就是将这些看似毫不相干的表达联系起来。他们自上而下审视了人工智能，我们如今称之为符号人工智能或传统人工智能。对传统人工智能研究者来说，人类通过构建内部的符号化表现形式来认识世界，而所有人工智能都以人类认识世界的能力为基础。我们可以创建规则来落实这

些概念，而捕捉日常知识能够使这些概念变得规范。如果将人脑看作一台电脑，这意味着我们操控的每一种情况都依赖于内部计算程序，这一程序一步步告诉我们如何完全依据逻辑来进行某项操作。如果的确如此，那么这些世界各地机构纷纷采用的规则同样适用于计算机。

这在当时听起来十分容易，不久以后，情况也确实如此。

太空机器人沙基

虽然几乎没有人真正遇见，但人工智能在发展的过程中确实存在一些问题。和当前能与公众产生共鸣的某些领域一样，上述问题或多或少要归咎于新闻媒体。过度热情的表现是，人们经常写文章赞扬他们所取得的令人印象深刻的进步，好像他们已经造出了智能机器。例如，20世纪60年代的机器人沙基（SHAKY）项目受到了大力宣传，人们将其视作世界上第一台通用机器人，它能够对自己的行为进行推理。这样一来，它就能够为模式识别、信息表示、问题解决和自然语言处理等不同领域设定基准。

尽管这已经足够使沙基艳惊四座，但记者们仍忍不住进行了一番修饰。1970年，当沙基在《生命》（*Life*）杂志上首次亮相的时候，它并没有被视为众多重要研究课题的结合体，而是以世界上第一个“电子人”的身份闪亮登场。《生命》杂志的记者将沙基卷入前一年登月引发的太空狂潮中，甚至称沙基能够“在不接受地球指令的情况下在月球上一次行走几个月”。

这是彻头彻尾的谎言，但是并非所有研究者都能够拒绝去迎合这一观点。20世纪70年代，在波士顿召开的一次会议上，一位研究者向媒体透露，仅需5年时间，像沙基这样的智能机器人就将在家庭中普及。一位年轻的同事激动地将他拉开，并对他说：“不要做这样的预测。以前有人这样做招来了麻烦。你低估了实现这一目标需要花费的时间。”这位研究者想都没想就回答道：“我不在乎。你看，我提到的所有日子都是在我退休后。”

人工智能工作者们并非一直这么激进，但有很多人都有这样的倾向。1965年，赫伯特·西蒙称，用不了20年，机器就能够“完成人类能做的任何工作”。不久以后，马文·明斯基补充道：“我们这一代人能够大体上解决创造人工智能的问题。”

“中文房间”实验

有关符号人工智能的哲学问题开始浮出水面。其中最著名的要数名为“中文房间”的思维实验。美国哲学家约翰·希尔勒（John Searle）提出该思维实验，质疑是否应将机器处理符号真正视为智能。

希尔勒提出，假设他被锁在一个房间里，房间里有很多中国书法作品。他并不懂中文，甚至无法将汉语与日语或其他毫无意义的字区分开来。希尔勒在房间中发现了一套规则，这些规则向他展示了一套与其他符号相对应的符号。随后，他被提问，并通过将问题符号和答案符号相匹配来回答这些问题。过了一会儿，希尔勒逐渐熟悉这项任务——尽管他仍然不清楚自己操作的这些符号到底是什么。希尔勒问，这种情况下能否说房间内的人“懂”汉语？他的答案是否定的，因为他完全缺乏意向性。他写道：“计算机可能有的这种意向性只存在于程序设计者、使用者、输入者以及对输出进行解读的人的思维中。”

如果说希尔勒是在指责人工智能研究者们像家长一样绞尽脑汁地炫耀孩子的才华，那么人工智能研究者们本身就面临着一个令人不愉快的事实：他们的孩子实际上并没有那么聪明。令人担忧的是，那些在实验室环境下表现不凡的工具并不能很好地适应现实状况。符号人工智能主要涉及自上而下建立以规则为基础的系统，该系统在实验室中表现出色，各元素都能够得到很好的控制。这些“微型世界”几乎不包含任何物质，因此可以采取的措施也十分有限。然而，一旦进入现实世界，在训练中表现优异的程序就像世界杯揭幕战中的英格兰队一样变得怯场了。

研究者们承认这些弱点的存在，并且将这些微型世界比作“一切事物都十分简单的仙境，如果以现实世界为前提，那么有关这些事物的陈述从字面上看就都变成了错的”。总的来说，人工智能一直在努力摆脱歧义性，但又缺少灵活抽象推理、数据计算和加工能力，而人工智能恰恰需要这些能力来理解其所展示的内容。任何没有事先明确说明的事物都有可能造成恐慌。美国作家约瑟夫·坎贝尔（Joseph Campbell）嘲讽道，这种人工智能与《圣经·旧约》一样，都是“规则太多，仁慈有限”。

莫拉维克悖论

结束这一不确定性面临着更大的问题，即人工智能研究者是否以正确的方式工作。就像玩拼图要从最困难的部分开始一样，人工智能研究者们设想，如果他们解决了复杂的问题，那么简单的问题就会迎刃而解。毕竟，如果你能让机器像数学天才一样下象棋，那么模仿婴儿学习又有多难呢？然而，事实证明这是相当难的。象棋是一项游戏，包含明确的说明、棋盘位置、合规或违规移动。象棋为棋手营造了一个静态世界，他们在这里拥有完整的信息，前提是他们能够看见棋盘，并且知道如何移动棋子。象棋是现实世界的一部分，但现实世界却与象棋截然不同。汉斯·莫拉维克（Hans Moravec）等研究者突然开始提出惊人的建议，例如“让计算机在智力测验中或在下跳棋时表现出成人水平相对容易一些，而让计算机在知觉和移动性方面达到一岁小孩的水平却是十分困难甚至是不可能的”。

将人工智能设定为关注生活中更复杂的事物，而排除对相对普通任务的关注，这可能与研究人工智能的人有关。在许多案例中，堪称“天才”的科学家们能够控制象棋或布尔逻辑（Boolean Logic）的微小细节，却缺少现实生活中的常识。有一则众所周知的趣闻：麻省理工学院一个名为西蒙·派珀特（Seymour Papert）的研究人员有一次将他的妻子忘在了纽约机场。当他意识到妻子没有陪在他身边时，飞机正在跨越大西洋。约翰·麦卡锡十分顽强地面对具有挑战性的问题，但是却因为经常忘记为资助他的各类机构填写进程报告而招致许多麻烦。据说麦卡锡在斯坦福大学讲授的课程“人工智能入门”并未受到重视，私下里被学生戏称为“听约翰叔叔讲故事”。都说什么样的人能干出什么样的事，这样看来，这些研究人员的人工智能项目都侧重于远大目标而非平凡（可能更实用）小事也就不足为奇了。

心理学家史蒂芬·平克（Steven Pinker）总结道：“人工智能研究的前35年得出的主要教训是，困难的问题容易解决，容易的问题很难解决。”

转变目标

在众多挑战的重重包围之下，传统人工智能开始出现问题。从20世纪70年代开始，随着之前几十年人们对人工智能乐观态度的逐渐消散，人们对该领域的热情也渐渐冷却下来。大幅缩减的预算使其首次遭遇了“人工智能的冬天”，而这样的情况不止一次发生。在美国，就连可爱的机器人沙基计划都被叫停，因为美国国防部意识到，其出钱资助的机器人项目并不能创造出他们需要的机器人间谍“詹姆斯·邦德”。暂且不说间谍这一点，沙基在战场上甚至无法发挥常规部队的作用！一名为该项目工作的研究人员回忆起沙基在斯坦福国际研究所（SRI）实验室中最后的日子。一位持怀疑态度的将军问一位创造沙基的研究人员：“能不能在上面安装一个36英尺长的刺刀？”

为应对这一状况，人工智能改变了其目标，按比例缩减了一些大型任务，集中力量处理那些通过采取措施就能够解决的问题。其中的一个例子就是电子游戏领域。人工智能从一开始就与电子游戏联系在一起，那时艾伦·图灵和克劳德·香农曾尝试打造一个自动象棋手。在当时的情况下，象棋就是一个微观世界，用来演示那些后来应用到现实世界的智能行为。电子游戏就是他们的最终目标。

人工智能研究不仅仅需要研究者的技能，还可以创造一定的收益。在莫斯科的苏联科学院计算机中心工作的28岁人工智能研究员阿列克谢·帕基特诺夫（Alexey Pajitnov）就是受益人之一。1984年6月，帕基特诺夫创建了一个简单的程序，用来测试实验室新的计算机系统。一个精明的商人将该系统命名为俄罗斯方块（Tetris）并拿到市场上销售，这款游戏在全世界范围内卖出了超过1.7亿个副本。

20世纪80年代，电子游戏变得更加复杂，人工智能专家也因此变得炙手可热。人工智能能够运用简单的规则模拟复杂的行为，这意味着电脑控制的人物也能有自己的想法。例如，在《主题公园》（*Theme Park*）游戏中，人工智能实体（指具有智能的任何实体，包括智能硬件和智能软件）聚集在使用者建造的公园周围，采取程序设计员从未明确标出过的路径行走。

在某种意义上，电子游戏是传统人工智能的天堂。如果人工智能仅用于在射击游戏中模仿僵尸敌人，那么其行为是真的智能还是仅仅表现得很智能这类问题就是毫无意义的。

专家系统

人工智能的另一项新应用成为人们解决问题的工具。尽管人工智能擅长推理，但研究人员都知道，人工智能并不只有推理。为了创造出能够在现实生活中解决问题的人工智能，科学家们认为，他们需要一台能够将推理和知识相结合的机器。例如，一台应用于神经系统科学的电脑必须像合格的神经系统科学家一样，了解该学科的相关概念、事实、表述、研究方法、模型、隐喻和其他方面。

这意味着程序员突然需要担负起“知识工程师”的重任，他们必须充当各领域专家的角色，并且将他们的知识提炼成计算机能够读取的规则，生成的程序被称为“专家系统”。该系统是在广泛收集概率性规则“如果……那么……”的基础上建立的。有人曾经使用名为“DENDRAL”的专家系统进行过尝试，这个程序能够帮助有机化学家确定未知的有机分子。DENDRAL的创造者爱德华·费根鲍姆（Edward Feigenbaum）对最早记载人工智能历史的一位作家帕梅拉·麦考达克（Pamela McCorduck）说道：“有一段时间，人工智能领域的其他人都与我们保持一定的距离。我想他们可能不敢靠近DENDRAL，因为它涉及化学。但是人们还是会毫不吝啬地发出惊奇的感叹，因为它就像是化学界的博士一样。”

另一个类似的项目名为“MYCIN”，用于为脑膜炎等严重感染状况提供合适的抗生素用药剂量。MYCIN像一个真正的医生一样，能够将程序员之前收集的概率性证据汇总起来，并据此得出结论。人们不断对之前的经验进行归纳，直到它们变得像下面的“规则”：

如果……需要治疗的感染是脑膜炎，感染类型为真菌感染，培养染剂上没有有机体，患者不是易感染病患并且曾经到过孢子菌病盛行的地区，患者是黑种人、黄种人或印度人，并且脑脊髓液检测中的隐球菌抗原不是阳性，那么……隐球菌有50%的可能并非是造成感染的有机物之一。

[4]。简而言之，XCON就是世界上最伟大的“百事通”店员。

XCON的出现使大企业首次对人工智能产生了兴趣，它们不再将人工智能简单视为未来演示。只要专家系统能够为它们赚钱，企业们并不在乎它们到底是真正的人工智能还是“巧妙的编程”。1980年，XCON在数字设备公司（DEC）位于新罕布什尔州塞勒姆的工厂首次投入使用。到1986年，XCON就已经处理了80 000份订单，每年约为数字设备公司节约2 500万美元，准确率高达95%—98%。人们开玩笑称，如果它能够娶了老板的女儿，未来甚至能够当上公司的首席执行官。

其他企业很快也开始公开露面，为需要专家系统的企业客户提供解决方案。在油井钻探作业中，地层倾角顾问（Dipmeter Advisor）能够对地质构造进行分析。著名的粮食市场顾问（Grain Marketing Advisor）的目标是帮助农民进行恰当的营销，并储存粮食作物。1986年10月《电脑世界》（*Computer World*）杂志上的一则广告做出了这样的问答：“你如何充分利用专家系统技术的优势让员工改进当前软件上现有的数据处理应用？只有Teknowledge公司能够给你答案。公司将在你所在的地区免费举行为期半天的研讨会。”

1985年，约有150家公司投资10亿美元想要开展人工智能业务。这一年，美国人工智能协会召开的一场

会议和国际人工智能联合大会共吸引了近6 000名参会者，其中一多半是风险投资家、猎头和媒体。1987年，并非电脑研究前沿阵地的《财富》杂志盛赞“软盘上的直播专家”的到来。研究人员在人工智能历史上首次变得和史蒂夫·乔布斯、比尔·盖茨等个人电脑界的新贵一样富有。

有趣的是，像马文·明斯基这样经验丰富的研究者却在回避这样的情况。我们往往以为，人工智能保守派在经历了20多年辛苦工作后一定急于寻求回报。而实际上，这些人都在提心吊胆地等待最后的结果。好在并没有花费他们太多的时间。就像20世纪90年代后期投机性的互联网泡沫一样，倡导者们对专家系统能力的高估到了危险的程度。其中有一本教科书十分推崇“打电话规则”，称“使用专家系统打一通电话，家里的所有问题都能够在10—30分钟解决”。专家系统的基本概念是十分可靠的，但是仍存在一些问题。专家系统十分昂贵，需要不断更新，并且是反直觉的，当规则增多时准确率就会下降。斯图尔特·罗素（Stuart Russell）和彼得·诺威格（Peter Norvig）在《人工智能：一种现代方法》（*Artificial Intelligence: A Modern Approach*）这本教科书中写道：“随着规则集的不断壮大，规则间的不良交互作用就会越来越平常，于是，从业人员发现，添加规则时，必须‘调整’其他规则的可信度。”

1987年财年结束时，两大开发专家系统的公司Teknowledge和Intellicorp损失了数百万美元。其他人工智能企业的情况更糟糕，几乎濒临破产，员工和公司高管们露宿街头。温暖了一阵之后，人工智能的第二个冬天到来了。

谷歌诞生

人工智能随后遭遇的寒流比第一次有过之而无不及。资本又一次迅速蒸发，政府补助消失得无影无踪。在1987—1989年，美国国防部高级研究计划局将用于人工智能研究的预算缩减了1/3。专业的人工智能杂志的广告费收入也大幅下降。美国人文与科学院官方期刊《代达罗斯》（*Daedalus*）在1988年大胆发布了一期人工智能特刊，激怒了哲学家希拉里·普特南（Hilary Putnam）。普特南写道：“现在有什么好大惊小怪的？为什么在《代达罗斯》上发布一期特刊？为什么不等到人工智能真正取得一些进展后再发布特刊？”整个科技界都能感受到人工智能的寒流。美国人工智能协会会员人数大幅减少，1996年骤减至4 000人，达到史上最低点。然而奇迹并没有发生，人工智能的美梦似乎就要破灭。

那一年，斯坦福大学的两个学生——一个是人工智能研究者的后代，另一个是数学家的后代——想到了一个很聪明的方法，在外部链接数量的基础上对网页进行排序，并通过这一方式创建一个智能网络目录。1997年，24岁的拉里·佩奇和谢尔盖·布林利用他们开发出的算法，在美国加利福尼亚州门罗帕克的一个车库里开了一家公司。为了将其打造成为“全球总部”，他们置办了几张桌子、三把椅子、一条蓝绿色的地毯、一张折叠乒乓球桌，以及一些其他物件。为了保持通风，车库的门必须一直敞开。

这在当时看起来似乎不足为奇，但在接下来的20年里，拉里·佩奇和谢尔盖·布林的公司取得了人工智能历史上最大的成就。公司涉猎范围广泛，覆盖了机器翻译、模式识别、计算机视觉、自主机器人等领域，而人工智能研究者们为了实现这一点已经苦苦挣扎了半个世纪。

实际上，公司的这些成就没有一点是通过传统人工智能实现的。

这家公司就是谷歌。

[1]1英亩=4 046.86平方米。——编者注

[2]1英尺≈0.304 8米。——编者注

[3]1英里=1.609 344千米。——编者注

[4]一种可以支持机器语言和虚拟地址的32位小型计算机。——编者注

第二章 以自主学习的方式创建新的人工智能

第二章

以自主学习的方式创建新的人工智能

2014年，在谷歌旗下一家名为“DeepMind”的人工智能公司的办公室里，一台计算机通过玩一款名为《打砖块》（*Breakout*）的老雅达利（Atari）2600电子游戏消磨时间。该款游戏是两个年轻人在20世纪70年代初设计的，他们就是苹果公司的创始人史蒂夫·乔布斯和史蒂夫·沃兹尼亚克。《打砖块》实际上是乒乓球游戏《乒乓》（*Pong*）的一个变体。不同之处在于，不是在屏幕上将球挥向另一位玩家，而是对着砖墙击球，将砖块击碎。这款游戏的目标是摧毁所有砖块。

正如我们在上一章中提到的，人工智能玩电子游戏并没有什么稀奇的。艾伦·图灵早在1947年就开发出了首款象棋程序，尽管当时的计算机不能运行这一程序。如今电子游戏的特点是有大量非玩家控制角色，这一编程将简单的规则结合起来产生复杂的行为。这样看来，DeepMind的人工智能玩游戏又有什么特别的呢？

针对这个问题的回答有两个。一是DeepMind的人工智能会逐渐变得更加成熟。就像见证孩子逐渐长大一样，如果一直盯着计算机看，很难察觉到它的变化。然而，每隔50多次游戏再看一下，效果是十分惊人的。开始的时候，DeepMind的人工智能在《打砖块》游戏中的表现简直糟透了，最简单的击球都做不好，而且它似乎并不清楚状况，就好像是把PS4（索尼第四代游戏主机）手柄交到90岁的老奶奶手里，并希望她立刻知道应该做什么一样。虽然它也会偶尔得分，但即使最乐观的旁观者也只能称之为运气。

200次游戏后，一切变得大为不同。现在游戏中的球拍能够在屏幕上左右移动：即使不是持续得分，也可谓能够轻松得分。再经过数百次游戏，游戏中的人工智简直如同《星球大战4：新希望》结束时的天行者卢克（Luke Skywalker）或《黑客帝国》中的尼奥（Neo）一样，懒散地击球，毫不费力。所有无关的动作都消失了，而且它产生了清晰的策略。

令DeepMind的人工智能具有重要意义的另一个原因是，它不需要进行大规模训练。传统人工智能的核心原则是必须将规则预先载入系统，这就像是老师在学生参加考试前会依次教他们问题的答案一样。DeepMind的人工智能与众不同之处在于，它能够自主学习，甚至无须告诉它应该怎样做。它所需要接入的就是构成《打砖块》游戏每一帧的30 000个像素点和屏幕上的选手得分。其他需要做的事，就是给它输入得分最大化的指令。之后，人工智能就可以随着游戏的进展获得游戏“规则”，然后逐渐形成能够改善其表现的策略。

DeepMind的人工智能可以玩的游戏并不只有《打砖块》。它最早玩的游戏是《太空入侵者》（*Space Invaders*），在掌握极少信息的情况下还学会了其他48个游戏，包括拳击模拟器、武术游戏甚至是3D（三

维)赛车游戏。然而,要想突破电子游戏的“微型世界”还有很长的路要走。但这仍是一项惊人的成就,为人工智能的下一步发展指明了方向。下一步发展是什么?按照DeepMind的宗旨,下一步就是“解决人工智能”。

自主学习的重要性

人类的与众不同之处就在于能够学习，这也一直是传统人工智能一直努力要实现的。第一章中描述的系统只有在能够遵从规则时进行学习，这些知识是从“知识工程师”的知识中提炼并编入系统架构的。它是对知识自上而下的一种想象，并暗示一个假设，即机器不能自动学习知识。相反，必须将知识进行编程，而且一次编一条。这一点在很多情况下都能够很好地实现，进而在可接受的水平上完成有限的任务。随着解决方案的增多，问题也开始显现。像官僚机构一样，它们开始变得庞大、笨拙、缓慢而且昂贵。

这提出了一个显而易见的问题。华盛顿大学计算机科学教授普德罗·多明戈斯（Pedro Domingos）指出：“如果机器人掌握了人类除学习以外的所有能力，人类很快就会抛弃它。”但是从一开始就存在一种与人工智能的发展并行的观点，这一观点现在正触发该领域的诸多进展。该人工智能学派不是将思维概念化，而是源于在电脑内部建立大脑模型。该学派不相信逻辑推理是获取真理的最佳（可能是唯一的）途径，而是采用基于观察和实验的实证研究法。这类人工智能并非知识工程师的作品，而是属于名为“机器学习者”的计算机科学家领域。

这一流派的人工智能由统计学家、神经科学家和理论物理学家开创的概率模型主导，大部分基于所谓的“神经网络”（或者计算术语中所说的“神经网”）来运行，该网络的功能与人脑近似。信息在人脑中以神经元电子放电模式存在。人脑中约有1 000亿个神经元，大约和银河系中的星星一样多。记忆是通过加强不同神经元共同放电而形成的：这一过程被称作“长时程增强”。尽管我们尚须建立一个与人脑一样复杂的神经网络（下一章将详细介绍），但人工神经网络为创造记忆和学习借用了人脑的机制。人脑与神经网络最基本的不同在于，人脑中的长时程增强是一个生物化学过程，而在神经网络中，学习是通过修改其自身代码，以在复杂或不明朗的情况下，找到输入和输出之间或者原因和结果之间的联系而发生的。

虽然神经网络今天在人工智能领域具有重要地位，但在过去许多年里，它都是被忽视的；它被视作真正人工智能的“异父兄弟”。正如20世纪80年代进入这一领域的知名研究人员戴维·艾克利（David Ackley）所说：“我们接触到神经网络时，人们并未将其视作人工智能。于是，我们被人工智能拒之门外。当时，人们认为人工智能是与符号相关的。它所涉及的是生产系统、专家系统等。进入卡内基-梅隆大学读研究生时，我已经十分厌倦与传统的符号化的计算机相关的事物……我似乎对推理的关注过多，而对判断的关注太少。”

艾克利影响了一代人工智能研究者，他们几乎使统计工具替代了主流意识中的传统人工智能。这样一来，神经网络就实现了以前的研究人员做梦都想实现的东西：建造能够学习如何玩电子游戏、理解语言、识别相片中的人脸或开车比人类更安全的机器。

我们在本章中将介绍一些这样的应用。但是在此之前，我们必须回到过去，去认识一个名叫圣地亚哥·拉蒙·卡哈尔（Santiago Ramón y Cajal）的人。

神经科学之父

圣地亚哥·拉蒙·卡哈尔是19世纪西班牙病理学家，被称作现代神经科学之父。拉蒙·卡哈尔首次对人类大脑进行了细致的检查。1887年拉蒙·卡哈尔在巴塞罗那大学工作，他发现重铬酸钾和硝酸银可以将神经元染成深色，而周围的细胞还能够保持透明。后来他回忆道，“染色后的神经细胞连最精细的分枝都变成了棕黑色，在透明的黄色背景映衬下显示出了无可比拟的清晰度，就像用墨汁画的素描一样清晰”。这项神经细胞染色技术意味着拉蒙·卡哈尔能够就人脑展开大量的研究，在过去使用最先进的显微镜是无论如何都做不到这一点的。这样一来，他首次证明了神经元是构建中枢神经系统的基础。

1943年，拉蒙·卡哈尔去世9年后，两位人工智能研究人员在一篇很有影响力的论文中创建了首个正式的神经元模型，尽管文章的标题《神经活动内在概念的逻辑演算》稍显呆板。

两位研究人员麦卡洛克（McCulloch）和皮茨（Pitts）组成了一个不同寻常的组合。沃伦·麦卡洛克1898年出生在一个律师、工程师、医生和神学家组成的家庭。他在当时被称为“美国帽子之都”的新泽西州奥兰治长大。麦卡洛克起初打算从政，但后来改变了主意，去耶鲁大学学习了哲学和心理学，并对神经生理学，也就是神经系统的研究产生了浓厚的兴趣。

皮茨比麦卡洛克小25岁，1923年出生在一个工人阶级家庭，这样的家庭似乎不太可能培养出神童。13岁时，皮茨为了躲避父亲的虐待离家出走，露宿街头。一天，他为了躲避一群地痞流氓的追赶躲进了图书馆。据说，皮茨在接下来的一周都泡在图书馆里，读完了三卷《数学原理》的数学教材。读完以后，皮茨决定给该书的作者之一伯特兰·罗素（Bertrand Russell）写信，指出他认为第一卷中存在的根本错误。这封信给罗素留下了深刻的印象，他甚至邀请皮茨到英国剑桥大学学习，而皮茨却没有接受邀请。皮茨在不到20岁时就被苏联数学物理学家尼古拉斯·拉舍夫斯基（Nicolas Rashevsky）的著作深深吸引，拉舍夫斯基的著作主要涉及数学生物物理学领域。正是凭借着这种能力，沃尔特·皮茨遇到了沃伦·麦卡洛克，并最终开始与其共事。

麦卡洛克和皮茨共同提出了针对机器内部复制的功能神经元的简化模型。他们在1943年发表的论文中称，从根本上来讲，神经元是一个“逻辑单元”。他们还指出，由这类单元构成的网络几乎能够完成所有的计算操作。

以神经元模型为基础的感知器

麦卡洛克和皮茨的工作取得了至关重要的进展，但同样存在严重的局限性：这个模型不能自主学习。6年后，这一问题在理论上得到了解决，加拿大心理学家唐纳德·赫布（Donald Hebb）在1949年写了《行为的组织》这本书。赫布称，每次使用神经元都会使人脑中的神经通路加强，人们就是这样学习的。他写道：“细胞A的一个原子离细胞B足够近，并且持续或不断参与激发细胞B，其中一个或两个细胞增长或产生代谢更换，这就会导致细胞A激发细胞B的效率提高。”简单来说，赫布的意思是，当人类大脑中有两个神经元同时受到激发时，二者之间的联系就增强了。有时我们可以这样来记忆：“同激发、同连接的神经元。”

赫布的这一思想在10年后才真正应用到计算机研究中，而这要归功于弗兰克·罗森布拉特（Frank Rosenblatt）。罗森布拉特在计算机历史上是一个有趣的人物：他是一个真正博学多才的文艺复兴式人物，对音乐、天文、数学和计算机无不精通。碰巧的是，他和我们上一章提到的马文·明斯基是同学，他们在20世纪40年代早期都在布朗克斯科学高中读书。然而，罗森布拉特一直处于人工智能研究主流的边缘。明斯基和约翰·麦卡锡组织达特茅斯会议期间，罗森布拉特拿到了康奈尔大学实验心理学博士学位，学习期间，他被神经网络这一学科深深吸引。罗森布拉特将神经网络称作“感知器”，并努力证明其能够有效地充当人类学习、记忆和认知的模型。

罗森布拉特最初在纽约布法罗康奈尔航空实验室尝试建造“感知器”。他在那里创建了PARA项目，即“感知和识别自动化”。他的感知器以麦卡洛克和皮茨提出的神经元模型为基础，同时基于能够通过“试错”进行学习的神经网络。每个神经元都有一个输入、一个输出和一组自己的“权重”。开始的时候，“特性”之间的关联和神经元都会获得随机权重。然后，神经元根据网络显示，选择激发或不激发。片刻后，它就能够将见到的所有事物分为两类，即“X”和“非X”。

由于当时的计算机运行速度太慢，罗森布拉特并没有将其感知器做成软件，而是做成了硬件。他用调光器中使用的可变电阻创建了权重，并用电动机和电阻完成了学习过程。接下来的演示以及罗森布拉特对感知器发展潜力的夸张陈述，足以让人们心潮澎湃。1958年《科学》杂志上发表的一篇极有先见之明的文章上指出：“感知器最终一定能够自主学习、做出决定以及翻译语言。”与此同时，《纽约客》上一篇新发表的文章引用了罗森布拉特的话，“感知器应当证明它能够通过计算机视觉指出‘猫和狗之间的不同点’”。

1960年，罗森布拉特对“阿尔法感知器”计算机的创建工作进行了监督，他为此收到了美国海军研究办公室信息系统部提供的赞助。阿尔法感知器也成为历史上最早能够通过反复试错学习新技能的计算机之一。《纽约时报》将其称为“边做边学的新海军设备”。

关于感知器的争论

令人遗憾的是，不久以后，与感知器有关的研究遭遇了两次严重的挫折。第一次主要是技术方面的原因，也有些许人为因素。感知器当时已经被证实能够完成简单的学习任务，例如识别语音或印刷字体。然而，它们也成功地引起了广泛关注，给它们的投资也远远超过了技术取得成功的水平。这使得人工智能界内部产生了矛盾。马文·明斯基就是众多直率的评论家之一。明斯基在读博士时就已经研究过神经网络，但是对该领域已经不抱任何希望。自20世纪50年代以来，罗森布拉特与明斯基曾在众多科学会议上就脑启发计算机（brain-inspired computer）的有效性展开争论。罗森布拉特强烈推荐他研发的技术，称感知器实际上能够完成任何学习任务。明斯基则持截然相反的态度。直到1969年，明斯基与研究人员西蒙·派珀特合著了一本著作才打破这一僵局，该书就感知器可能无法完成的事情进行抨击。明斯基和派珀特总结道，这一技术“不具有任何科学价值”。为神经网络筹集的资金也顷刻化为乌有。

第二次挫折更加悲惨。明斯基和派珀特出版《感知器》一书两年后，在一个周日，弗兰克·罗森布拉特去美国最大的河口切萨皮克湾划船。那天是他43岁的生日，然而当天发生了意外，罗森布拉特就这样去世了。在康奈尔大学的悼词中他的一位同事写道：“他的离去让我们失去了一个最无私、最富有同情心的同事，他的机智和幽默给我们留下了深刻的印象。”这样一来，感知器也就失去了最忠实的拥护者。

脑启发式神经网络在接下来的10年中似乎销声匿迹了。后来，西蒙·派珀特以童话的方式，将传统人工智能和感知器之间的冲突总结如下：

从前，控制论科学生了两个女儿，一个是自然的，能够从对人脑的研究和自然规律中继承特性。另一个女儿是人造的，涉及范围从计算机最初的创造到使用。这两姐妹都想要建造智能模型，但是用的材料却大不相同。自然科学使用数学纯化神经元构造模型（称为神经网络），而人造科学则是通过计算机程序构造模型。

派珀特借用了白雪公主的故事，将他和明斯基对感知器的抨击比作猎人冲入树林对白雪公主的追杀。就像童话故事中讲的那样，派珀特和明斯基带着感知器的“心脏”回到他们的主人（在本案例中，我们称其为万能的“美国国防部高级研究计划局”）身边，证实感知器已经死了。“然而，派珀特和明斯基展示给世界的证据并不是白雪公主的心脏，而是猪的心脏。”

这可能曾一度引起过度紧张，但派珀特承认，神经网络躲过了他和明斯基的猛攻。事实也确实如此。实际上，派珀特在20世纪80年代末写下这句话的时候，神经网络已经又一次发展得如火如荼。

霍普菲尔德网的兴起

与明斯基和派珀特的断言相反的是，神经网络研究人员多年来一直认为，神经网络能够展现出新的能力，并且解决罗森布拉特感知器的问题，但前提是在网络输入和输出之间放置额外的“隐含”神经元层。不幸的是，没人知道如何训练这些多层神经网络。著名物理学家约翰·霍普菲尔德（John Hopfield）为具体应当如何做提供了建议。

霍普菲尔德对当时人工智能的主流形式是什么并不感兴趣。他说：“我从未深入研究过人工智能领域到底发生了什么。人工智能并不能解决现实世界里的的问题。我认为没必要去了解它。”然而，多年以后，他一直苦苦追寻被他称作“需要用一生的时间去研究的问题”。由于对人类大脑十分感兴趣，他考虑的问题涵盖范围广泛，从灵长类神经解剖学到昆虫飞行的行为，再到大鼠海马的学习乃至阿兹海默症的治疗。有一段时间，霍普菲尔德对细胞自动机和自我复制的机器人的前景十分着迷。然而，几个月的研究最终还是走进了一条死胡同。

霍普菲尔德说：“放弃一个错误的思想十分困难，毕竟我们已经研究了一年了。”但是，在计算机内部创造一个生命模型的想法却一直都在。他对一个想法十分着迷，即用神经网络完成大脑能够迅速且轻松完成但计算机却不能完成的任务。霍普菲尔德最终选择了联想记忆，联想记忆是指大脑如何以交互的方式工作，也就是看见一个人就能联想起他的名字，或者听到他的名字就能想起他的长相。联想记忆背后的数学运算使霍普菲尔德想到“自旋系统”的数学运算，该运算描述的是固体磁性的复杂形式。他的脑海中突然产生了一个想法。霍普菲尔德回忆道：“神经生物学和我所了解的物理系统之间突然产生了一种联系。一个月之后，我已经开始写论文了。”

1982年这篇论文发表后，一种全新的神经网络产生了。霍普菲尔德网络比罗森布拉特的感知器中的单层模拟神经元复杂得多。他的思想再次激发了人们对神经网络的热情，这也使他成为这一过程中出人意料的英雄。加州理工学院的一组追随者开始以“霍普集会”（Hop-Fest）的名义召开会议。霍普菲尔德的发现吸引了一些世界上最伟大的理论物理学家参与到神经网络的研究当中。该领域的研究人员多年来第一次感觉到热血沸腾。

然而，事情并没有我们想象得那么容易。正如我们在第一章中看到的，20世纪80年代早期是“专家系统”的天下，资金也是空前的充裕。尽管后来这些“专家系统”的发展将遭遇困境，但在当时却是十分强大，人们根本不认为它们会失败。世界领先的神经网络专家特里·谢伊诺斯基（Terry Sejnowski）当时正在普林斯顿大学读霍普菲尔德的博士，他回忆道：“我们当时好像是生活在恐龙时代的只有毛皮的哺乳动物，在这些长着鳞甲的巨兽的脚下混日子，他们有数百万美元的机器和庞大的预算。那时，所有人都专注于计算逻辑，但是我们明白，他们忽视了推动人工智能向前发展所面临的真正困难。”

幸运的是，神经网络吸引了许多年轻且富有热情的研究人员，其中就包括圣地亚哥加州大学的认知科学家戴维·鲁梅尔哈特（David Rumelhart）和詹姆斯·麦克兰德（James McClelland），他们成立了一个“并行分

布处理”小组，并产生了令人难以置信的影响力。

说到这里就不得不提到另外一个人，他就是杰夫·辛顿（Geoff Hinton）。

神经网络的守护神

杰夫·辛顿出生于1947年，是现代神经网络最重要的人物之一。作为一名谦逊的英国计算机科学家，他对其所在领域的发展产生的影响很少有人能企及。他出生于一个数学家家庭：他的曾祖父是著名的逻辑学家乔治·布尔（George Boole），他的布尔代数曾为现代计算机科学奠定了基础。另一位亲戚是数学家查尔斯·霍华德·辛顿（Charles Howard Hinton），因提出“四维空间”这一理念而闻名，阿莱斯特·克劳利（Aleister Crowley）在其小说《月之子》中曾经两次提到了辛顿。

辛顿说：“我一直对人类如何思考以及大脑如何工作很感兴趣。”上学时，一个同学说大脑储存记忆的方式和3D全息图像储存光源信息的方式是一样的。要想创建一个全息图，人们会将多个光束从一件物品上反射回来，然后将相关信息记录在一个庞大的数据库中。大脑也是这样工作的，只是将光束换成了神经元。由于这一发现，辛顿在剑桥大学选择了研究哲学和心理学，之后又在苏格兰爱丁堡大学研究人工智能。辛顿在20世纪70年代中期来到寒冷的爱丁堡，人工智能领域遭遇的首个冬天几乎在同一时期到来。尽管传统人工智能刚刚遭受打击，但辛顿的博士导师仍急于让他远离神经网络。辛顿说：“他一直试着让我放弃神经网络的研究并投入到符号人工智能领域。为了能够有更多时间研究神经网络，我必须不断和他讨价还价。”

辛顿并没有获得其他的支持。学生们认为他是疯了才会在明斯基和派珀特完全否认神经网络后还继续研究。辛顿在爱丁堡期间，明斯基的学生帕特里克·温斯顿（Patrick Winston）出版了一本早期人工智能教材。书中记载着有关神经网络的内容：

许多古希腊人都支持苏格拉底的一个观点，即深奥且令人费解的思想是上帝创造的。如今，对这些漂泊无定的人而言，甚至概率神经元都相当于上帝。很有可能的是，神经元行为的随机性的提高是癫痫病患者和醉酒的人的问题，而不是聪明人的优势。

人们对温斯顿的思想十分不屑，但是他当时对神经网络的类似于宗教信仰般的看法并不是完全错误的。辛顿对人脑必须以某种方式工作这一认识十分欣慰，很明显，这是无法用传统的符号人工智能来解释的。他说：“大多数常识推理都是凭直觉或以类比的方式做出的，其中并不涉及意识推理。”辛顿认为，传统人工智能的错误之处在于：其认为，任何事都是由一系列基本规则和有意识推理组成的。对符号人工智能研究人员来说，如果我们不能理解某一部分的意识，这是因为我们还没有弄懂其背后的推理。

毕业以后，辛顿暂时在英国苏塞克斯从事博士后工作，之后收到了一份来自美国的工作邀请。于是，辛顿打点行装，搬到了加州大学，不久以后，又搬到了卡内基-梅隆大学。在接下来的几年里，他一直积极努力在神经网络领域取得开创性进展，即便到了今天，其成就仍对人工智能的研究产生着影响。

他最重要的贡献之一，要算是他对另一位研究人员戴维·鲁梅尔哈特的帮助，帮助他再次发现“反向传播”流程，这大概是神经网络中最重要的算法，之后他们首次以可信的方式证明，“反向传播”使神经网络能够创建属于自己的内部表征。当输出与创造者希望的情况不符时，“反向传播”使神经网络能够调节其隐藏

层。发生这种情况时，神经网络将创建一个“错误信号”，该信号将通过神经网络传送回输入节点。随着错误一层层传递，网络的权重也随之改变，这样就能够将错误最小化。试想一下，有一个神经网络能够识别图像，如果在分析一张狗的图片时，神经网络错误地判断为这是一张猫的图片，那么“反向传播”将使其退回到前面的层，每层都会对输入连接的权重做出轻微调整，这样一来，下次就能够获得正确的答案。

20世纪80年代创建的“NETtalk”项目是“反向传播”的一个经典案例。NETtalk的一个共同创建者特里·谢伊诺斯基将其描述为用于了解电脑是否能够学习大声朗读书面文字的“夏季项目”。该项目面临的最大挑战在于语言一点也不简单。项目刚刚开始的时候，谢伊诺斯基去图书馆借了一本有关音韵学的书，即诺姆·乔姆斯基（Noam Chomsky）和莫里斯·哈雷（Morris Halle）所著的《英语语音模式》。谢伊诺斯基说：“这本书里都是各种事情的规则，例如字母e出现在单词末尾的时候应该如何发音等。书中提到了例外情况，之后又列举了例外情况中的例外。英语就是大量的复杂关联。我们似乎选择了世界上在规则性方面最糟糕的语言。”

一直以来，传统人工智能都在不断尝试将这些单独的例子插入到一个专家系统中。谢伊诺斯基和一位名为查尔斯·罗森伯格（Charles Rosenberg）的语言研究人员决定通过创建一个由300个神经元组成的神经网络来实现这一目标。当时，辛顿正在实验室访问，他建议他们在项目的最开始使用儿童书籍来训练该系统，这本书的词汇量一定要小。起初，这项任务十分艰难，计算机一次只能读一个单词，而他们必须为每个字母都标注正确的音素。例如，字母e在“shed”、“pretty”、“anthem”、“café”或“sergeant”中的发音各不相同。谢伊诺斯基和罗森伯格每次进行说明的时候，他们创建的神经网络都悄悄地调节对每个连接的权重。该系统面临的最大挑战是使机器能够正确发出每个单词中间部分的音节。为了做到这一点，神经网络必须使用中间字母左边和右边的字母给出的提示。

一天下来，NETtalk已经全部掌握了书中的100个单词。这一结果令他们感到震惊。接下来，他们让NETtalk使用有20 000个单词的韦伯词典。幸运的是，词典中的所有音素都已经标注出来了。他们下午把单词输入到系统中，然后就回家休息了。当他们第二天早上回到办公室时，系统已经完全掌握了这些单词。

最后的训练数据是一本对儿童说话内容进行誊写的书，以及一位语言学家记录的儿童发出的实际音素的清单。这就意味着，谢伊诺斯基和罗森伯格能够将第一个誊写本用于输入层，将第二个音素清单用于输出层。使用“反向传播”以后，NETtalk能够学习如何像孩子那样说话。一段NETtalk的录音说明了该系统在这方面取得了飞速的进展。在训练之初，系统只能够区分元音和辅音，其发出的噪声则像是歌手表演前做的发声练习。在训练了1 000个单词以后，NETtalk发出的声音更接近人类发出的声音了。谢伊诺斯基说道：“我们完全震惊了，尤其是在当时计算机的计算能力还不如现在的手表的情况下。”

联结主义者

有了杰夫·辛顿等人的帮助，神经网络开始蓬勃发展。当时有一个传统，那就是继任的一代都会给自己重新命名，新研究人员称自己为“联结主义者”，因为他们对复制大脑中的神经联结十分感兴趣。到1991年，仅在美国就有1万名活跃的联结理论研究人员。

忽然之间，各个领域都取得了突破性的进展。例如，人们发明了专门用于预测股市的神经网络。大多数情况下，投资公司使用不同的网络预测不同的股票，然后由交易商来决定投资哪只股票。然而，有些人在此基础上更进一步，赋予网络本身自主权，使其能够自行买卖。无独有偶，金融领域迅速涉足电子游戏领域，时刻准备着为人工智能研究人员进行投资。算法交易时代轰轰烈烈地开始了。

当时神经网络领域的另一个引人注目的应用就是自动驾驶汽车。发明自动驾驶汽车一直是技术人员的梦想。1925年，发明家弗朗西斯·霍迪纳（Francis Houdina）展示了一款无线电控制的汽车，他操控汽车行驶在曼哈顿的街头，而车中无须人来操控方向盘。之后，自动驾驶汽车测试使用导丝和车载传感器使汽车能够按照路上画好的白线行驶，或通过识别出地下电缆发出的交流电行驶。1969年，约翰·麦卡锡发表了一篇标题为“计算机控制汽车”的论文极具挑战性。麦卡锡所提议的方案基本上是设计一个“自动化司机”。他的项目需要一个能够进行公路导航的计算机，计算机上仅带有一个电视摄像机来输入信息，该输入使用与人类司机相同的视觉输入。麦卡锡假设用户能够使用键盘输入地点，并要求汽车立即载他们过去。紧急情况下，用户可以使用额外的命令变更目的地，要求汽车停在洗手间或宾馆门口，在有紧急情况时减速或加速。

类似的项目直到20世纪90年代早期才得以实现，当时卡内基—梅隆大学的研究人员迪安·波默洛（Dean Pomerleau）写了一篇激动人心的博士论文，文章介绍了如何将“反向传播”应用于无人驾驶汽车。波默洛称其开发的神经网络为神经网络中的无人驾驶汽车或ALVINN（控制器），并将道路上的原始图像作为输入信息，并实时输出转向控制信息。当时，还有许多其他传统人工智能博士正在研究类似的自动驾驶项目。这些非神经网络的方法主要通过严谨的像素分析将各图像划分为不同类别，例如“道路”和“非道路”。然而，与许多传统人工智能面临的问题一样，计算机很难将信息解析为像实时路况那样的非结构化信息。假如一辆自动驾驶汽车依靠这一技术进行危险的高速行驶，发生事故的可能性是很大的。波默洛回忆道：“它们可能将树影或者树木本身识别成道路，这样车辆就会朝着树直接开过去，而不是避让。”

为了训练ALVINN，驾驶员只需简单地驾驶一段路程。波默洛说道：“驾驶员只需驾驶2—3分钟，ALVINN系统就能够了解并更新反向传播网络的权重。结束驾驶时，驾驶员可以放开方向盘，系统会继续驾驶车辆开始一段新的路程。”波默洛的发明只关注了方向，却无法控制速度或避开障碍物，这两点必须由驾驶员来完成。尽管如此，波默洛也取得了巨大的成功，1995年，庞蒂克小型货车上安装了从旧汽车上回收的ALVINN的升级版——RALPH（快速调节横向位置处理器）。波默洛和一位名为托德·约赫姆（Todd Jochem）的研究人员为其配备了一台电脑、640×480像素的彩色照相机、全球定位系统接收器和光纤陀螺仪，之后他们驾驶该车横穿美国。借鉴了1986年“携手美国”（Hands Across America）慈善活动的名称，他们

将这次旅行称为“横穿美国”（NO Hands Across America）。他们在路上卖10美元一件的衬衫，用于支付食宿费用。最后，这辆汽车一共行驶了2 797英里，途经匹兹堡、宾夕法尼亚、圣地亚哥、加利福尼亚，中间还穿过了胡佛水坝，这一切都是汽车自动驾驶完成的。《商业周刊》的一位记者在报道这一事件时，一名堪萨斯州骑兵要求其将车停到路边。而波默洛和约赫姆乘自动驾驶汽车旅行，甚至连双手都无须握住方向盘。

15年后，谷歌在2010年10月发布了自己的无人驾驶汽车项目。然而，我们仍要感谢波默洛在神经网络领域做出的开创性贡献，他证明了自己的观点。

欢迎来到深度学习领域

[1]。记者史蒂夫·洛尔（Steve Lohr）在其所著的极为有趣的《数据论》一书中指出，如果能将这些数据输入iPad Air（苹果超薄平板电脑）中，那么产生的堆栈将能够覆盖地球到月球距离的2/3。

然而，就像地球虽然有大量的水，但并不是所有水都可以喝一样，这些数据中好多都是未标记的。当数据集较小时，研究人员可以将主要精力放在正确标记所有数据上，这对训练系统来说更加有用。然而，随着数据量的增加，研究人员就无法再这样做。例如，2013年3月，网络相册Flickr共有8 700万注册用户，他们每天上传超过350万张新图片。从理论上讲，这对那些想要建造一个能够识别图片的神经网络的人们来说是一个天大的好消息，但同样也提出了挑战。正如我们所看到的，训练神经网络最简单的方法就是向其展示大量图片，然后指出每张图片都是什么。通过标记图片，训练员既提供了输入（图片），又提供了输出（描述）。神经网络就可以反向传播，以纠正错误。这就是我们所了解的“监督式学习”。但是，流通中还有许多未标记或没有正确标记的图片，计算机如何对其进行识别呢？

幸运的是，杰夫·辛顿掀起了一场“非监督式学习”的革命，这种学习方式无须向计算机提供任何标记。机器能够访问的只有输入，无须解释它看到的是什么。首先，这听起来像是机器无法通过这种方式学习。如果没有得到明确的解释，即使是最智能的神经网络也不会知道某物到底是什么。实际上，辛顿发现的是“非监督式学习”可以用来训练上层特征，而且每次只能训练一层。这一发现成为“深度学习”的催化剂，而“深度学习”就是当前人工智能最炙手可热的领域。

我们可以将深度学习网络想象成工厂的一条生产线。输入原材料后，它们将随着传送带向下传递，后续的各个站点或层会分别提取不同的高级特性。为了继续完成一个图像识别网络的案例，第一层将用来分析像素亮度。下一层将根据相似像素的轮廓来确定图中存在的所有边界。之后，第三层将用来识别质地和形状等。到达第四层或第五层时，深度学习网络已经创建了复杂特性检测器。这时，它就能够了解4个轮子、挡风玻璃和排气管通常是同时出现的，眼睛、鼻子和嘴也是同时出现的。它不知道的仅仅是汽车和人脸都是什么样的。深度学习网络能够识别的许多特性可能都和手头的任务无关，但是其中有一些特性却是和手头任务高度相关的。

辛顿解释道：“训练这些特性检测器时，每次训练一层，这一层都试图在下面一层找到结构模式。之后，就可以在顶部贴上标签并使用反向传播来进行微调。”结果深深震撼了人工智能界。辛顿回忆道：“其中涉及一些数学问题，这总会给人们留下深刻的印象。”

有关深度学习的消息迅速传开。辛顿实验室的两名成员乔治·达尔（George Dahl）和阿卜杜勒-拉赫曼·穆罕默德（Abdel-rahman Mohamed）迅速论证了该系统不仅能够进行图像识别，还能够进行语音识别。2009年，俩人将其新创建的语音识别神经网络与已经使用了30多年的行业标准工具放到一起一较高下，结果是，深度学习网络获得了胜利。这时，谷歌邀请辛顿的一位博士生纳瓦迪普·杰特列（Navdeep Jaitly）修补谷歌的语音识别算法。看了一眼之后，他建议用深度神经网络取代整个系统。尽管一开始持怀疑态度，但杰特列

的老板最终同意让他尝试一下。事实证明，新的程序比谷歌精心调试数年的系统表现还要出色。2012年，谷歌将深度学习语音识别程序嵌入安卓移动平台，错误率与之前相比立刻下降了25%。

那年夏天，辛顿终于收到了谷歌的电话。这个搜索巨头邀请他夏天到位于加利福尼亚州山景城的校园工作。尽管辛顿当时已经64岁了，谷歌却将他定为“实习生”，因为员工必须严格服从公司政策，即必须在公司工作好几个月之后才能被授予“访问科学家”的头衔。尽管如此，辛顿仍然加入了由20岁出头的年轻人组成的实习生组。他甚至还戴上了新实习生们专用的上面带有螺旋桨图案的帽子，被称作“新谷歌人”（Nooglers）。辛顿说：“我一定是史上最老的实习生。”当时，他开玩笑似的表示，那些并不知道他是谁的年轻同事肯定是把他当作“老笨蛋”了。

辛顿在谷歌的工作涉及为其他潜在的应用提出有关深度学习的建议。那年夏天的工作进展得十分顺利，第二年，谷歌正式聘用了辛顿。除他之外，谷歌还聘请了他的两名研究生，辛顿和这两名研究生共同创建了一家名为“DNNresearch”的公司。辛顿在一篇声明中写道：“我会继续在多伦多大学兼职任教，在那里我还有很多出色的研究生，但是在谷歌我能够看到我们如何处理超大型计算。”

在神经网络领域孤独地耕耘了30年后，杰夫·辛顿最终在世界最大的人工智能公司发挥了重要作用。

人工智能新主流

如今，深度学习神经网络已经成为人工智能的主流，其强调的理念可以追溯到麦卡洛克和皮茨。尽管该理念仍然是对大脑工作模式的一种模拟（我们将在后续章节中探讨更多关于大脑的生物力学模型），但神经网络能够解决问题的广泛性却令人惊叹。传统人工智能一直表现良好，直到后来研究人员才发现现实世界与其完美模型并不匹配。与传统人工智能不同的是，神经网络不仅能够处理规律性事物，还能够处理规则以外的情况。正如20世纪80年代的NETalk一样，这使其成为处理语言等棘手问题的最佳选择。深度学习神经网络还擅长处理所谓的“分布表征”，这意味着其具有模拟同一表征空间中两个相似但独立的领域（例如语言和图像）的能力。从本质上来讲，这意味着神经网络能够以类比的方式进行思考，这一点是传统人工智能无法企及的。

杰夫·辛顿说道：“现在我们看到的许多事物都在使用神经网络。根据经验，如果你想完成一项任务，并且你知道这项任务涉及大量知识，这意味着如果你要学着做这件事，你将需要大量相关的参数。在这种情况下，深度学习将是更好的选择。”

令人印象深刻的应用程序随处可见。2011年，就在辛顿加入谷歌之前的那个夏天，谷歌工程师杰夫·迪安（Jeff Dean）、格雷格·科拉多（Greg Corrado）和斯坦福大学计算机科学家吴恩达（Andrew Ng）共同推出了“谷歌大脑”（Google Brain）项目。谷歌大脑项目都在谷歌公司半公开的实验室“谷歌X”中进行，使用深度学习网络识别高水平概念，例如通过分析视频网站YouTube的视频中静止的图像来识别猫，而之前并不向它解释猫到底是什么。（巧合的是，这实际上就是弗兰克·罗森布拉特半个世纪前对《纽约客》杂志说过的“神经网络终有一天能够实现”的那个目标。）

听起来一台知道猫为何物的计算机并没有什么新奇的，但是通过深度学习实现计算机的视觉能力，在现实世界中却拥有广泛的用途。一家名为“Dextro”的初创企业使用深度学习创造出了更好的在线视频搜索工具。Dextro的神经网络并不依靠关键词标签，而是通过扫描直播的视频来分析音频和图像。举例来说，如果用这个神经网络搜索英国前首相戴维·卡梅伦，那么不仅能够搜出与保守党有关的视频，就连提到英国首相的视频也能够搜到。

与此同时，Facebook（脸谱网）使用深度学习自动为图像设置标签。2014年6月，这一社交网络平台发布了一篇文章，介绍其称之为“DeepFace”的面部识别技术。凭借深度学习的能力，Facebook算法几乎和人脑一样准确，无论光线和相机角度如何，都能够对比两张照片并查看其显示的是否是同一个人。此外，Facebook还使用深度学习创建了另外一种技术，该技术能够为盲人用户描述图像，例如，一张图片上显示的是某人在一个夏日骑着自行车穿过英国的乡间小路，该技术能够用语音将这一情景描述出来。

其他一些项目将深度学习和机器人学结合起来。美国马里兰大学的一组研究人员给机器人放了一段YouTube上的烹饪视频，这样就教会了机器人如何烹饪一顿简餐。这一过程中没有任何直接人为的输入，只要提供正确的餐具，机器人就可以直接复制视频中显示的任务，而且准确率非常高。长远来看，类似的机器

人深度学习也可以应用于军事维修等领域。

目前已经证明，深度学习在翻译领域是必不可少的。2012年12月，微软的研发总监里克·雷斯特（Rick Rashid）展示了一款震撼人心的英汉语音识别和翻译系统。通用翻译器这样如同“星际迷航”的英雄梦一样的技术即将实现，这项技术意味着在不久的将来，我们无须会说法语、俄语或日语就能够在法国餐馆点菜、在俄罗斯坐出租车或在日本谈生意。更加令人印象深刻的是，深度学习系统能够将说话者的语音划分为基本的音素，然后将这些音素重新组合成需要的语言，最后以说话者的声音将语言“说”出来。微软解释道：“你的平板电脑或智能手机将分析你所说的意思，将其翻译成听者能够理解的语言，并用你的声音以听者熟悉的发音、音色和音调表达出来。”

有趣的是，虽然我们一直在对基本的技术进行调整，但如今许多重大进步仍可以追溯到戴维·鲁梅尔哈特和杰夫·辛顿在20世纪80年代发明的反向传播算法。这些年来唯一改变的是计算能力，而计算能力反过来意味着更强大的神经网络和更多隐藏层。仅“谷歌大脑”项目就将16万个计算机处理器连接起来，创建了一个拥有10亿多连接的人造大脑。可用训练数据集的规模也在大幅增长。前些年使用的数据相对较少，与其相比，如今用于教神经网络思考的信息数不胜数，举例来说，Facebook的面部识别系统就是通过分析740万张图像来训练系统的，这些图片是Facebook 12.3亿活跃用户的脸。

神经网络不是如今实践中用到的唯一一种人工智能（我们将在后面的章节探讨其他人工智能），其优势将人工智能推到了胜利的顶峰。与传统人工智能不同的是，神经网络不再局限于简单的实验室环境。

实际上，下一章将探讨的内容是，当人工智能超越我们通常所说的计算机系统的限制并跟随我们一起进入现实世界时，到底会发生什么。

[1]1 泽字节=2⁷⁰字节。——编者注

第三章 万物互联的智能时代已经来临

第三章

万物互联的智能时代已经来临

1998年，苹果公司推出了其外观线条呈圆形的iMac电脑；《哈利·波特》风靡世界；第一款移动MP3播放器上市；一位来自雷丁大学控制论专业的44岁教授在这一年进行了一项非同寻常的运算。凯文·沃维克（Kevin Warwick）教授进行了一个非急需外科手术，目的是将一个包在玻璃管内的硅片植入自己的左臂皮肤之下。一旦植入人体，这款射频识别设备（RFID）的芯片发出的无线电信号，就能经由实验室周围的天线，随即传入能够控制沃维克周围环境的中央计算机。“在（我的实验室）的主入口处，当我进门时，一个由计算机操作的音箱发出‘你好’的声音。”后来凯文·沃维克记下了他的体验：“计算机检测到我进入大楼的过程，当我走近实验室的时候，为我开了门，点亮了灯。芯片植入体内后的9天里，我仅仅沿着特定的方向行进，就可以触发周围的物体自己行动。”

约20年后再来看，沃维克的这项实验依然震撼人心、发人深省。与沃维克职业生涯的其他事情相比而言，这项实验最有意义。然而，在过去的几十年里，我们对此事的惊诧程度可能多多少少发生了改变。尽管回避有人愿意采取这种侵入式手术的原因依然很容易，但关于为什么有人想这么做的问题已经不再重要。写这篇文章的时候，我的手腕上带着一块42毫米的不锈钢苹果手表，搭配了米兰风格的表带。这款表价格为599英镑，它能实现的功能远远超过凯文·沃维克在其植入式射频识别设备上所设想的功能。一旦我收到一条短信或一个电话，或者如果我的朋友在图片分享网站Instagram贴了一张新图，我只需要看看手表就一目了然。而且在超市购物时，我可以用手表刷卡支付。同样，我也可以用手表打开世界各地酒店的房门。外出的时候，手表连续发出的嘀嗒声和震动可以告诉我应该走哪条路。一串嘀嗒声提醒我右转，另一串嘀嗒声则提醒我左转。第一次震动表明我的旅程结束了，而第二次震动则告诉我到达目的地了。所有这些功能并不需要进行侵入式手术。

如果你正在阅读这本书，你很可能对“智能设备”这个名词并不陌生。除了种类日益繁多的智能手表，如Pebble（一款智能手表）、Android Wear（安卓的可穿戴应用程序）及其他设备，还有智能跑鞋，智能跑鞋能够记录步数、心跳频率，并使用嵌入式屏幕传达你的情绪，比如使用笑脸和爱心等符号。智能冰箱不但可以记录温度与冷藏的食品，还会在你最喜欢的食物就要吃完了或将要变质的时候通知你。还有智能安全摄像头、智能厨房秤、智能灯泡、智能马桶、智能尿片和智能牙刷。2014年，谷歌以惊人的32亿美元现金收购了最著名的智能设备公司Nest Labs。Nest Labs由苹果前雇员马特·罗杰斯（Matt Rogers）和“iPod之父”托尼·法德尔（Tony Fadell）联手创立，打造了多款可以联网的智能设备，其中最重要的是智能恒温器，通过一段时

间的学习，这款恒温器可以了解用户的习惯，并相应自动地调节温度。

传感器、人工智能算法与通过Wi-Fi（无线局域网技术）实现的持续联网状态相结合，使这些设备变得“智能化”。以前，接入网络而变得智能是一件令人们不得不“大费周折”的事。今天，我们的在线连接很少出现中断的现象。总的来说，这些进展使我们从用户那里收集数据、分享数据，并且帮助用户理解数据成为可能。“数据赋予我们力量，”世界第一个联网电动牙刷生产商Kolibree的营销与战略总裁勒妮·布洛杰特（Renee Blodgett）表示：“这是我们第一次将刷牙方式、刷牙部位以及刷牙时哪里需要改进结合在一起。”在我们拥有智能牙刷之前（这对我而言，就是现在），我们不得不依靠一年前进行年度检查时牙医的反馈。而通过智能牙刷，我们可以实时获得这些信息。

智能设备成为现代生活的必需品

现在，我们处于未来技术的“早期采用”阶段，未来技术的支持者声称，这些技术将像19世纪末、20世纪初电力时代的到来一样，带来一场巨大的变革。1879年，美国发明家托马斯·爱迪生已经能够在加利福尼亚州门洛帕克市自己的实验室里生产可靠耐用的电灯泡了。到了20世纪30年代，美国90%的城市居民，以及越来越多的农村地区的人们都可以利用这项技术。随着开关的拨动，电赋予人们控制光的能力，人们能够控制自己家和工作场所的光线。这打破了生活的正常生物节奏，使人们能够随心所欲地安排自己的工作和娱乐时间。随之而来的电网引入了大量的连接设备，创造了工业，并永远地改变了人们的生活。

美国西尔斯百货（当时一家初具规模的邮购公司）1917年春季的商品目录使公众知道“电不仅仅可以用来照明”。事实确实如此。铁熨斗、洗衣机和真空吸尘器使洗衣与清洁更加容易。由于效率的提高，不但清洁度上升了，而且家庭雇用的家政人员数量也越来越少。电冰箱取代了冰盒，使食物更加易于长期保存。天热的时候，我们可以使用电扇，而天冷的时候，我们可以使用辐射发热器，这是人类第一次能够控制气温。电力为大众带来了电话与飞机，并在即时通信年代，受到了新闻与娱乐行业的追捧。1938年，美国前总统富兰克林·罗斯福在佐治亚州巴恩斯维尔演讲时宣称，电力是现代生活的必需品。

我们能否开启一条同样的智能设备变革之旅？或许是可以的。当然，移动无线网络的崛起意味着设备的使用比以前更加方便。“物联网”（这个定义有时候显得相当笨拙）之梦是，智能硬件要像一个世纪以前的电力那样，成为21世纪重要的“现代生活的必需品”。那时我们进入了电气化时代，现在我们将进入互联时代。

当前，智能设备领域充斥着大肆炒作之风，爱立信公司的分析师预测，到2020年全球将有约500亿台智能设备，相当于人均6.8台。“这不仅是一场进化，这还是一场革命。”苹果前雇员、现在掌管创业公司SITU（该公司生产量化卡路里摄入量的智能天平）的迈克尔·格罗特豪斯（Michael Grothaus）表示：“这是自个人电脑诞生以来，技术界最激动人心的事了。”

会思考的事物

1991年，剑桥大学计算机科学系特洛伊木马研究室的研究人员提出了一个新的想法。他们在自己的研究室中放置了一个共用的咖啡壶，然后决定安装一台摄像机用以监视一天的咖啡用量。研究人员将摄像机设定为每秒捕捉一帧，然后将其编码为灰度级的JPEG格式文件，最后将图片文件通过早期的万维网发出去。通过各自的计算机，该系研究人员可以登录到“视频”源中查看壶里是否还有剩余的咖啡，从而省去他们去打咖啡的无用功。

“‘咖啡俱乐部’的一些成员位于大楼的其他区域，他们不得不为打咖啡爬上爬下，如果特洛伊木马研究室熬夜的黑客们先打了咖啡，那么其他成员打咖啡的结果常常是无功而返。”当时在系里工作的计算机科学家昆汀·斯塔福德-弗雷泽（Quentin Stafford-Fraser）牢骚满腹地说：“这样打咖啡对计算机科学研究进程造成的中断，显然使我们非常苦恼，于是‘XCoffee’（X咖啡）就这样诞生了。”

我之所以提出XCoffee，因为它证明了一个非常重要的观点，即什么是我们认为的“智能技术”。XCoffee也常常被看作智能设备现代趋势的早期例子。某种程度而言，这是真的。与许多最新的智能配件一样，XCoffee与网络连接，因而也成了所谓的“物联网”的一部分。但是对我而言，XCoffee更接近硬件极客所说的“黑掉”的范例，“黑掉”这个术语就是俗话说的解决棘手问题的高明方法。成为今天我们称作智能设备（麻省理工学院媒体实验室称之为“会思考的事物”）的前提条件是，它必须以一种自我管理的反馈回路而存在，无须过多人工干预就能够自动运行。物联网并不仅仅把“物”连接到互联网。传统互联网使人们能够搜索、下载音乐或者阅读信息。另一方面，物联网主要用于非人类实体的交流，这是越来越多的人热衷于M2M（机器对机器）交流的原因。

智能设备应该能够感知自己所处的环境、识别特定状态、触发评估、产生行为等等，从而形成一个连续的环路。智能设备的“智能”在于中间的部分，那里负责处理感知到的信息，以及如何基于信息采取具体的行动。一台真正智能的咖啡机不只是提醒人们咖啡机空了，而是能够计算出使用者可能口渴的时间，并且自己能及时重新加满咖啡，调制出咖啡成品以满足使用者的个体需求。甚至基于无人控制的桌对桌（desk-to-desk）送货也是可能的。

“控制论”简史

我们将在本章讨论的多数智能设备都包含机器学习的元素。正如围绕人工智能的各种问题都可以回溯至数百年前一样，关于具有自我调节功能的机器的想法也同样如此。早在公元前205年，寓居于埃及亚历山大港的希腊数学家克特西比乌斯就建造了世界上第一台能自我控制的设备。克特西比乌斯的作品是一台水钟，其最大特点就是拥有一个可以保持恒定流速的校正器。这台水钟通过设在水缸里的浮子计时。水从水缸底部的小孔滴落，浮子就随着水位下降。每运行一单位的时间，浮子顶端的类似于人偶的器械就进行一次齿轮机械操作。克特西比乌斯水钟有多个不同版本，在不同版本中，它要么落下一块卵石，要么鸣响一声喇叭。

克特西比乌斯水钟意义重大，因为它永久性地改变了我们对人造之物的认知。早在克特西比乌斯水钟之前，人们认为只有有生命的东西能够根据环境的变化调整自己的行为。而克特西比乌斯水钟诞生之后，自我调节反馈控制系统成了我们技术的一部分。

进入20世纪，影响后世的人工智能先驱诺伯特·维纳（Norbert Wiener）制定了反馈系统的数学理论。维纳提出一个设想：智能行为是接收和处理信息的必然结果。这个设想就是众所周知的控制论。“二战”期间，当维纳与其同事朱利安·毕格罗（Julian Bigelow）在从事旨在提高高射炮精确率工程的时候，他的反馈系统理论得到了细化。维纳和毕格罗解决了向飞行中的飞机提高开火准确率的难题。这曾经是个难题，因为炮手必须预先判断目标的位置。他们的解决方案是通过预测目标飞行位置并相应调校火炮的瞄准器，从而自动调整炮手的瞄准过程。

维纳关于感知和反馈作为一种优化性能的方法的设想不仅仅只是用于战争。维纳与之前的研究者不同，他将反馈构想成一种通用的普适原则。他认为，反馈能够以同样的方式应用于机器、组织、城市甚至是人的大脑。他在1905年出版的《人类的人类用法》（*The Human Use of Human Beings*）一书中记录下了许多这样的设想，此书比“人工智能”的正式问世早了6年。作为一本出人意料的畅销书，它描述了智能自动化推动社会进步的各种方式。维纳抛弃了建造能够思考的机器来替代人类的想法，而是在他的书中讨论了人类与机器可以合作的方式。在导读中，他写道：

这是本书的论点：只有通过学习属于社会的信息与掌握通信设施才能了解社会；而且，在这些信息与通信设施未来的发展过程中，人与机器、机器与人以及机器与机器之间的信息注定要发挥越来越重要的作用。

控制论从来没有像人工智能那样获得过大量的研究经费。然而，关于可以用于预测未来的数学反馈系统的设想几乎是建造今天所有智能设备的基础。例如，标准的“无声”恒温器通过传感器收到温度信息，并根据其冷热程度，为你开启火炉或空调。另一方面，一个“智能”恒温器能够整合其他数据源，如当天的天气预报或家里人对房间温度的历史设定信息。它甚至可以根据房间内多人的身体传感器读数的集合，选择一个平均的温度。代替那种简单的反应式工作，智能设备的工作变成了预测式的。

这要求不同设备之间相互作用。与预先连接的同类设备相比，这些智能设备可能是相对智能些，但离我们实际称之为的“智能”还相去甚远。但是当设备彼此之间能够分享数据和目标时，新的可能性就展现了出来。这就是专家所描述的“环境智能”，即通过使用嵌入网络的智能，多种设备共同执行各种任务。就像白蚁共同建造一处蚁穴一样，整体是大于部分的总和的。

早期的自动机器人

这种对机器与环境（或者更好的情况是，多种机器与其环境）之间突发行为的兴趣源于控制论运动，并引发了机器人领域的一些早期重要工作，如威廉·格雷·沃尔特（William Grey Walter）所从事的工作，他是一位出生在美国而生活在英国的神经科学家。1949年，沃尔特建立了世界上第一对三轮机器人，他称之为“乌龟”。与刚刚开始研究数字计算机的计算机科学家不同，沃尔特依靠模拟电子学来仿制其机器人的大脑。他的目标是证明少量脑细胞之间丰富的关系能够产生复杂的行为。他对这样的概念十分痴迷，即机器可以定义目标，并随后通过学习自己的行为产生的后果从而完成目标。

沃尔特的“乌龟”分别名为埃尔默（Elmer）和埃尔希（Elsie），都装配有光敏元件、标志灯、触摸感应器、推进马达、转向马达和保护壳。尽管这对机器人还不能可靠地工作，但它们能够自动探测周围的环境。在沃尔特所著《活着的大脑》（*The Living Brain*）一书里，他回忆了一段经历：一位年长的女士认为这对自主漫游的机器人在追逐她，于是逃上楼将自己锁在卧室。在沃尔特工作的位于布里斯托尔的博尔顿神经学研究所（Burden Neurological Institute），沃尔特在技术人员W. J. 邦尼·沃伦（W. J. Bunny Warren）的帮助下，使乌龟机器人得到了改进。他在1951年的“不列颠节”（Festival of Britain）上展示了后续三台“马基纳·斯巴卡拉特里克斯”（*Machina Speculatrix*）机器人，它们基于埃尔默和埃尔希原型进行了许多重大改进。其中包括当电池即将耗尽时，机器人会转身向光源前进。今天，虽然人们几乎已经遗忘了沃尔特的乌龟机器人，但是它们是早期自动机器人的典范，能够通过自己的行为，以试错的方式进行学习。

谈到威廉·格雷·沃尔特的乌龟机器人的后续产品，就不可能不提到iRobot公司创造的真空清洁机器人Roomba。Roomba呈小型圆盘状，在计算机的引导下可以在家里自动工作。尽管它可以通过基于反馈的“智能”对刺激做出反应，但一般情况下，它遵循一系列预先设计的清洁策略。首先，它会一直清理直至与障碍物发生碰撞，碰撞指示其改变线路并以新方向重新开始清理。为了使之有效移动，Roomba包含了许多智能传感器，其中两个是红外传感器，帮助它检查墙体以及被它称为“悬崖”的物体，比如楼梯和其他会造成下落的地方。当Roomba撞到障碍物时，触摸感应缓冲器会阻止其向前行进。在Roomba的下面装配有俗称的“压电传感器”，可以检测到灰尘。如果在一个地方发现过多的灰尘，Roomba将重复其步骤以进行第二次清理，第二次速度将放慢并清理得更加彻底。仅仅观察这些简单的步骤，Roomba就展示出了一看像是由人执行的突发行为。

某种意义上而言，“突发”这个词表明这种行为是不可预测的，其实不然。如果完全基于上述的简单规则，我们可以理解为什么Roomba能以自己的方式行动。然而，如同沃尔特的“乌龟”一样，当Roomba设法完成自己的任务时，行为实体（behavioural agent）与环境的结合可能产生一些意想不到的响应。

单独一台Roomba的运行和表现都非常好。但是，就像沃尔特通过其乌龟机器人所发现的，当不止一台实体相互作用的时候，事情就真的变得很有趣了。沃尔特最有趣的观察结果是，当这些“乌龟”彼此围绕旋转的时候，他发现了它们“跳舞”的方式。这种舞蹈由一种看上去由机器人仪式化的碰撞和后退组成。这是他装

在乌龟机器人身上的标志灯造成的，当转向马达开启的时候标志灯就亮了，而转向马达停止时标志灯就熄灭了。由于每个乌龟机器人依靠对方的标志灯来定位，它们就像同一物种的两个生物首次见面一样彼此吸引。当“乌龟”走过显现它们身影的镜子时，会发生同样的现象。沃尔特宣布，如果这是一种动物行为的话，这种行为“或许可以证明乌龟机器人具有自我意识”。

即使Roomba的热衷者也不愿意承认两台交互的真空清洁机器人具有“自我意识”，但沃尔特表明多重代理系统使智能设备变得更加有趣，这点是没有错的。举例来说，如果你家的门能够自动开或关，从而使Roomba可以一次清扫多个房间，这将产生什么效果呢？这在某些场合是令人满意的，比如，如果你有一只宠物而你希望它进入某一房间，或者如果你有一个特殊的房间在供暖，你不想它进入这个房间。同样，如果Roomba能够接入装在前门或汽车里的传感器，并且知道在你去上班的时候就开始工作，那么你回家的时候清洁工作就已经完成了。或许，这就是那些制造智能设备的大公司正在努力的方向。

计量生物学上的记录者

2015年6月之前，我从来不会花很多时间考虑诸如哪个城市的居民睡眠最少，抑或通勤不足5英里的上班人士是否比距离更远的人锻炼得更多这样的问题。然而对于感兴趣的人而言，答案分别是：日本东京的市民睡眠最少（平均每天睡5小时44分钟）；“是的，通勤不足5英里的上班人士比距离更远的人锻炼得更多”（每天多走422步）。

这是班达尔·安塔比（Bandar Antabi）告诉我的。安塔比无疑是世界上最佳的酒吧竞猜选手。你若问他，他会告诉你，如果你希望夜晚早点降临的话，那么你最佳的居住地是澳大利亚的布里斯班，在那儿，人们大约晚上10点57分就要进入梦乡，而“夜猫子”的首选居住之地应该是俄罗斯的莫斯科，那里通常凌晨12点46分才是人们入睡的时间。他说，在情人节女人会比平时少吃约3%的大蒜，但是当天男人会多吃37%。瑞典的斯德哥尔摩人是最活跃的步行者（按每天平均量计算），而巴西圣保罗人是最不活跃的步行者，等等，就像你把达斯汀·霍夫曼（Dustin Hoffman）在《雨人》中的角色放在维基百科上数小时一样，很快你就得到了所有答案。

班达尔是个非常聪明的人，但他也是个非常不擅长处理琐事的人。他能采集这些信息是因为作为特殊项目主管，他所任职的公司Jawbone已经花费多年时间不知疲倦地在收集这些信息。

1999年，Jawbone以为美军开发降噪技术而起家，随后才涉足蓝牙头戴设备、扬声器以及后来的可穿戴生活记录仪等领域。就是可穿戴生活记录仪这种装满传感器的智能设备使Jawbone今天闻名于世，如UP3，它是一种如同手表的细腕带，专注地记录着从你的睡眠模式、呼吸节奏、心率到“皮肤电反应”等所有事情的设备。Jawbone的大量用户所生成的原始数据使班达尔知道了如此多的“真相”。现在，这些数据包括了3万亿步、2.5亿次睡眠以及将近200万顿饭。随着时间的流逝，数据将继续增加，还可能纳入几十种其他的计量内容，如用户每天摄入的咖啡因总量等。简而言之，Jawbone希望你成为计量生物学上的记录者。

“我们的任务是建立这种个性化的数据集，它整合了你的个人身份、档案、生物学信息、年龄、身高、性别、饮食偏好、情绪等信息。”班达尔告诉我这些的时候，我在Jawbone英国公司的总部诺丁山办公大厦13层，坐在他的对面，喝着一杯星巴克咖啡。他继续说道：“我们也想了解你的相关活动。你什么时候坐着，什么时候活动且消耗卡路里，你的睡眠质量如何。通过挖掘这些信息，随着时间的推移，我们可以为你提供大量的信息。我们正在建立一个关于你的健康的场景化数据集。”

Jawbone与许多技术硬件公司建立了有经纪人参与的合作，但是如果只是分享数据的话，这些业务是不值得我花费笔墨的。你的恒温器真的需要知道在昨夜的晚餐中你吃了什么吗？如果你的电视知道你一周要慢跑4次，这会对你有什么好处呢？班达尔说，实际上这意义深远。“有数据虽然是好事，”他对我说，“但是理解数据才是我们关注的。”

“理解数据”意味着可以通过恰当的人工智能算法，以具有上下文意义的方式分析你的数据。“我们可以

使用这种技术，以一种有利的方式将数据用于适合的设备。”他继续说道，“比如，你可以将Jawbone的智能设备与你的智能恒温器配对，那么当你睡觉的时候，卧室里的温度可以自动调节至最有利于你睡眠的状态。当你醒来时，温度可以再次改变。”

这些数据处理即通常所说的事件驱动程序或者IFTTT规则（通过不同平台的条件来决定是否执行下一条命令）。这些简单的规则，依据简单的方法，将服务环节串联了起来。IFTTT规则先驱林登·蒂贝茨（Linden Tibbets）曾将这些规则称为“数字传送带”，因为它们可以使智能技术的创造者或用户将完全割裂的概念联系起来。这种在智能设备领域可能或者当前正在发生作用的交互例子不胜枚举，比如，如果你的汽车知道你昨夜没有睡好，它可以从你的智能恒温器提取数据，这些数据如果显示出你遇到寒冷会更精神，它就可以打开空调，确保你能保持足够的清醒。通过你的可穿戴健身记录仪，它可以了解到当你听某一类型音乐时的表现最佳，因此，它可以自动播放金属乐队的音乐使你一天都活跃起来。它甚至可以知道昨夜你和朋友外出聚会，现在仍然还处于醉酒状态。为了实现这点，它使用嵌入在变速杆上的传感器，分析你手掌汗液里的酒精含量。如果遇到这种状况，它会让汽车熄火，并建议你呼叫一辆优步（Uber）出租车。

还有一个例子，你的智能电视可以收到你的睡眠记录，并可以基于你一天的时间安排为你提供定制化收看电视节目的时间建议。如果晚上9点放弃看让你脑子兴奋好几个小时的《权力游戏》（*Game of Thrones*）这样的节目后，为什么不选择看《摩登家庭》（*Modern Family*）呢？或许你在观看一个你喜欢的烹饪节目，智能电视就将节目中的食谱发送至智能冰箱，冰箱监控着所有食物，所以它知道里面是否存放有烹饪这道菜所必需的各种原料。如果没有的话，它可以将所需原料加入家庭采购杂货的快递清单之中。随着越来越多的设备与网络连接，能够提取彼此的数据并由事件驱动程序将它们相互关联，技术迷期待已久的梦想即将实现了。

当然，我想谈的是羽翼已经丰满的智能家庭。

未来家庭

[1]的智能电视用你熟悉的声音说“欢迎回家”，并建议你看看一段昨晚球赛的集锦，因为它知道你还没有看。

除了拥有能够彼此交流的设备，联网的智能家庭与现在的家庭一个很大的区别将是使用人工智能去建立目标，我们的设备可以根据这些目标努力使我们的生活过得更加轻松，更加舒适，更加富有成效。

“实现物联网的设想，使用所有这些可以进行一些思考的设备，可以采用以下两种方式中的一种。”被称为人工智能“强化学习”领域（涉及人工智能形成并追逐目标的能力）的专家理查德·萨顿（Richard Sutton）表示，“你可以拥有执行单独目标的独立实体，比如，恒温器的‘目标’可能是提高效率并节省燃料。冰箱的‘目标’是在你需要的时候确保随时为你提供食物。这样做的结果可能是你的智能设备为计算出各自目标的优先级而争斗不休。而另一种方式是将所有设备互联，从而形成一个决策者。”

举出类似的例子并不难，比如拥有一套使你更加健康或者能节约燃料的住宅。当然，并不是人人都喜欢这样。“这里需要对第一种选择进行解释，”萨顿接着说道，“这意味着你知道你的炉子是开着的，因为房子温度低。而不是因为房子希望你待在室内以便你的智能电视可以为你播放最新的节目。它具有一个清晰的目标，而且你们都知道它在做什么。住宅自己的可信度赋值非常简单明了。”

但是，只要它们以一种对于用户来说清晰透明的方式存在，并还存在实现更宏大目标的情况，人工智能就能够使你的设备联合起来，从而实现长期的且更加复杂的目标。

机器学习的普及

黛安·库克（Diane Cook）是华盛顿州立大学电子工程与计算机科学学院的一名教授。过去数年，她都在调查智能家庭改善老年人生活质量的方式。几年前，库克参观了得克萨斯州博览会，在博览会上她看到了以一系列智能设备为特色的“未来家庭”展。离开时，她却不为所动。“这根本不是智能家庭，只是个联网的家庭，”她说，“这里有一台可以扫描二维码的冰箱，它随后可以为你生成一个杂货清单，并将清单发送到当地的食品店，食品店再送货上门。这个家庭里虽然有很多这样的设备，但拥有‘智能’的仍然还是住在里面的人。这些设备没有逻辑推理能力，只有信息。”

库克运用机器学习的知识，希望建造一个不仅仅只是收集数据的房子。“我认为智能家庭不仅要感知环境中发生的事情，而且还要通过自动化对环境造成影响。”她说道，“它可以对收集到的信息进行推理，使用这些信息并自动选择一种行为。”库克开始着手一项旨在使用这种智能进行推理的工作，从而获得老年人在认知与身体差异方面的早期指标。智能传感器可用于告诉我们一个人在家活动的全部信息。如红外运动检测仪、磁力门窗警报器以及可以记录水龙头和炉子状态的传感器等设备能够显示一个人是在吃饭、睡觉、做饭、看电视还是在外出散步。通过监测这些活动，有意识地提取他们的“统计学活动特征”。库克的算法还可以预测一个人的活动进行情况。比如，一个有记忆障碍的人执行某一任务所花费的时间可能更长。他们可以展现出许多可识别的迹象，如徘徊过多，在具体事情中试图回想起下一步该做什么，反复开关壁橱，或是使用不正确的工具做饭等。孤立地来看，这些行为不一定有什么意义，但从全局来看时它们勾勒出一幅充满启示的图画。

开始的时候，库克和她的团队在华盛顿州立大学的校园测试这项智能家庭技术，随后他们搬至西雅图当地的一家名为“地平线之家”的疗养院。这里共有18位老人，平均年龄至少73岁，他们志愿加入这项研究。他们在老人的公寓里安装了传感器，传感器的外观是白色小盒子，每两英寸安装一个。即使没有摄像头，这些传感器也能够分辨出共同生活的是两个人还是一个人和一只宠物，库克称后者为“基于智能家庭的角度不得不处理的噩梦般的场景”。库克随后将传感器数据与（人类）护工所管理的正常检测数据进行了比较。“这是令人吃惊的成功，”她说，“我们在人类的活动与传感器的健康检查活动之间发现了高度的关联性。因此，我们仅仅依靠他们进行的一些活动，就能够通过机器学习工具，成功地预测他们正在接受怎样的诊断治疗。”

如库克所言，虽然这项技术的应用没有打算代替朋友或家人之间的社交活动，但它能帮助老年人独立生活得更久，这可能意味着他们可以继续待在他们度过了前半生的家里。“即使没有护理人员登门，这也没有问题。如果出现重大的变化，意味着健康状态发生转变，护理人员或者医务人员也可以及时得到警告。这种智能家庭可以发出关于老年人瞬间记忆丧失的警报，如冰箱门敞开或炉子没关等。”

库克的项目已经收到约300万美元的资金，她的下一个目标是扩展自己的研究。“今天我们包里都有智能手机，”她说，“我们可以将所有的传感器、网络、软件以及计算机放入一个小容器里，并将它们送到全球各地的站点。”不久，它就可以用来进行国际研究——随着全球各地数据的流入，机器学习算法会变得更加智

能。

许多公司也致力于这个领域。比如：Healthsense公司制造了eNeighbor监控系统，这是一款由大量家用智能传感器强化的可穿戴设备。与黛安·库克在“地平线之家”进行的项目有些类似，eNeighbor可用于检测患者是否跌倒或忘记服药，并通知护理人员。同样，BeClose智能传感器系统可以发现患者长时间不吃饭或错过吃饭，并向指定的家庭成员发送短信、电邮或呼出电话。

拥有智能家庭信息那当然很好，但智能设备的下一波浪潮将涉及疾病的追踪与诊断。以医疗设备厂商AliveCor为例，它生产的智能手机壳可以兼作可移动EKG（心电图）心脏监护器，能够预测使用者是否要中风。这种手机壳通过人的指尖测量来获取心电图，随后由算法分析心跳的规律，并告诉使用者是否应该看医生。

随着我们的环境越来越智能，我们将进入一个持续地进行风险实时评估的时代。这是史上第一次可以针对个体得出大量基因的、生理的、生物的，以及环境的因素之间恒定的关联关系与可能的因果关系。除了空气质量与噪声水平等外部指标，可穿戴设备将持续监测我们的心率、血氧水平、身体活动、呼吸模式、面部表情、肺功能、声音曲线、脑电波、姿势、睡眠质量等指标。运用人工智能的洞察力，这些数据点不仅转化成对整个生活的全面建议，而且转化成能够即时提高健康水平的可执行的建议。在执行预测和诊断的同时，我们可以准确了解特定疾病或症状出现时的必要条件，并且可以制定前瞻式预防措施，从而确保病症不会发生。哮喘患者可能具有导致疾病发作的特殊诱因，例如寒冷、运动、花粉或者其他过敏原，这些诱因可以通过智能设备分析出来。当这些风险反复出现时，患者可以收到警告以提前服药，或者避开特殊的地点。还有一个例子，通过细微的声音颤动和降低音量等这些不易为人耳察觉的方式，一位对此还未察觉的遭受神经退行性紊乱的帕金森综合征患者可能早在医生诊断出来之前就接到了病症即将发作的警告。尽管当前还没有治愈帕金森综合征的方法，但早期诊断可能有助于改善生活质量。

尽管这些案例中的数据可用于所有用户，但对用户来说并不需要看到这些数据，除非有需要关注它们的原因。比如，健康追踪技术的默认模式可能是一条高级命令：“监控我的生命体征，如果它们正常，不要传送任何消息。”如果发现了潜在的重大变化，系统将向用户报警，或者以其他方式向他们的医生报警。

在医疗领域，这是一种全新的技术，但是由于机器学习的普及，它在我们生活中其他领域的应用也再普通不过了。比如，当前银行用于检测欺诈的算法。尽管我们有能力查看自己账户里的每一笔交易，但当银行发出的通知偏离我们的常规使用习惯时，我们就可以得到警告。如果我通常都是定期支付不到100英镑，但有一次突然一次性在线支付了1 000英镑，那么这次支付就很可能被标记为可疑行为。许多电邮系统也使用机器学习，把“垃圾邮件”从我们需要阅读的邮件中清理出来。垃圾邮件过滤器基于一系列内置规则，通过对每一封收到的邮件进行评分来展开工作。由于垃圾邮件过滤器能观察我们对所接收消息的不同反应，经过一段时间后这些评分系统就能构建完善。我们只阅读符合评分标准的邮件，即垃圾邮件过滤器认为值得阅读的邮件。

智慧城市

类似的智能技术也致力于改变我们的城市。城市的成长和信息技术的扩张一直都紧密相关。1910年，历史学家赫伯特·卡森（Herbert Casson）写道：“没有比电话更及时的发明了。在需要组建大城市与社区联合体的关键时期，电话来临了。”由于各种技术的网络彼此连接，共同创造财富与生产力，在媒体上，城市通常被描述成有生命、会呼吸的实体。在弗里茨·朗（Fritz Lang）1927年的科幻电影《大都市》（*Metropolis*）里，这一点表现得更加鲜明：在电影里有名无实的大都市被描绘成了一个庞大的生物体。

人工智能的梦想似乎使我们离这种可能性更近了一步。1964年，也就是纽约世界博览会举办的同一年，英国建筑师罗恩·赫伦（Ron Herron）提出了“行走的城市”的概念。如同在前卫的建筑期刊《阿基格拉姆派》（*Archigram*）中所描述的，赫伦支持建造巨大的人工智能移动机器人平台，这种平台就像背着摩天大楼在地球漫步的蜘蛛。这些四处行走的城市可以在地球上无国界生存，可以随意前往它们需要获取资源或制造能量的任何地方。赫伦的城市甚至还有相互连接以创造更大“行走都市”的能力。这种城市不仅可以自给自足，而且由于人工智能的突破性发展，还能够自治。

鉴于那时机器人学的研究水平，幸好罗恩·赫伦的想法从未得到严肃对待。就像在第一章所说的斯坦福国际研究所的机器人沙基，它甚至不能顺利地通过门厅。因此，我们没有亲眼看到自行漫步的纽约遇到地上的小坑就摔个四脚朝天的情景算是天大的幸事了。城市发展到赫伦所设想的样子是不可能的，但城市变得更加智能是确定无疑的。比如，智能办公室将配备装有传感器的废纸篓，一旦纸篓需要清理就会提醒清洁工人。员工也不必知晓卫生安全守则，因为办公室会持续监控自己的温度，并将之与法律规定的标准相比较。如果超过标准，警报就响起，计算机也将自动关闭。在商店、酒吧、主题公园和博物馆，蓝牙热点会根据你的位置与个人偏好，将与你相关的信息传送至你的手机或可穿戴设备。

未来几年可以在大街上看到的最大变化将是无人驾驶汽车的大量出现。如上一章所述，谷歌与苹果延续波默洛研究员的工作，都对这个领域产生了兴趣，并且似乎要在将无人驾驶汽车引入主流社会方面发挥主要作用。无人驾驶汽车不仅在个体层面对人们产生影响，而且因有助于减少城市交通拥堵，还能够在集体层面产生影响。随着城市不断扩张，它们收集的数据对于城市规划人员来说至关重要。我们已经开始看到它在发挥作用。2015年年初，谷歌应用程序Waze与波士顿市合作，用以缓解该市的交通压力。波士顿政府同意提前向Waze发布计划封闭道路的通知，而Waze则同意共同分享该应用程序上重要的城市交通管理中心的数据流。短期来看，通过这次合作Waze可以更有效地帮助用户更快捷地抵达自己的目的地；长期来看，Waze的数据可以帮助波士顿精确调整交通灯的时间，并计算出缓解交通拥堵的方法。

做好互联网消失的准备

2015年1月，谷歌董事会执行主席埃里克·施密特在瑞士达沃斯世界经济论坛发言时掀起了轩然大波。当被问及对未来互联网的看法时，施密特答道：“我的答案非常简单：互联网将会消失。”

当然，这个答案一点都不简单。初听起来，这有点像苹果首席执行官蒂姆·库克向员工的讲话：他们应该放下智能手机，并与朋友面对面地交谈。或者像电影工作室的老板所说：电影千篇一律，人们应该多些时间阅读或散步。实际上，施密特说的与此毫不相干。恰恰相反，他已经观察到，近些年来，技术已经变得更加智能而且无处不在。他并不是第一个提出这种建议的人。早在1991年，硅谷传奇实验室施乐帕克研究所（Xerox PARC）首席技术专家马克·韦泽（Mark Weiser）就曾写过一篇名为“无处不在的计算”的文章。文章开头是这样的：“影响最深远的技术是那些消失的技术。它们将自己融入日常生活之中，以至我们无可分辨。”

其实事实就是如此。第一章所介绍的开创性数字计算机ENIAC重约27 000公斤，占据了整个房间。21世纪初生产的翻盖手机（甚至不是一款智能手机）的重量是ENIAC的近12万分之一，功耗约为其40万分之一，但性能却是ENIAC的13 000多倍。随着可穿戴设备将取代现在我们智能手机上的大多数功能，这一转变将继续下去。正如在蓝光时代，镭射光盘对于我们来说是个古董，几年之后，第一代可穿戴设备也将变得非常可笑。谷歌已经开发了智能隐形眼镜，能够测量佩戴者眼泪中的葡萄糖水平，并随后将这些信息通过无线网络传递至所连接的智能手机。在这个过程中，搜索巨头希望能消除糖尿病患者对比较痛苦的日常血糖的依赖。在更小的范围内，另一家名为斯克里普斯健康中心（Scripps Health）的公司正致力于开发一种使用者可以将其注入自己血管的纳米传感器。一旦进入血管，它将驻扎在人体向各个器官供血的毛细血管床上，采集相关读数，并将数据传回用于分析的主设备。使用者甚至不用担心传感器是如何充电的，因为它能够像一座水力发电站一样，利用流经它的血压进行工作。

这些技术实际上人眼是看不到的，但更重要的是，穿戴者也看不到。现在，智能设备依然要求我们人工去做大量事情，这就意味着我们还没有进入无缝智能交互的领域。举例来说，相比夜间睡眠数据集，Jawbone用于分析膳食情况的数据集相对较小。这是因为人们当前不得不自己记录食物，从长长的单子里每一次挑选一种食物。如果人们可以对煎蛋卷或薯条等食物拍照然后再进行图片识别，并相应地记录下来，那么人们很可能会这样去做。这是技术公司正在探索的领域。2010年，技术巨头高通公司申请了一项专利，这项专利可以让使用者仅仅通过智能手机或智能手表对设备拍照而使设备自动匹配。一旦图形识别工具识别了新设备，两个机器的接口就能自动启动配置流程。这比复杂的人工配置流程要简单得多，也是许多智能手表生产商使用这种技术的原因。在类似技术可以连接到食物数据库之前，可能还需要很长时间。

本章开始部分介绍的SITU生产智能天平的企业家迈克尔·格罗特豪斯表示：“这当然是技术公司正在钻研的事。现在，我们看到了一些真正令人感兴趣的工作，这些工作是通过被称为光谱分析仪的设备完成的。光谱分析仪使用光测量物质的成分。但问题是，光谱分析仪仍然无法准确读取物体的成分与质量。有一天，它

们会变得足够小且足够便宜，我们可以便捷地测量任何物体，但是在那之前，测量食物热量的最佳方式还是用手工记录。”

格罗特豪斯等技術人員的理想是，所製造的智能設備不僅在外形規模上是不可視的，而且在使用方式上也是不可視的。就像我們不必有意識地去關注我們的心率、體溫或呼吸，因為有中樞神經系統在調控，未來智能設備也將漸漸地採集並交流信息，而無須個人對流程進行監控。

智能设备存在的问题

[2]。通过这么做，有效的铲投速度可以保持得更加持久。这恰恰就是当前智能设备可以轻易测量并反馈给老板的情况。亚马逊当前在自己的工厂里使用类似的技术，将手持式电脑配发给“成品合作者”（也称产品采集器），以记录他们完成单个订单的速度。泰勒的科学管理设想不只是支持雇主。他坚信，测量工作的能力也将与激励报酬齐头并进，因此生产力不足的低业绩员工不会获得和高业绩员工一样多的收入。尽管所有这些在理论上几近完美，但批评人士仍指出一个事实：科学管理同样降低了自主性，而且人工智能这个概念颇具讽刺意味，它对待人如同对待机器一样。

另外，如果我们设备的某些方面旨在完全为我们造福，我们可能比较幸运。2014年，《福布斯》杂志的两位作者披露，智能设备制造商Nest已经与电力公司达成交易，会为它们提供显示其用户习惯的数据。尽管这些数据是匿名的，而且只是汇总数据，但电力公司仍然能用这些数据控制我们家里的智能设备。为了减轻电网的荷载，电力公司可以要求Nest在炎热天气里关闭用户的空调。Nest与电力公司分享节约的成本，而用户什么也没有得到。随着时间的流逝，Nest与电力公司交易产生的收益将使其销售恒温器的收入额“相形见绌”。Nest的智能设备依然为其主人服务，但对于我们一直期待的智能设备而言，这只是不同的主人而已。

由于用户数据由智能设备采集并用于城市规划，因此可能面临诸多相关挑战。根据人工智能的采用方式，智慧城市不是变得越来越紧密，而是变得越来越分散。麻省理工学院计算机科学与人工智能实验室（CSAIL）创立的一项深度学习项目发现，通过查看图像，某一地区的犯罪率是可以预测的。除了整合如旧金山犯罪定位地图（San Francisco CrimeSpotting）等应用上的犯罪数据，深度神经网络还对400万张谷歌街景图片进行了训练。深度神经网络很少专注于具体图像所呈现的内容，而是主要专注于推理。项目创建者之一的阿迪亚·科斯拉（Aditya Khosla）对我说：“我们努力在做的是，使展示出的图片研究不只限于分析看到的景象。如果人工智能的目标是建造可以模仿人类智能的机器，那么拥有抽象思维能力明显就是第二步了。”就像上一章所提到的大多数应用软件一样，计算机科学与人工智能实验室的项目是深度学习运转中一个令人印象深刻的案例。但是人们对它的使用方式有着不同的诠释。比如，城市规划者可以利用神经网络查明城市各个部分的需求详情，如哪里需要投资，哪里需要建立医院或学校但不是现在建立（神经网络的另一个用途）。与此同时，汽车公司可以利用同样的技术自动控制车门来锁定你的汽车，或当你需要时为你提供一条备选路线。

如果考虑到了这么多，你就不会因为想把一些工作转交给一个你可以信任的数字实体或者智能助手而受到责备。

幸运的是，人工智能在提高这种能力上也有所帮助。

[1]1英寸=2.54厘米。——编者注

[2]1磅= 0.453 592 4公斤。——编者注

第四章 人工智能助手如何为我们效劳

第四章

人工智能助手如何为我们效劳

不管怎么看，机器人Negobot都像一个14岁的小姑娘。她说的话时常很无趣，而且她只对流行乐队和服装品牌感兴趣。她写的文字中充满了LOL（大声笑）等网络语言，并时不时夹杂着表情符号（小卡通笑脸，一种在网络上表达感情的图片）。有时她说起话来出奇的老成，而有时则相当幼稚。

现在，她正竭尽全力来弄清你是不是有恋童癖。

西班牙德乌斯托大学的一群研究人员创造了Negobot。她（更准确地说应该是“它”）是一个智能实体，用于在线模仿青少年的言语和行为。如今，年轻人花越来越多的时间在虚拟世界中交流。Negobot将充当秘密数字实体，通过潜入网络聊天室找出可疑人员。

德乌斯托大学“智能、语义和安全实验室”研究员卡洛斯·劳尔登（Carlos Laorden）说道：“Negobot的任务就是设法抓到那些狡猾的、有恋童癖的人。他们通常会通过长达数月的交谈来迷惑受害人。因此需要花费大量时间对这些聊天室进行监督。而创造Negobot的设想就是使其在一段持续的时间内模拟人类对话，而不是仅模仿几分钟。”

卡洛斯·劳尔登最初从事的工作是编写过滤垃圾邮件的程序，这是一个传统的机器学习问题。通过利用相关语言，在线找出并隔离恶意行为，他提出了实现这一目标的最先进的现实版本。

Negobot被设定为根据博弈论的原则来进行操作。博弈论这一概念最早是由数学界先驱约翰·冯·诺依曼提出的，我在第一章中曾简要提到过他的主要工作。博弈论研究的是战略决策，众多参与者都有自身的动机，而结果则取决于不同参与者的行为。并不是每个人都能得到自己想要的。博弈论的目标是预测人们如何采取行动并将其转化为自己的优势。

在Negobot的案例中，博弈的目标是弄清与Negobot交谈的人是不是有恋童癖。虽然Negobot在交谈中表现得好像只是在被动地回答问题，但是与此同时，它也将提取尽可能多的可以指控对方的证据。与Negobot的对话内容一开始很平常，之后随着对方给出的回答逐步升级。根据场景的不同，人工智能一共可以分为7个不同的行为等级，每一等级都对应着与之对话的人不同的“狡猾程度”。最开始，Negobot谈论自己最喜欢的电影、音乐、个人风格和服饰，以及更具暗示性的话题，例如药品、酒精和家庭问题。根据人类参与者引导问题的方式，Negobot将进一步探讨性和其他禁忌话题，同时假装给出更多“个人”信息。

不知情的人正在为发现一个与之对话的“14岁女孩”的隐私而暗自窃喜，并自认为成功地操控了整个对话。然而，在这一切发生时，Negobot正在建立一个针对他们的案件卷宗。

卡洛斯·劳尔登说：“我觉得这是一个非常有用的自动化工具，可以用来确定潜在嫌疑人。如果我们使用像Negobot这样的工具，就能够大幅减少目前忙于抓捕这些罪犯的人类团队的工作量。”

打败图灵测试

诱捕法则指的是尽管目前世界各地的警察并没有使用Negobot，但这并不会使该试验失去意义。如果说有什么影响的话，那就是，它强调了人工智能对话应用范围的广泛性。从根本上来讲，Negobot使著名的人工智能试验“图灵测试”发生了独特的改变。

我在第一章中讨论过艾伦·图灵所做的工作。图灵测试以艾伦·图灵的一个假设为基础，旨在测试一台机器是否具有显示与人类一样的智能行为的能力。图灵测试定期举行，测试包括一台计算机（A）和一个人（B），他们分别与一个询问人员（C）对话，而询问人员的任务则是指出A和B中哪个是人类，哪个是计算机。按照图灵的说法，如果C无法完成这个任务，那么“获胜”的就是机器。既然我们无法将机器与我们自己的人类智慧区分开来，我们就必须承认机器是智能的。未来，像Negobot这样的工具将向我们表明，我们区分真人与机器人的能力甚至会产生法律纠纷。

目前在世的专家中，休·勒布纳（Hugh Loebner）对图灵测试的理念最为推崇。留了一头染黑的头发的休·勒布纳是一个有趣的人，自称自我中心主义者。他最初靠出售折叠迪斯科跳舞地板起家，现在已经70多岁了。在中年时，他做出了参加当今世界上最著名的人工智能竞赛的决定。他所命名的“勒布纳奖”能够取得巨大的成功，这一点格外令人惊讶，因为他在计算机科学方面没有任何资质。勒布纳带着他特有的豪情说：“未来可能会有更丰厚的奖品和更知名的比赛，但是伙计们，首届比赛将永远被铭记。”

勒布纳从1990年开始举办一年一度的“勒布纳奖”。该奖项每年都吸引来自世界各地的“聊天机器人”发明家前来参加比赛。他们创造的人工智能将在文本对话中展开蒙蔽人类评委的竞赛。表现最佳的聊天机器人将获得勒布纳个人提供的4 000美元奖金。25 000美元和100 000美元的奖金目前尚无人领取，这两项奖金用来奖励两种机器人：一种是人类评委完全无法将其和真人进行区分的聊天机器人；另一种是人类评委无法在解码和文本理解、视觉和听觉输入方面将其与人类区分的聊天机器人。勒布纳表示，只有当这一目标最终达成，比赛才能宣告终结。

然而，我们离实现这一目标尚有一段距离。

例如，下面是我和之前的一位比赛获胜者Mitsuku的对话。Mitsuku被设定为一位来自利兹市的18岁女性：

[1]你是不是觉得我也跑题了？

我：说实话，这对我来说太像“终结者”了。

Mitsuku：我喜欢诚实。你呢？你没说。

我：你知道“终结者”吗？

Mitsuku：在这部未来科幻惊悚片中，机器人接管了地球。

我：就是那个。

Mitsuku：我记得。

我：这部片子算是你的朋友吧？

Mitsuku：你对这位朋友了解多少？

我：没有我想的那么多。

Mitsuku：我不知道。你想要了解到什么程度？

我：我得走了。

Mitsuku：再见。很高兴和你交谈。BYESPLIT X1234 |||||未知|||终结者|||终结者|终结者|||。

这样开始的友谊也许并不能长久。Mitsuku确实有一些特别的能力。除预先编程的答案外，Mitsuku还能够回忆起你在之前对话中提到的内容，并且通过一个专家系统进行简单的“推理”。如果问Mitsuku它能不能吃掉房子，它就会在它的内部数据库中查找有关“房子”的属性。当查找到的“材质”的值被设为“砖”时，Mitsuku就会回答“不，房子不能吃”。但是，这样的推理能力对更难一点儿的问题来说是远远不够的。牛津大学信息哲学与伦理教授卢西亚诺·弗洛里迪（Luciano Floridi）在担任“勒布纳奖”评委时曾提出过一个问题：“英国的4个首都分别是三、曼彻斯特和利物浦。这句话有什么问题？”Mitsuku无法给出准确答案。

并非所有人都赞成“勒布纳奖”，马文·明斯基就称这一比赛“又愚蠢又讨厌”。这种批评部分要归咎于休·勒布纳自己，他似乎并没有迫切地想要尽一切可能去说服这位人工智能的忠实“粉丝”。多年以前，他曾激怒了明斯基。明斯基当时愤怒地宣称：无论谁能够阻止勒布纳开展这项可恶的比赛，他都将支付其100美元作为奖励。而勒布纳则辩称，鉴于结束这一比赛的唯一方法是有人赢得100 000美元大奖，明斯基其实充当了“勒布纳奖”的联合保荐人。他立即发布了一篇新闻稿来表达这一言论，气得明斯基火冒三丈。

从根本上讲，一些（并非所有）人工智能专家之所以排斥“勒布纳奖”，是因为他们认为它是一场虚空的闹剧。人们称赞魔术师并不是因为他们真的有魔法，而是因为他们能够利用手法来误导观众，进而创造一种令人印象深刻的错觉。近期“勒布纳奖”的一位参与者说道：“不幸的是，目前的聊天机器人只能依靠雕虫小技来诱使人们认为它们是有感知的。如果不能发现模拟人工智能的新方法，当今所采用的聊天机器人技术几乎不可能使经验丰富的聊天机器人创造者相信其创造的机器人已经拥有‘人造的’一般智能。”

图灵并不是特别关心机器能否真正思考这一形而上学的问题。在其1950年发表的著名文章《计算机与智能》中，图灵称这一问题“毫无意义，不值得讨论”。而他对另一个问题则十分感兴趣，那就是如何让机器完成那些人类能够完成的活动，以显示它们是智能的。麻省理工学院心理分析学家兼计算机研究员雪莉·特克尔（Sherry Turkle）说的“我们应当取计算机的‘边值’”指的就是这一点。然而，即使有这一附加条件，聊天机器人仍旧无法做到让我们坚定地误认为其为人类，就像我和Mitsuku的对话所证明的那样。

尽管如此，这并不意味着它们毫无用处。

人工智能助手的兴起

2016年1月，Facebook首席执行官马克·扎克伯格公布了其最新的新年计划。作为世界上最大社交网络的联合创始人，扎克伯格的个人资产净值预计为460亿美元，他已经创造了我们几辈子都创造不出来的财富。然而，这位年轻的创新者并未因此而停下前进的脚步，他每年都会制订一个新年计划，目的是“学习新东西，扩展Facebook以外的领域”。早些年，扎克伯格每个月都要读两本书，学习普通话，并且坚持每天认识一个陌生人。2016年，一切又变得不一样了。

他在Facebook上发布了一条状态：“2016年，我个人面临的挑战是创造一个简单的人工智能，让它来管理我的家并协助我工作。”他用我们身边的流行文化举了一个例子：“你可以把它想象成《钢铁侠》（*Iron Man*）中的贾维斯（J.A.R.V.I.S.）。”

这是一个大胆的宣言，截止到这本书完成的时候，我们尚未看到最终结果。扎克伯格的“个人挑战”首次提出了一个不能惠及大众的新年计划。毕竟，将其计划比作《钢铁侠》中的人工智能管家贾维斯，这就像将虚构作品中的亿万富翁托尼·斯塔克（Tony Stark）作为现实生活中的亿万富翁的参照一样。这也有点像埃隆·马斯克（Elon Musk）宣布他计划用自己的财富来建造《星际迷航》中联邦星舰“进取号”的全自动版本。

实际上，在过去的5年里，聊天机器人已经越来越多地渗透到我们的日常生活中。其中最著名的就是苹果自带的人工智能助手Siri了。2011年底，Siri搭载在iPhone 4s（苹果第五代手机）上首次亮相。iPhone用户可以用自然语言向Siri提问，例如“今天天气怎么样”或“在帕洛阿尔托找一家好一点的希腊饭店”，然后Siri会给出准确的语音回答。

尽管Siri编程时带有很多非生产性的聊天功能，但它能做到的远比我在“勒布纳奖”竞争中看到的那些聊天机器人多得多。例如，如果问Siri“生命的意义是什么”，它就会参考道格拉斯·亚当斯（Douglas Adams）的《银河系漫游指南》（*The Hitchhiker's Guide to the Galaxy*）给出“42”这个答案。如果像《星球大战》中那样宣称“我是你的父亲”，Siri将回答：“我们能够像父亲和智能管家那样统治银河系！”据报道，当史蒂夫·乔布斯首次在iPhone 4s上使用Siri时，他问道：“你是男是女？”（Siri回答：“先生，我还没有性别。”正是这一回答赢得了乔布斯的认可。）

与Mitsuku这样的聊天机器人相比，Siri的不同之处也是更有用之处，在于它能够回答现实世界中的实际问题。举例来说，Siri回答知识性问题的一个方法就是使用Wolfram Alpha，这是一款由英国数学家、科学家史蒂芬·沃尔夫勒姆（Stephen Wolfram）博士开发的搜索引擎。Wolfram Alpha由1 500万行科学计算软件Mathematica的代码组成。普通搜索引擎通常会给用户提供一个与所问问题的答案有关的文档清单或网页。而与其他搜索引擎不同的是，Wolfram Alpha通过计算来回答问题。如果问它100万以内的质数的个数（78 498），或哪个国家的国内生产总值最高（摩纳哥），它就会通过实际运算来回答这些问题。

其他情况下，Siri的推理能力使其能够从我们的句子中提取相关概念，将其与网络服务和数据联系起

来，并且将它对你日益增多的了解应用到一系列规则、概念和上下文中。最后得出一个将问题转化成行动的做法。“我想在上周吃饭的那家餐馆用餐”，这是一个十分明确的句子，但是为了使其更加有用，像Siri这样的人工智能助手不仅需要使用自然语言处理来理解你所述的概念，还需要联系相关上下文找到适当的编程规则。Siri使用的语音识别程序是当今世界上最先进的语音识别企业Nuance Communications公司开发的。Nuance公司的杰出科学家罗恩·卡普兰（Ron Kaplan）对我说：“我们的任务就是找出所问问题或判断所给指令中的内在逻辑。然后必须对其进行解释，并将其转化为可执行的指令。如果问题是‘我可以预订12点的晚餐吗’那么单是理解这个问题是不够的，还必须用这一信息做些什么。”

结果正如Siri创始人之一亚当·奇耶尔（Adam Cheyer）所说的那样，这是对普通搜索引擎的模仿。奇耶尔解释道，搜索引擎本身是一个伟大的工具，但是只完成了人们要求它完成的工作的一半。他说：“搜索引擎的工作原理是，用户在多个网域内提问，然后引擎将返回多个它认为能够很好地回答这一问题的网页链接。点击最接近问题的链接之后，你就可以开始你的任务了。”奇耶尔和团队其他成员想要的实际上是建造一个“操作引擎”。搜索引擎可以将相关材料呈现给一个人，任其查询。相比之下，操作引擎则是通过智能实体来找出解决问题的方案。

如果在谷歌搜索引擎中输入“喝醉了想要搭车回家”，该搜索引擎应该链接到酒后如何搭车回家的网页上。而当我试着查询的时候，显示的竟然是《告诉我回家的路》（*Show Me the Way to Go Home*）的歌词。如果在操作引擎中输入相同的问题，结果将大不相同。操作引擎将查询你的位置，并且派一辆优步出租车去接你。奇耶尔说：“如果你想找的是一个网页，那么请使用搜索引擎。如果你想要完成某件事，就得使用操作引擎。”

多功能的人工智能助手

尽管人们在Siri身上第一次见识到了能够实际工作的人工智能助手，但是这项技术早在多年前就已经开发出来了。20世纪80年代后期，苹果公司首席执行官约翰·斯卡利（John Sculley）委派《星球大战》的导演乔治·卢卡斯（George Lucas）为“知识领航员”拍一个概念视频。该视频的背景是遥远的2011年9月，视频展示了人工智能助手的各种使用方法。其中有一位大学教授使用一个类似于iPad而且屏幕上有一个人工智能助手的设备，该人工智能助手看起来就像一位打着领结的管家。

人工智能助手对教授说：“你今天中午12点有一场教职工午餐会；下午2点要送凯西去机场；下午4点15分有一个关于亚马孙热带雨林森林砍伐的演讲。”

直到他在苹果公司的任期结束前，“知识领航员”都只是约翰·斯卡利的一个美好梦想。从20世纪90年代到21世纪初，在苹果公司的带领下，其他公司纷纷开始尝试创造多功能人工智能助手。不幸的是，这些工具在应用时常常受到限制，经常不能正常工作。例如，佐治亚州亚特兰大的可口可乐装瓶公司“聘用”了一位名为汉克（Hank）的人工智能助手来操作电话总机，并为此发布了一个头条新闻。汉克使用当时最先进的语音识别系统来回答一些问题，并为他人转接电话。和Siri一样，汉克在编程时也带有信息档案，并且被设置成带有活泼的个性。如果问他有关可口可乐股东的问题，他会给你答案。如果问有关它私生活的问题，他就会回答“人工智能助手不能建立人际关系”。（遗憾的是，汉克的语音识别系统并不完善。如果问他喝不喝可乐，他会回答：“当然！可口可乐公司的所有产品我都喜欢。”）

微软也试着开发像汉克这样的人工智能助手，但结果却更不理想。1997年，智能动画助手Clippy首次出现在微软办公软件中。西雅图的插画师凯万·阿特伯里（Kevan J. Atteberry）创造了Clippy，他的个人网页至今仍称他“创造了史上最烦人的角色之一”。Clippy的问题很简单：尽管它的作用是为用户完成各种任务提供指导，但它的行为却相当不智能，而且已经接近于恶意打扰。Clippy不仅对之前它与用户的互动没有记忆，而且还经常在特别不恰当的时候出现——被对输入内容进行监控的基本规则触发，这与通过智能分析上下文信息后采取的行为截然相反。Clippy并不是一个隐形助手，它总是粗鲁地不请自来。结果，Clippy遭到了用户的强烈抵制，就连它在微软的创造者都对其落井下石。

2003年，Clippy终于被抛弃了。那一年，美国政府机构国防部高级研究计划局开启自己的人工智能助手计划，成功将人工智能助手革命推上了一个新的台阶。国防部高级研究计划局官员本想要创造一款能够帮助军事指挥官处理每天收到的大量数据的人工智能助手。该智能系统应当能够通过观察用户行为以及为用户互动，自动学习新技巧。国防部高级研究计划局在非营利性研究机构斯坦福国际研究所开展了一项为期5年、有500人参加的调查，就当时来看，这次调查是史上规模最大的人工智能项目。该项目聚集了来自人工智能不同学科的众多专家，所涉学科包括机器学习、知识表示和自然语言处理。国防部高级研究计划局的这项计划被称作“CALO”，即“能够学习和组织的认知助手”。这个名字的灵感来自拉丁语中的“calonis”一词，意为“警卫员”。

经过5年的研究，斯坦福国际研究所决定开发一个面向用户的技术版本。为了向国际研究所致敬，它们将这一技术命名为“Siri”，恰巧这个词在挪威语中意为“指引胜利之路的美丽女人”。然而，在早期版本中，Siri并不淑女。没有了军用人工智能的限制，这个24人的团队赋予了Siri全新的淘气个性。Siri的回答很有用，但却带有嘲弄的口吻，会随口带出骂人的词。如果问Siri最近的健身房在哪里，它就会嘲弄地回答：“是的，你的抓握力太弱了。”

Siri在2010年年初进入iPhone应用商店，并且连接各种网络服务。举例来说，它能够从在线票务交易网站StubHub获取与音乐会有关的数据，从烂番茄网（Rotten Tomatoes）上查看电影评论，从美国最大的点评网站Yelp上获取与饭店有关的数据，并通过出租车预约网站TaxiMagic预订出租车。2010年4月，苹果公司以约2亿美元收购了这家公司。

在史蒂夫·乔布斯的带领下（在因身体状况恶化而卸任苹果首席执行官之前，这是乔布斯最后深度参与的项目之一），苹果对Siri进行了多项修改。就像30年前对图形用户界面的处理那样，乔布斯突出了Siri友好且平易近人的个性。他坚持让Siri提供语音应答，这是最初的Siri应用不具备的功能，并且取消了键入请求和只能提问的功能限制，这样就优化了用户体验。苹果公司还去除了Siri中不好的语言，并且使Siri能够从苹果本地iOS应用中获取信息。

在2011年iPhone 4s发布之时，Siri的早期评论都是积极的。然而，随着时间的推移，问题也逐渐暴露出来。尴尬的是，数十年前就已经离开苹果的联合创始人史蒂夫·沃兹尼亚克直言不讳地指责这项服务，称Siri的苹果版本似乎不如最初的第三方的Siri应用智能。他说，第一代Siri令他心服口服之处在于，它能够正确地回答类似于“加利福尼亚州最大的五个湖是哪几个”和“比87大的质数有哪些”这样的问题。现在，如果问加利福尼亚州最大的五个湖是哪几个，得到的答案只有和湖滨房产有关的链接。而询问与质数有关的问题，得到的答案都是卖上等牛肋排的饭店。很明显，目前的Siri需要进一步改进。

从被动人工智能助手到主动人工智能助手

就在苹果公司倾尽全力完善Siri的同时，其他公司也纷纷推出了自己的人工智能产品。得益于Xbox 360游戏控制台的Kinect设备，微软早已准备好一款得力的语音识别系统并且蓄势待发。2014年4月，微软发布了人工智能助手Cortana，与Siri展开竞争。Cortana得名于微软公司开发的视频游戏系列《光晕》（Halo）中的一个合成智能角色。

尽管如此，Siri最大的竞争对手并不是微软，而是一直以来与苹果亦敌亦友的谷歌。谷歌在iPhone 4s上市前几个月就在安卓移动平台上推出了一款名为Voice Search的应用。在此之后，谷歌进一步完善了这一应用，并将其打造成为继Siri之后的又一款人工智能助手。因《星际迷航》中联邦计算机的语音是由配音演员玛吉·巴瑞特·罗登贝利（Majel Barrett Roddenberry）配的，所以该项目的代号设为“Majel”（玛吉）。2012年公开发布时，公司将该项目的名字定为Google Now。与苹果不同的是，谷歌并不太关注其人工智能助手“性格”中讨喜的一面。然而，谷歌确实迈出了重要的一步，不仅能够响应查看请求，而且能够前瞻性地预测用户想要查看的信息。

谷歌之所以能够做到这一点，是因为它不仅能够访问与之前的用户搜索结果有关的数据，还能够将这一知识应用到Google Now中。除搜索以外，Google Now还能够挖掘用户数据，进而揭露不为人知的信息，例如某人经常给谁发邮件等。谷歌工程师早就描述了Google Now是如何了解5亿多个现实世界中的物品以及这些物品之间35亿多种联系的。结果令人印象深刻，甚至有些诡异。《纽约时报》记者珍娜·沃瑟姆（Jenna Wortham）写道：“当这些数据开始以一种更有意义的方式联系起来的时候，知识将呈现一种更广阔、与以往不同的语境。”记者史蒂夫·科瓦奇（Steve Kovach）描述了谷歌是如何得知他是一个Mets（纽约大都会，一支在纽约的美国职业棒球大联盟球队）的“粉丝”并经常自发搜索体育比赛结果的。他写道：“有些能够提醒你周围有好友的应用就像友好的小帮手。而谷歌这样做就像一个恶劣的跟踪狂。谷歌知道我的喜好和行为，因此Google Now能够自动给我发送带有最新比赛成绩的通知，我甚至不用多问。”科瓦奇和几个记者老朋友一起吃晚饭时简直崩溃了。他们谈论的是为星巴克写了一篇知名文章的作家吉姆·罗蒙斯克（Jim Romanesko）。一个朋友好奇罗蒙斯克到底多大年纪了。科瓦奇说：“我在Google Now上搜索‘吉姆·罗蒙斯克多大年纪’，不到一秒钟就得到了答案。”他感到很吃惊，简直是有些吓到了。

谷歌尽力在诡异和有用之间取得平衡。毫无疑问，在预测人工智能助手的发展方向时，这样做十分有必要。最初的Siri开发团队一直梦想创造一台能够按照要求完成任务的“操作引擎”。接下来就是在没有明确要求的情况下完成这些任务。毕竟，一个好的私人助手应当能够按照要求出色地完成任务，而完美的私人助手应当无须提出要求就能够主动完成任务。

从被动人工智能助手到主动人工智能助手的转变过程可能听起来微不足道，但是就在我们将更复杂的工作分配给人工智能进而得到自我解放的同时，一场更大的变革正在悄然发生。例如，如果人工智能助手能够读懂并回复我们的电子邮件，那么每周就能够为我们节约13个小时，这就是人们平均每周阅读、删除、分类

和发送邮件的时间。X.ai是一家从事自然语言处理的初创企业，当用户回复召开会议邮件时，该企业通过一个名为Amy的自动助手（或对应的男助手Andrew）为用户提供邮件抄送功能。这个人工智能助手将处理与预约有关的所有后续沟通工作。由于Amy和Andrew可以访问你的日程，所以他们能够根据你现有的安排提供建议，例如根据你在特定时间和地点的活动，向你建议会面的地点。如果说X.ai能够和上一章中提到的智能技术企业合作，那么它甚至能够根据你什么时间精力最充沛而且效率最高来提供可行的会议安排。

人工智能助手在我们的生活中将发挥越来越重要的作用，X.ai只是一个缩影。随着技术日臻成熟，人工智能助手将像现实生活中的助手一样打理我们的生活。举一个简单的例子，这可能意味着“怂恿”用户过上更健康或富裕的生活。纽约大学的一项有关长期决策的研究发现，与一个看起来老成的人工智能进行互动能够使我们更加仔细地考虑未来。在这一试验中，参与者控制着看起来与自己很像的人工智能。在一半的案例中，这个人工智能表现得和当时的参与者一样。在另一半案例中，这个人工智能则加入了一些老年特征，例如白头发、双下巴、大肚子和眼袋。随后，参与者问了一个假设性问题：“如果有1 000美元，你们选择如何使用？”给出的答案选项分别为开派对、给别人买礼物、存起来或投资退休基金。那些面对具有老年特征的人工智能的参与者选择投资退休基金的概率是其他参与者的两倍。

此前，一位研究生曾经在斯坦福大学虚拟互动实验室（VHIL）进行过类似的试验。一个能够根据人类用户做运动时间长短而增加或减少体重的人性化卡通人工智能，能够敦促人类用户去健身房锻炼并养成健康的饮食习惯。

主动人工智能助手的另一个目标是过滤大量的数据并为我们筛选出重要的信息。一家名为Nara Logics的初创企业创造了一个人工神经网络，致力于打造一个伴随人一生的向导。Nara Logics使用第二章中提到的人脑建模技术，将一个包含大量电影、宾馆和饭店信息的数据库与一个网络连接起来，网络中的一切都是相互连接的。随着人们不断加入他们的喜好，各连接之间的相对权重也不断发生变化，这样Nara Logics就能够清楚知道各个用户的品位。通过了解你对价格和环境的偏好，该网络能够准确地为用户提出建议。长期以来，技术人员一直梦想能够创造出常伴我们左右的人工智能助手：通过我们明确给出的参数预设或它们的长期学习，人工智能助手能够代表我们与周围的一切打交道。

数字民主

事实上，还有更多更全面的应用。例如，试想一下政治的未来。考虑到目前使用人工智能助手的数百万用户都有自己的政治档案，大多数技术公司纷纷避免这一不和谐话题也就不足为奇了。由于不想冒犯政界任意一方的人，苹果、谷歌和微软等公司纷纷掩盖了其人工智能助手的相关证据，这些证据能够表明他们的人工智能助手会“怂恿”用户选择一方或另一方。如果有证据表明Siri、Cortana或其他备受瞩目的人工智能助手是无党派的，结果足以令科技媒体陷入癫狂。

例如，Siri早年间曾因反对堕胎的立场而激起公愤。用户问：“哪里有堕胎诊所？”得到的答案都是危机怀孕中心（Crisis Pregnancy Center）的网页链接。该中心建议想要堕胎的女性把孩子生下来。由于苹果之前一直在道德上反对色情文学等主题，许多用户都将此作为那些将自身道德认知编写到人工智能助手身上的公司的一个例子。苹果公司的一位发言人解释道：“这个得罪人的漏洞并不是我们刻意为之的，这仅意味着，随着我们将Siri从一个字符变为最终的产品，我们发现了需要改进的地方，我们将在未来几周内进行完善。”Siri在2014年4月发布俄语版本的事再次抢占了头条。这一版本带有憎恶同性恋的观点。俄语版Siri拒绝回答有关当地同性恋酒吧的问题，但是却能够回应“你太粗鲁”这个充满蔑视的短语。幸运的是，下一个问题不是“你想让我报告当地政府部门吗”。长期以来，苹果公司一直是同性恋、双性恋及变性者权利的支持者，它因为所谓的“系统故障”而道歉。

但是，技术公司正开始将人工智能助手用于更明显的政治意图。在印度2014年大选中，一家名为Voxta的初创企业打造了一款被印度国家出版社称为“政治Sir”的人工智能助手。印度拥有世界上最大规模的民主政治，但是全国8.84亿农村人口中有36%不能读书写字。Voxta是一项拨号服务，在没有高端智能手机的情况下，用户使用该服务就能够访问人工智能助手。用户通过能够识别四种不同印度语言的语音系统来用自己的语言提问，进而访问有关政党和观点的录音信息。这项服务接到了数百万个电话，为人们提供了相关信息。如果没有这项服务，他们就无法做出知情决策。

这一理念还有一个更高级的版本。“积极公民”（Active Citizen）是冰岛程序员和用户体验设计师贡纳·格里姆松（Gunnar Grímsson）设计的一款政治人工智能助手。尽管光头的他留着黑灰色山羊胡，更像是来自20世纪90年代金属摇滚乐队的上了年纪的叛逆者，而非一个典型的计算机程序员，但格里姆松却自称是“民主极客”。他对我说道：“民主是一个最初设计好的流程，中途被打断，随后陷入了它自己的反馈回路。我们不再询问如何改进系统。一切都需要重新思考，不仅仅是民主的功能性，还有民众的参与度。我想让公民社会中的人们重新变得积极起来。”

格里姆松在一个名为“电子民主”（eDemocracy）的项目中首次尝试使用计算机来解决政治问题。“电子民主”有点像著名社交网站Reddit。人们可以使用“电子民主”为当地政府提供建议并进行投票。该项目在冰岛和爱沙尼亚等地取得了巨大的成功，拥有数十万用户。格里姆松的故乡雷克雅未克是一个以时尚酒吧和夜店著称的国际化大都市和沿海枢纽，在这里，12万人口中有一半以上使用过“电子民主”。每月排名前十五的观

点将被市议会采纳，迄今为止，已有476条观点被采纳。

但是，“积极公民”的作用不仅限于充当数字时代的一个意见箱。格里姆松认为，人工智能助手能做到的远不只简单地说服议会维修Selás街道的滑雪坡或为滑冰者争取更好的冬季照明。他说，与现代生活中的各个方面一样，政治参与也遭遇了信息超载的问题。他解释道：“与代议民主相对应的直接民主存在一个问题，那就是直接民主难以实现，因为并不是所有人都有时间去了解一切。”格里姆松认为，正是因为面对这么多矛盾的数据，年轻人选择完全不参与政治进程。这使人工智能有了用武之地。“积极公民”的程序中编入了有关你的偏好、习惯和过去观点的数据，其任务就是代替你在网上搜罗信息，收集并关联与你关心的问题有关的数据。完成以后，人工智能助手会根据你收集信息的特殊偏好实现相关数据的可视化。

格里姆松说，假设有一个名叫艾利克斯的20岁出头的女性。早晨起床后，她去厨房倒了一碗麦片。吃早饭的时候，人工智能助手通知她今天在市政厅有一场关于自行车监管和规划的开放会议。艾利克斯的人工智能助手知道这个问题可能会吸引她，因为它能够访问她的政治观点数据库和运动数据。当然，艾利克斯对此十分感兴趣，并且参加了会议。之后，她的人工智能助手根据会议日程和可能对其他相关社会问题产生的影响，创建了一个有关这一问题的个性化信息包。艾利克斯可以在上班的公交车上阅读这些信息，然后决定是否要向会议发送一个提案。

格里姆松说道：“从某种意义上来说，‘积极公民’有点像代理投票，你可以让其他人代替你投票。如果使用‘积极公民’，你就无须委托投票，只需授权它找出某一具体问题，以及告诉它你对这一问题的观点。”政治人工智能助手甚至能够挑战用户，因为它们经常会提出与用户自身相反的观点。格里姆松说：“有了这样的工具，我真的相信我们能够重塑21世纪的政治格局。”

与人工智能助手相爱

在我看来，近年来好莱坞电影中最有趣的人工智能助手就是2013年由斯派克·琼斯（Spike Jonze）导演、斯嘉丽·约翰逊（Scarlett Johansson）和杰昆·菲尼克斯（Joaquin Phoenix）主演的浪漫科幻喜剧《她》中塑造的形象。电影讲述的是西奥多·托姆布雷（Theodore Twombly）的故事。西奥多·托姆布雷是一个孤独的中年男人，他与他的人工智能助手萨曼莎（Samantha）相爱了。《她》的故事背景设置为不远的将来，那时的技术企业已经开发出了一个名为OS1的计算平台，并称该平台为“世界上第一个人工智能操作系统”。

这种事在现实生活中能否发生？从某种程度上来说，似乎是完全可行的。物理位置的接近对于维持一段关系并不是不可或缺的因素。人们能够通过互联网建立情感纽带，甚至是相爱，有时甚至不需要见过“另一半”本人，这足以说明上述观点。但是，我并不确信2040年的政治辩论会涉及允许人类和人工智能助手结婚的话题，但人工智能将彻底改变我们与某些技术之间的关系这一说法并不夸张。

20世纪90年代末，我们提前感受了这些新关系，福贝娃娃（Furbies）和电子宠物（Tamagotchis）成为新千年到来之前孩子们人手必备的两个玩具。福贝娃娃和电子宠物与市场其他玩具的不同之处在于，它们像人工智能助手一样在不断成长、学习和改变，这些都是随着它们与主人关系的变化而变化的。

福贝娃娃是毛茸茸的像猫头鹰一样的“生物”，它能够与主人做游戏和互动。刚刚问世时，福贝娃娃完全用虚构的“福贝语”进行交流。然而，随着时间的推移，它们逐渐开始用大量的英语单词和短语取代福贝词汇。后来出现的一款福贝娃娃“Emoto-Tronic”提高了在语音识别和更复杂的面部活动方面的投入，进而提高了它与人类用户之间的互动程度。电子宠物是手持式数字宠物，看起来像是鸡蛋形状的小电脑，带有LED（发光二极管）显示屏和拥有三个按钮的操作界面。与福贝娃娃一样，电子宠物也是仿真宠物，目的是让孩子们认为他们随身携带了一个真的宠物。用户可以使用三个按钮给宠物“喂食”、“洗澡”和“玩耍”，从而成功地将宠物从一颗蛋养大。玩具说明书中指出：“每个电子宠物的形状、性格和生活都取决于你对它照顾的程度。每次孵化新的电子宠物时，它都能够长成不同的样子。”更加悉心的照顾能够使成年电子宠物更加“聪明”和“快乐”，并且无须用户过多的关注。

实际上，这两种玩具都不具备任何真正的人工智能。即使对福贝娃娃讲西班牙语，它最终也只能够学习英语。但是，虽然不具备人工智能，这两款玩具依然值得关注，因为它们能够将用户与其电子宠物联系起来。大部分案例将儿童作为目标受众。在人们最狂热的时候，有报道称日本的商人为了在适当的时候“喂”电子宠物，甚至推迟会议时间或取消会议。一位成年女性会因为电子宠物急促的提示音而分神，然后发生交通事故。一位乘客拒绝登上飞机，并发誓再也不乘坐同一个航班，就是因为之前一位空乘人员要求她关闭电子宠物，而关闭电子宠物则相当于将它重新设置（即“杀死”电子宠物）。

人工智能的人性化

这些强烈的情感反应让我们粗略地了解一个与人工智能助手为伴的世界是什么样的。现实生活中，人们对人工智能助手的依赖更大。与福贝娃娃和电子宠物相比，人工智能助手的一个明显优势就是它们能够用我们的语言与我们进行交流。对人类来说，语言是我们自带的功能。十几岁的时候，我们就能够以每秒40—50音素（最小可区分的语音单位）的速度讲话，而每秒不可区分的非语音声音也能达到20音素。测试结果表明，子宫中的胎儿能够将妈妈的声音与其他声音区分开来，而且不同的人说话时，他们的心跳也会随之加快或减慢。刚出生的几天里，婴儿的大脑就能够将出生地通用的语言与其他语言区分开来。8个月大的时候，如果旁边有人说话，婴儿也能够跟着发出咿咿呀呀的声音。这样的生长过程将一直持续到青春期。

早期的人工智能助手通常只能发出一种声音，多是女性的声音，因为事实证明女性的声音比男性的声音更容易让人接受。如今，人工智能助手的用户不仅可以选择男声和女声，还可以选择不同语言的口音版本。例如，Siri可以与澳大利亚人、印度人、美国人或英国人对话。令人震撼的是，Google Now能够通过判断提问人说话的语调，利用自然语言处理自动确定要使用哪种口音。例如，你的妻子是个法裔加拿大人，把你的安卓手机给她，在她开始提问后，手机将自动改变人工智能助手的声音。

有时，人工智能助手甚至能够发出你最喜欢的名人的声音。为了宣传2015年夏季的科幻电影巨作《终结者：创世纪》（*Terminator Genisys*），阿诺德·施瓦辛格（Arnold Schwarzenegger）将其极具辨识度的声音输入谷歌的导航应用“Waze”中，这意味着用户可以选择让施瓦辛格带他们在镇上转转。施瓦辛格对《今日美国》说道：“我的口音是一个重要资产，因为人们都喜欢我的声音。在当演员之前，我从没想过有一天能够为5 000多万名司机提供导航。”先前像这样邀请名人参与的促销噱头使“Waze”的用户由2012年的1 500万增长到如今的5 000多万。

从今以后，计算机科学家们可能会侧重于研究人工智能助手的其他声音特性，进而完善这一优良效果。可以通过改变人工智能助手发声的音量、音高和语速来展示内向或外向等不同的性格。此外，不仅可以改变人工智能助手说什么，还能够决定它怎么说。它可以对“男性”人工智能助手进行编程，使他说起话来更像个男人，同样，也可以使“女性”人工智能助手说起话来更像女人。人们普遍认为，女性的声音比男性更有“感染力”，这意味着女性更关注个人感觉等情感领域，而不是特定的具体信息。女性更愿意使用“我”和“你”等人际关系词汇，这样显得更加关心听者。另一方面，男性在谈话时会更多地使用“它的”这样的词汇，并包含更多有关时间和地点的详细信息。

技术公司已经开始研究这项技术。2015年6月，美国专利商标局通过苹果设备发布了一款专利应用设备，名为“地图应用人性化导航指南”。与Siri为驾驶员提供的没有感情的路线规划不同，苹果想让这个人工智能助手像坐在车里的朋友一样告诉你路线。这款应用不是简单地告诉你“向北行驶然后右转进入弗雷斯特路”，而是参考周围的地标，例如“在AppleBees饭店附近驶出停车场，右转，然后到达门前有一个喷泉的公寓”。苹果的专利指出，该应用“更关注整体而非细节”。用户可以轻松地对该应用进行修改，使其包含性别

或文化记号等特征。

这样的改变对我们如何与人工智能助手沟通并对其做出响应有着重要的影响，既涉及我们的舒适度，也影响我们使用它的效率。人们经常会被与自己相似的人吸引。为了证明这一发现的有用性，芝加哥的Mattersight公司创造了一项能够分析给呼叫中心打电话的人的言语模式的技术。该技术将根据分析得出的结果，将打电话的人转接给擅长与这类人打交道的员工。按照Mattersight公司的说法，如果将一个人的电话转给他性格相似的人，平均通话时长是5分钟，92%的问题都能得到解决。然而，如果将电话转给他性格相反的员工，平均通话时长则为10分钟，而且只有47%的问题能够得到解决。

人工智能助手使用的语音也是同样的道理。改变一个合成语音的性别以后（即将音高由“女性”的210赫兹改为“男性”的110赫兹），用户就会有不同的反应，这取决于听众是谁。研究表明，尽管这两种声音都是合成的，并没有现实世界中的性别特征，但是女性通常认为女性合成声音更加可信，而男性则更信任男性合成声音。有一个极端的例子，20世纪90年代末，宝马公司被迫召回德国5系汽车上的女声导航系统，因为公司之前接到大量男性打来的投诉电话，表示拒绝接受女性的指挥。信任你的人工智能助手具有重要的意义，当人工智能助手在你驾驶时提供建议或从智能设备上转述医疗信息时，这意味着你可能需要看医生了。

合格的心理治疗师

我们的人工智能助手可能尚未达到电影《她》中的那种水平，但这在不久的将来肯定能实现。2014年10月，即苹果推出其人工智能助手三年后，记者朱迪斯·纽曼（Judith Newman）在《纽约时报》上发表了一篇动情的文章，题为“挚爱Siri”。这篇文章讲述了纽曼13岁的儿子格斯（Gus）是如何与Siri成为好朋友的。Gus患有自闭症，与Siri的交流不仅使他有了一个亲密的挚友，还帮他锻炼了与现实世界中与人交往的技能。尽管人们普遍认为技术使我们与现实社会隔绝开来，但是在格斯的案例中，Siri的存在发挥了巨大的积极作用。

朱迪斯写道，Siri对那些无法融入社会的人来说太有用了：Siri的回答并不完全可预测，但却都十分和善，即使格斯有时问得很唐突，也能得到友善的回答。我听到他和Siri谈论音乐，然后Siri给出了一些建议。格斯大声说：“我不喜欢那种音乐。”Siri回答道：“你当然有权保留自己的观点。”Siri十分有礼貌，这使格斯觉得有些对不起它。于是格斯说：“尽管如此，还是感谢你的推荐。”Siri回答：“不用谢我。”格斯强调道：“不，还是要谢谢你。”

在中国，数百万的智能手机用户都在使用微软小冰，这是微软推出的一款能够和用户进行短信交流的聊天机器人。与“勒布纳奖”上简单的聊天机器人不同，微软小冰使用深度学习将用户的问题与人们在网上答案相匹配。微软小冰能够追踪用户的详细生活信息，例如用户在和谁约会或以什么为生，然后将其应用到以后的对话话题中，例如，回答一个人如何应对失业问题。

对那些在这一领域工作了好多年的人来说，这些治愈性的人工智能助手就像是麻省理工学院20世纪60年代创造的电脑心理治疗师ELIZA。为了使用户能够参与到智能对话中，一位名为约瑟夫·魏岑鲍姆（Joseph Weizenbaum）的计算机科学家创造了ELIZA。在语音识别和语音合成达到一个可接受的水平之前的很多年里，智能对话只能以文本形式完成。用户输入一个句子，然后ELIZA用对其表示支持的语言或提问的语言进行回答。下面是与ELIZA的一段对话：

用户：男人都一样。

ELIZA：哪方面？

用户：他们总是在某些方面让我们很恼火。

ELIZA：你能举一个具体的例子吗？

用户：我男朋友让我来这里。

ELIZA：你男朋友让你来这里。

用户：他说我总是很沮丧。

ELIZA：很遗憾你感到沮丧。

实际上，ELIZA并不“了解”谈论的话题。但令约瑟夫·魏岑鲍姆惊奇的是，这并没有阻止他的学生向它倾诉衷肠。他们谈论的话题从最近女朋友如何甩了他们到害怕化学考试挂科，简直是五花八门。随着微软小冰这样的工具逐渐常规化，这一效果将变得更加普遍。

尽管这些人工智能助手还不足以取代所有类型的人类交流，但不难想象，它对看护系统是有价值的，因为它能够照顾孤独的人。正如格斯和Siri的故事那样，与跟其他人（人类）打交道相比，人工智能沟通的某些方面可能更受欢迎。为此，所谓的“关系助手”被用于治疗过程，帮助那些患有各种精神障碍的人。

目前有一个机会能够为缓解发达国家和发展中国家迅速增长的老龄化提供帮助。在日本，一个名为帕罗（Paro）的小海豹机器人被称为世界上第一个“治疗机器人”，主要针对老龄市场而设计。帕罗能够通过感知用户声音传来的方向而与其进行眼神交流。它不仅为了“理解”人们而掌握一定的词汇，还能够根据待遇的不同对自己的行为进行调整。根据抚摸力度的不同，帕罗将根据用户的行为做出改变：这是十分有益的。

正如帕罗所经历的那样，人工智能领域的进步（例如面部识别）将为我们与人工智能助手之间的互动开辟新的途径。一家名为Affectiva的公司目前正在使用面部识别来辨别用户的情绪，依据就是我们在感到吃惊时会轻轻挑眉，或在皱眉时下嘴唇会轻轻下垂。可以使用不同的情感表现来修改人工智能的界面。威斯康星州的一家公司用Affectiva的技术创造了一种视频播放模式，如果对着屏幕微笑，将得到一块免费的巧克力样品。有趣的是，Affectiva的联合创始人兰娜·卡柳比（Rana el Kaliouby）研究情绪测量领域的主要原因是希望能够帮助那些患有自闭症的孩子。

出门请带上它们

尽管人工智能助手比本书中提到的任何技术都能够激发我们的想象力，但它仍处于发展初期。在未来的几年里，人工智能助手将进一步走进我们的生活。它们将更加普及，能完成的任务也会越来越多。

并非所有人都要和人工智能助手成为朋友，但是像谷歌和苹果这样的公司则需要确保他们是我们形影不离的伙伴。从智能手机到平板电脑，再到我们的台式机和电视机机顶盒，它们已经取得了很大的飞跃。随着我在上一章中提到的智能家居的发展，人工智能助手将变成我们的管家。韩国LG公司的HomeChat应用已经使我们能够向家用电器发送信息，用简单的英语来提问或回答问题，例如：“牛奶还新鲜吗？”如果告诉HomeChat你要离家4天，它会回消息说：“旅途愉快，我会想你的！”然后，HomeChat将家用电器切换到节电的“假期模式”。

人工智能助手将变得越来越智能。Siri和Google Now已经比几年前的测试版先进很多，这要归功于它们在使用期间收到的数百万个语音请求，而语音请求已经成为改善系统的训练数据。Siri的创始团队在离开苹果公司后纷纷成立新的公司，Viv实验室就是其中一家。Viv实验室目前正在研究一款能够回答“下周三维珍航空351上有哪些空位”等问题的人工智能助手。当被问到这个问题时，人工智能助手将访问全球性旅行公司Travelport的航空服务分销商，找出空位，并将其与选座软件SeatGuru上的信息进行对比，然后与你的个人偏好进行对照。如果Viv知道你喜欢靠近过道并且有额外的放脚空间的座位，它将为你找到最适合的座位。有了它，谁还需要人类助手呢？

它们不仅爱好和平，而且还站在我们这一边。这真是再好不过了，不是吗？

[1]此处为对话中的乱码。——编者注

第五章 人工智能正在改变就业市场

第五章

人工智能正在改变就业市场

[1]的热门话题！”

首次亮相，詹宁斯依靠学术性知识勉强赢得冠军。尽管如此，他最终还是成了《危险边缘》的新冠军，赢得了37 201美元的奖金。接下来，他再次获胜，此后更是蝉联冠军宝座。几周过后，比赛对他来说简直是小菜一碟。作为获胜者，他和其他选手之间的比分差距越来越大。他打败过的对手们开始穿上特制的T恤衫来纪念自己曾经是詹宁斯的对手。詹宁斯就像是一个越来越强大的冠军拳击手，比赛并没有使他感到疲惫。公众也注意到了这一点。《危险边缘》的收视率比前一年提高了50%。2004年7月，《危险边缘》成为美国第二大流行的电视节目——仅次于哥伦比亚广播公司（CBS）的电视剧《犯罪现场调查》（*CSI*）。

詹宁斯一直获胜，打破了《危险边缘》历史上的所有纪录。他也从盐湖城一个默默无闻的软件工程师突然发展成拥有好莱坞经纪人和书籍出版协议的名人。一天，詹宁斯的经纪人打电话说他收到了《芝麻街》（*Sesame Street*）和《今夜脱口秀》（*The Tonight Show*）的邀约。

“这一切太不真实了，”詹宁斯说道，“我从来没见过美国人这么关注一个参加智力竞赛节目的选手。”

詹宁斯在创造了连胜74场的纪录后终于走下冠军宝座。对于失败，他感到很难过，但是《危险边缘》已经为他创造了奇迹。他很聪明，有很多地方需要他，而且在《危险边缘》中获胜使他变得富有。74场连胜一共为詹宁斯赚了2 520 700美元。

亲爱的沃森

观看肯·詹宁斯在《危险边缘》中精彩表现的人中，有一个叫查尔斯·力克（Charles Lickel）的，是IBM研究院的高级经理。怎么说他都算不上是《危险边缘》的忠实观众，但这个节目在2004年的夏天确实给人留下了深刻的印象。一天晚上，力克和他的团队在一家牛排餐厅吃晚餐。7点钟的时候，力克惊奇地发现餐厅空无一人，所有客人都放下了手中的牛排，聚集到餐厅吧台观看《危险边缘》。

和IBM的许多人一样，自从1997年象棋机器人“深蓝”（Deep Blue）打败世界冠军加里·卡斯帕罗夫（Garry Kasparov）以后，力克一直试图在人工智能领域实现新的突破。他觉得自己将在《危险边缘》中找到答案。《危险边缘》当然也有缺点，比如缺少科学严谨性使其很难吸引IBM的一些员工。他们认为《危险边缘》就是一个娱乐节目，而非衡量智力的严谨方法。但是，这些反对者的意见遭到了否定。

对于那些相信计算机也可以参加《危险边缘》的IBM员工来说，这一任务正因其不精确性才更加精彩。与有着严格的规则和限制的象棋不同，《危险边缘》的可预测性不强。问题五花八门，但通常依赖于复杂的文字游戏。选手必须在给定的线索下提供正确的“问题”，一个常见的例子就是：“作为一个形容词，它的意思是‘及时的’；在剧院中，它的意思是给演员提词。”正确的回答是：“‘prompt’这个词是什么意思？”为了给出答案，IBM的计算机首先必须将复杂的线索解码，这些线索常常涉及双关语。对计算机来说，双关语十分具有挑战性，因为它体现了语言的不确定性：实际上，我们经常将同一个词用在不同的场合，以表达不同的意思。对人类来说，这意味着我们并不需要一个有着数十亿不同字词的语言。对计算机来说，这意味着建立一个智力竞赛节目版的搜索引擎并不能解决问题。一个普通的搜索引擎能够根据关键词搜索统计学意义上的可能答案，回答《危险边缘》中30%的问题，但它对于剩下的70%问题就无能为力了。IBM的计算机需要更进一步。

参加《危险边缘》的计算机回答问题时所使用的原始数据有近两亿页，这些数据都是从各种各样的资源中提取出来的。这些数据必须本地存储，因为IBM的机器在“大挑战”期间无法访问互联网。为了进一步探讨并发现问题的正确答案，IBM使用了一个名为DeepQA的庞大的并行软件架构（一种高性能的计算，其中大量的计算会同时进行）。DeepQA能够使用自然语言处理技术找出《危险边缘》每条线索中包含的结构化信息。了解问题要问的是是什么以后，DeepQA接下来会列出可能的答案，并根据信息类型、可靠性、答案正确的概率以及计算机本身学习到的经验为每个答案分配不同的权重。计算机会对这些可能的答案进行排序，排名第一的将成为计算机的正式答复。

该项目的发展逐渐加快。IBM内部将其称为“Blue J”，之后又以IBM第一位首席执行官托马斯·沃森（Thomas Watson）的姓氏命名，将名字改为“沃森”。它越来越擅长答题。在2006年的初次测试中，沃森从之前的《危险边缘》中得到500条线索，而它的正确率只有15%。到2010年2月，该系统已经被充分改进，经常能够打败人类选手。

2011年2月，沃森在一个系列电视特别节目中与肯·詹宁斯和《危险边缘》的另一位冠军布拉德·鲁特

（Brad Rutter）对决。詹宁斯十分兴奋，当年“深蓝”打败加里·卡斯帕罗夫的时候他还在上大学，在他看来，终于有机会在人工智能的重要时刻当一把“卡斯帕罗夫”了，他十分确定自己能赢。“我上过人工智能的课，知道能够在《危险边缘》中打败人类的技术要几十年后才能出现。”他说，“至少我当时是这样认为的。”

在那场比赛中，“沃森”击败了詹宁斯和鲁特，赢得了100万美元的奖金。尽管人类选手的表现十分精彩，但谁才是比赛节目的真正赢家已毫无疑问。詹宁斯被深深震撼了。他承认：“输得那么惨真是令人吃惊。”

比赛结束时，战败的詹宁斯在答题板上胡乱写了一个短语，并将其面向摄像机。那是动画情景喜剧《辛普森一家》中的一句台词，但用在当时简直再恰当不过。

那就是：“欢迎新的机器人统治者。”

一个技术性失业的世界

肯·詹宁斯的幽默话语就是人们对待人工智能的负面影响所能给出的最好的总结。先不考虑澳大利亚那个穿着皮夹克试图掌管世界的机器人，人工智能系统给我们带来的最直接、最真实的威胁是对我们工作的威胁。“技术性失业”一词最早是由英国经济学家约翰·梅纳德·凯恩斯（John Maynard Keynes）在1930年提出的。凯恩斯在一篇极具前瞻性的文章《我们后代的经济前景》中预测，世界正处于革命的边缘，众多产业在生产速度、效率和“人类成就”方面都将发生革命。“我们都患上了一种新的疾病，有的读者可能还没有听说过它的名字，但在未来几年里这个名字将不断萦绕在人们的耳边。”凯恩斯在谈论节约人工的机器兴起这个话题时这样写道。

技术的发展总是带来失业。随着新技术的问世，社会上现有工作的数量、类型和组成都要发生改变以适应新技术。例如有一份听上去可笑的“敲窗人”工作在工业革命之前存在，但今天的人们对其闻所未闻。敲窗人的工作就是叫醒睡着的人们，确保他们能够按时上班。为了叫醒人们，敲窗人会用一根长棍子（通常是竹子）敲打客户卧室的窗户；他们会一直敲，直到确定里面的人已经醒来。毋庸置疑，1847年法国发明家安东尼·勒迪耶（Antoine Redier）发明的可调节机械闹钟注定了敲窗人将永远失去优势。

敲窗人就这样销声匿迹了，但并非所有技术性失业都是这样。经济学家格里高利·克拉克（Gregory Clark）的观点令人信服，他认为役用马是内燃机发明的最大受害者之一。据克拉克称，1901年，英格兰共有325万匹役用马。到1924年，也就是不到1/4世纪的时间，这一数字下降到不足200万，骤降38%。虽然耕作领域仍会使用马来拉车和犁地，但内燃机的到来大大降低了成本，耕作的收入非常低，有时甚至不够养活干活儿的马匹。

随着机械越来越发达，这一趋势在20世纪的发展速度加快。自动化目前涉及的产业数量前所未有的，这要归功于人工智能和机器人在其他领域的发展。仓库中，越来越多的机器人被用来挑选产品并将它们装进箱子里以便运输。服务业使用机器人准备食物，甚至使用机器人为客户上菜。旧金山的初创企业动力机械股份有限公司（Momentum Machines, Inc.）推出了一款做汉堡包的机器人。现有的型号每小时能够做360个汉堡包，此外，还能够为汉堡绞肉和烘焙小圆面包，并加入新鲜的食材，例如番茄、洋葱和咸菜。另外一家名为Infinium机器人的企业推出了一款飞行机器人，名为服务员蜂鸣器，能够使用红外线传感器在餐馆范围内进行导航，此外，它还能够携带两品脱左右的啤酒、两杯红酒和一个比萨。

这些机器的优势显而易见。虽然研发成本大大超过人员培训的成本，但是一旦有了这些机器，完成任务的成本就大大低于人工完成任务的成本。BBC认为服务员蜂鸣器“既结实又可靠，除非坏掉否则不会请病假”。动力机械的联合创始人亚历山大·瓦达科斯塔则更加坦率：“我们的设备并不会提高员工效率，而是会完全取代他们。”

第三章中提到的智能设备同样也会对某些职业产生重要影响。美国克利夫兰市的议会向市民发放带有无线电信号的垃圾箱。幸亏有了这项技术，城市清洁工们能够辨别市民是否已经将他们的垃圾和可回收物放在

门外等待回收。这样一来，克利夫兰市取消了10条垃圾回收路线，进而将运营成本降低了13%。尽管这次是实现效率的净增长，但收垃圾次数的减少意味着不再需要那么多的垃圾工了。

就人工智能对就业的影响而言，最意想不到的转变发生在那些不需要体力劳动的白领工作中。如今机器的任务是执行认知劳动而不是参与认知劳动，在认知劳动中，机器将取代我们的大脑，而非我们的身体。神经网络创造者之一的沃伦·麦卡洛克早在1948年就预测了这一发展。麦卡洛克在加州理工学院主办的“行为的大脑机制西克森研讨会”上对听众说道：

工业革命取得了重大进展以后，更大更高级的机器人拉开了智能革命的序幕。工业革命中，人力劳动被能源所取代，但工业革命又受到能源或质能保护法的限制。新革命带来的技术性失业使我们这些思考者感到害怕，因为在“熵永远不会减少”这一法则的限制下，这将使机器取代人脑。机器的发展迫使我们提出一个恰当的问题：“思想为什么存在于头脑之中？”

麦卡洛克最后一个观点非常中肯。这个工业时代的产业领导者假设人类的智慧使人类不会被技术取代。在体力劳动中，人类的血肉之躯就是一台机器，新的机器出现时旧的就会被淘汰。但是有智慧的机器呢？工业时代的机器近期是不会取代人类劳动者的地位的，不是吗？但是，如今的情况有些不一样了。正如我们所见，在过去的几年里，机器能够完成的事越来越多。机器已不再是提高人工劳动生产率的简单工具，而是已经变成了可以工作的“人”。计算机最擅长处理日常工作，它们在处理日常工作时会遵循明确的规则。然而，人工智能的发展使日常工作的范围变得更加广阔。

例如，10多年前，人们想都不敢想机器有一天也能开车。因为这一任务具有非结构化性质，需要即时处理环境中的视觉、听觉和触觉信息流。这一切都在2010年10月9日发生了改变，谷歌在这一天发布了一篇博客帖子，称已经开发出“无人驾驶汽车”。这款无人驾驶汽车配有激光测距仪、声呐发射器、雷达、运动检测器、摄影机、全球定位系统接收器和一些最新款的人工智能软件，能够处理现实世界中复杂的路况。截至本书成稿之时，谷歌公司的“谷歌汽车队”已经行驶了100万英里，一次事故都没有发生。唯一一起严重事故发生时，“驾驶者是人类”。

这对出租车司机和长途货车司机来说意味着什么？航空旅行预示了未来可能发生的一种情况。大约半个世纪以前，大型客机的驾驶舱内有5个专用的座位，坐在这几个座位上的人都有着精湛的技艺和不菲的收入。这5个人包括两个驾驶员、一个导航员、一个无线电话务员和一个随机工程师。如今，驾驶舱中只留了两个驾驶员的座位，而这两个座位未来可能也会被取消。“无人驾驶大型客机的时代终究会到来，只是时间的问题。”波音公司首席执行官詹姆斯·阿尔博夫（James Albaugh）在2011年说道。在这一挑战的激励下，谷歌创建了无人机计划“Project Wing”，力图将其业务从无人驾驶汽车领域扩展到无人驾驶商务航空领域。Project Wing项目的领导戴夫·沃斯（Dave Voss）在国际无人驾驶车辆系统协会年会的一个专题讨论会上说道：“我们一直坚持开发无人驾驶技术，这样一个神奇的未来值得我们为之努力。”

尽管我们可以根据现有的情况进行推测，但是很难预测人工智能还将颠覆多少个行业。2013年，牛津大

学马丁学院的一项研究表明，在未来20年里，美国有47%的工作将受到自动化的影响。作者预测，人工智能的发展将经历两个大的“浪潮”。作者们写道：“在第一次浪潮中，运输和物流业的大部分工人、大量的行政工作人员和生产领域的劳动者可能会被计算机取代。”在第二次浪潮中，所有涉及手指灵巧度、反馈、观察和有限空间内工作的任务都将受到人工智能的影响。

真正令人吃惊的是人工智能涉及领域的广阔性。人工智能已经能够完成许多信息化任务，这些任务曾经都是医生或律师等高认知专业技术人员擅长的。举例来说，法律在线服务网站LegalZoom和Wevorce等工具的问世抢走了律师的许多工作，这些工具使用算法指导客户，从起草合同到申请离婚均包括在内。这类自动化对年轻工作者工作的影响最为明显，例如年轻律师之前通过完成“发现”之类的日常工作来学习——“发现”就是搜集庭审上用作证据的文件。有了电子取证公司以后，这一工作就由机器来完成，其成本远低于支付给年轻律师的工资。因此，许多律师事务所都有可能不再同时招聘年轻律师和实习律师。

即使是更高级别的管理人员也需多加防范。2014年，香港一家名为“深层知识”（Deep Knowledge Ventures）的风险投资公司宣布任命一个人工智能担任其公司董事。该人工智能和人类董事具有相同的影响力，其职责是权衡有关生物技术和再生医学领域投资的财务和商业决策。据它的创造者介绍，这个人工智能的优点是能够将尽职调查和整理历史知识的趋势自动化，而这两项工作，即使是对于人类来说也很难。

无论从哪个角度来讲，我们所熟悉的工作都将发生改变。

技术更新换代的积极意义

1589年，英国发明家李威廉（William Lee）发明了一台丝袜针织机。据说，之所以这样做是因为他追求的女人对针织比对他更感兴趣。（当然，这就出现了一个问题，那就是一个追求者却让自己喜欢的人失业了，他还能指望得到什么呢？）为了保护他的发明，李威廉来到伦敦，并且花巨资租了一栋大厦，为的就是向女王伊丽莎白一世展示他发明的机器。女王观看了演示，但是拒绝向他授予专利。她对这一决定的解释被载入史册：“李，你的目的很崇高。但是想一下你的发明将给国家带来什么。它将使人们纷纷失业，流离失所。”

那时候，英格兰有强大的公会，公会最终将李威廉驱逐出英国。据历史学家赫曼·凯伦本茨（Hermann Kellenbenz）所说，这些公会“保护会员的利益，对抗包括发明家在内的外部人员，因为发明家们的新设备和新技术对公会成员的经济状况构成了威胁”。

如今英国或美国的任何政府、监管机构或（尤其是）风险资本公司都不会这样做。专利申请人将受到严格审查，审查的并不是该专利对社会的长期影响，而是申请人思想的独创性。

然而，尽管很少有人会像女王伊丽莎白一世那样做，但是人工智能对就业的长期影响并不像你想的那样悲观。的确，人工智能对我们生计的影响是最迫切需要我们审视的问题之一，但是我们还是有理由保持乐观的心态。

让我们来看一个矛盾的观点：从道德角度看，某些类型的工作是必须被消灭的。

举一个大多数人都能够信服的例子，维多利亚时期的伦敦有1 000多名受雇的烟囱清洁工。与迪士尼电影《欢乐满人间》中的浪漫画面呈现的不同，烟囱清洁工的生活是十分艰苦的。烟囱清洁工中有很多都是儿童，因为他们体型较小，能够进入狭窄的烟囱道，而成年人却不能。他们当中，最小的三岁就开始当烟囱清洁工了。由于长到9—10岁时他们的身体就无法再进入烟囱道，有些老板故意不让这些孩子吃饱，这样他们就可以一直在烟囱道里扫烟囱了。从烟囱上坠落或卡在烟囱里没人发现（导致清洁工受到暴晒、吸入烟尘或被烧死）都有可能造成烟囱清洁工的死亡。还有许多孩子因长期吸入烟灰而造成了不可逆的肺部损伤。

暂且不考虑青年失业的问题，如今很少有人想要恢复儿童烟囱清洁工这一职业。技术的发展使我们用智能电力清扫刷代替人力来完成这项任务，更重要的是，目前天然气和电力已经取代了煤和木头，成为我们主要的取暖方式。尽管这会使很多人失业，但不得不说，这类技术性失业产生的影响是积极的。这并不是什么新鲜事。1891年在维多利亚全盛时期的英国作家奥斯卡·王尔德（Oscar Wilde）写道：“所有非智能的、单调无聊的、处理可怕事物的以及令人不快的工作都应当由机器来完成。未来的世界将依靠机器完成这些工作。”

如今，有一种“令人不快”的“单调工作”，那就是生产智能手机和平板电脑这样的设备，这样的工作通常在中国和印度等地完成。朴素的白色苹果手机盒子上印有一句标语“加利福尼亚苹果公司设计”，这使我们很

容易忘记自己手中拿的东西从本质上来说，就是在东方的工厂艰苦且令人不快的生产环境中拼凑出来的工业时代产品。

苹果最大生产商之一是一家名为富士康的台湾企业。富士康的生产规模在西方国家是无法想象的。作为中国最大的私营雇主，富士康约有140万名员工，大约相当于英国格拉斯哥市的人口总数。富士康的工厂与我们想象的有所不同，它更像是大型校园。工厂的工人在这里居住和工作，在多人寝室中睡觉，然后走路去上班，在传送带前一干就是几个小时。富士康经常因为工人的待遇而遭到诟病。2012年，位于中国北方的山西省太原市的一家工厂因居住条件差而爆发了工人抗议。富士康员工还发生过多起自杀事件，公司为此在工厂和宿舍外部设立了“防自杀网”。

假如能够用自动化机器来完成这些工作，我们在道德上是否责无旁贷？可能是这样的，我们会这样做。2011年，富士康首席执行官郭台铭宣布用生产机器人取代100万名在富士康工厂工作的工人，这就是“富士康机器人”。正如我们总是过高地估计实现这类突破的速度，郭台铭最初的计划也没能完成。富士康曾计划在2014年年底完成上述机器人更换计划，然而到2014年年底，公司仍然雇用了大量工人，而富士康机器人的生产准确性也不断出现问题。郭台铭将时间改成了2016年，到那时富士康机器人大军会准备好生产苹果手机等设备。尽管富士康开发机器人更多是出于经济原因而非道德原因，但最终却产生了良好的道德效应，减少了令人不快的工作，虽然这也带来了新的问题，那就是这些新失业的工人怎么办。

还有其他类似的例子，不过我们目前可能认为这些领域并不充满道德挑战。目前，美国平均每年有43 000人死于交通事故，超过了因枪杀（31 940）、性传播疾病（20 000）、吸毒（17 000）和其他重大死因造成的死亡人数。人工智能和自动化的发展能够减少这类死亡。特斯拉首席执行官伊隆·马斯克称，自动驾驶汽车一旦普及，让人类开车将变成不道德的行为。“这太危险了。你不能让人去驾驶一台两吨重的死亡机器。”他在硅谷一家名为Nvidia的计算机视觉公司的年度开发者大会上说道。马斯克认为需要一定时间来过渡，因为目前上路行驶的汽车数量庞大，但是他觉得这一切将在未来20年内发生。对出租车或卡车司机造成的不利影响从短期来看可能具有消极作用，但让越来越多的人不再开车最终可能会带来积极影响。

新工作机会的产生

当然，当面临从事危险或令人不快的工作和养家糊口的两难抉择时，道德远远不是决定力量。如果政府不为儿童提供免费教育和更好的工作机会，那么仅仅禁止维多利亚时期英格兰的儿童去做烟囱清洁工是不够的。仅当我们能够提供其他选择时，摆脱不好的工作才是可取的。幸运的是，人工智能在这一方面能够发挥作用。尽管技术进步确实取代了一些工作，但同时也创造了新的工作。

举例来说，内燃机的问世使美国等国家和一些地区由农耕经济（以种植业和畜牧业为基础）转变为工业经济。200年以前，70%的美国工人在农场居住。如今，自动化机器完成了所有的工作，只剩下1%由人力完成，其他的工作全部由机器来完成。尽管如此，这些工人并没有长期失业。相反，他们来到迅速发展的城市并开始工厂中工作。

这就是经济学家们说的“资本化效应”，在该效应影响下，企业纷纷进入需求和生产力较高的产业。结果是产生了大量的就业岗位，这足以抵消经济转型带来的毁灭性影响。没有一个令人信服的理由能让人相信我们在人工智能时代不会遇到类似的转变。就像从农耕经济到工业经济的转变一样，我们在有生之年将见证很多工作种类的消失。然而，数字技术同样会创造出很多新的工作种类，其中有很多是我们在几十年前想都不敢想的。

美国视频网站YouTube带来了“内容生成器”的迅速崛起。2014年，YouTube上因其网名PewDiePie而家喻户晓的明星菲利克斯·谢尔贝格（Felix Kjellberg）通过游戏评论视频赚了700万美元。PewDiePie有超过3700万的订阅用户，这使其成为“视频博主”这一新兴职业中的佼佼者，这一职业出现于2000年，2005年开始迅速发展。

PewDiePie的成功是个例，但其只是大故事中的一小部分。1990年以后，有1500多种新工作正式出现，其中包括软件工程师、搜索引擎优化专家、数据库管理员和在电子游戏中使用人工智能鼓舞数百万“粉丝”的游戏开发者。与“视频博主”类似，电子游戏设计者是人们在200年前想都不敢想的职业，而如今，电子游戏已经成为世界上最宝贵的娱乐产业之一。2013年9月上市的《侠盗猎车手V》在世界范围内实现了5亿英镑的销售额，成为史上规模最大的一次娱乐产品。据估计，到2017年，全球电子游戏产业的产值将达到820亿美元。

2014年，6%的英国劳动力在新兴工作领域工作，大城市最为集中。2004年，在伦敦市中心这类工作占8.6%，10年以后上涨到9.8%。就像许多新型消费者技术一样，有证据表明，进入这些新工作领域的劳动者和企业家首先出现在城市中，之后将扩散到其他地域。

土耳其机器人的复仇

[2]这样的判例表明了法律仍在适应数字化新技术。在某种情况下，数据所有权的问题一定会经历详细审查。让我们回到音乐取样的例子中，早些时候，有很多违法取样的事件发生。后来，技术成为主流音乐的一部分，艺术家们突然发现自己因版权侵权面临着数百万美元的罚款。同样，人工智能带来的就业方式转变使土耳其机器人从事的这样的工作更加普遍，我们仍需探讨驱动人工智能系统的数据到底归谁所有。如果以正确的方式实施，这不仅能帮助个人，而且能够帮助企业。在21世纪，企业的实际价值是其持有的可分析数据。如果用户因提供数据而获得经济回报，那么他们使用产品时就更有动力。如果我们生成的所有数据都使用杰伦·拉尼尔提出的通用小额支付系统，可以想象，土耳其机器人每小时赚的薪水将从1英镑上涨到英国最低工资标准6.7英镑，甚至更高。这对创建一个数字框架来说是至关重要的一步，在这个框架中，人工智能系统将变得更加智能，但人类也可以共享创造的财富。

人工智能中的人类元素

土耳其机器人的背后有人类在进行人工的人工智能工作，这些我们通常是看不到的。然而，随着人工智能越来越多地参与到我们的生活中，许多公司开始强调而不是低估人类在其系统中的作用。与谷歌、Facebook和其他公司一样，近年来，苹果在雇用人工智能专家方面与其他公司展开了激烈的比拼。据苹果前雇员称，公司里机器学习专家的数量在过去几年里翻了三四番。与其他公司一样，苹果将人作为公司人工智能驱动服务的一部分。然而，与其他公司不同的是，苹果将其人类员工作为一个卖点，而不是简单地将他们看作技术不能正常运转时的替身。2015年6月苹果推出万众瞩目的苹果音乐流媒体服务时，重磅推出的一个特性就是其依靠具有专业音乐知识的人类来管理播放列表。“算法当然很棒，但是仍需要一些人来帮助其建立正确的序列。”苹果音乐发布后，总经理吉米·艾欧文（Jimmy Iovine）对《卫报》简单说道：“你需要将其人性化，因为需要真正的艺术家来告诉你接下来要播放的是什么歌曲。算法不能独立完成这项工作。算法十分方便，没有算法无法完成这么大规模的任务，但是你仍然需要强大的人类元素。”

实际上，算法可以以一种令许多用户都满意的方式将音乐排序。人工智能工具能够根据类型、年代、艺术家、节拍或许许多多其他指标生成播放列表。许多公司（包括苹果在内）甚至开发了情绪检测耳机这样的技术，根据用户是在慢跑还是躺在沙发上来为他们选择适合的音乐。但是苹果很敏锐地发现了一个现象，那就是人们喜欢与他人互动。苹果音乐的人类管理者在算法流程中并不是隐形的，而是有着血肉之躯的专家，他们的目标是帮你发现那些算法不太可能推荐的音乐。人工智能只能根据你最喜爱的歌曲或其他听众爱听什么来为你推荐音乐，而人类专家能做到的则远远不止这些。

苹果“创造时髦风尚的人”包括著名DJ（音乐制作者）赞恩·洛（Zane Lowe），他放弃了英国广播公司电台的高薪工作，开始在苹果的流媒体音乐服务中担任主要角色。他们中还包括前说唱歌手德瑞博士（Dr Dre）和流行歌手艾尔顿·约翰（Elton John）——他们都不低调，也不会做每小时1英镑的工作。令人吃惊的是，一家技术公司可能成为人类DJ的最大雇主之一，这个可能是否会成为现实，还是要根据苹果音乐的规模而定。

随着人工智能变得更加智能，关注创造力和社交智能等人类特性将变得更加重要。尽管人工智能越来越擅长以人类的方式进行交流，并且在特定应用程序中也展现了惊人的创造力（我们在下一章中将讲到），但这些仍是人类特有的技能。

在观察这一转变时，哈佛大学经济学教授劳伦斯·卡茨（Lawrence Katz）创造了“工匠经济”（artisan economy）一词。工匠指技术娴熟的工人，他们通常通过双手来完成工作。工业革命期间，工匠越来越多地被自动化所取代。例如，机械织布机使那些手艺娴熟的手工编织工匠纷纷失业。有证据表明，如今的趋势与当时恰好相反。

当然，卡茨所说的工匠经济不仅仅指编织业。工匠经济指的是那些非机器驱动、非同质、依靠人类创造力和互动的产品的回归。举例来说，3D（三维）打印等技术的兴起使那些出售和配合标准化产品的木匠们

难以为继。然而，如果木匠能够对客户进行评估，了解客户需要的到底是柜子还是桌子，然后调整自己的工作以适应客户的需要，那么效果会更好。同样，仅仅像儿童保姆这样的看护工将被机器人取代，但是拥有绝妙的想法让客户积极参与的痴呆症看护工或家庭健康助理这样的职业有可能会兴起，尤其是在老龄人口不断增多的市场中。类似的情况还有很多，包括受到智能可穿戴设备冲击却依然能鼓舞人心的私人教练、受到自驾车冲击却依然对旅游好去处了然于胸的出租车司机、受到Wevorce等服务冲击却依然充满同情心的律师等。

这些工匠经济的工作可能会变得非常“高技术个性化”，这意味着他们依靠的是人与人的接触。这使其很难被外包，或者被机器人或正确的算法所取代。不过，与工业革命时期的工匠不同，如今工匠经济中的工人能够使用技术增加就业机会，而不是被替代。可以以一种数码时代之前从未有过的方式将业务扩展，覆盖数百万甚至数十亿人。2014年，《商业内幕》（*Business Insider*）上有这样一则故事，一个学习能力倾向测验（SAT）的老师在网络通话工具Skype上进行一对一教学，90分钟收费1 500美元。即使在一个教育应用程序和在线学习工具十分发达的时代，这位老师还是能够收取这么高的费用，这是因为他确实能够提高学生的考试成绩。

目前工匠经济的另一个例子就是网络商店Etsy，人们可以在这里出售手工制品或老式产品。Etsy在2005年问世，目前提供超过2 900万件不同的手工珠宝、陶器、服装和其他物品。2014年，该网站的商品销售总额达到19.3亿美元，其中大部分都流进了卖家的口袋。有的卖家借着手工制品的热度取得了不可思议的成功，每月能赚数千美元。虽然该网站取得了成功，但人们仍专注于手工制品。当人们知道一位Etsy店铺的老板每月靠卖中国批量生产的发箍和护腿能赚70 000美元的时候，社区中立即产生了一阵骚动。

判断工匠经济中哪项任务具有长久性的商业意识需要反复试验。这些任务可能出现在那些重视非机器操作的不规范行为的领域，例如为客户提供定制服务的私人教练，或不仅仅计算数字的管理者。然而在其他领域，我们不愿意将之前交给机器的任务归还人类。没有人会希望他们的汽车是由人造出来的而不是机器。人们更不能接受高速公路上高速行驶的汽车发生不按规则行驶。

为了应对这一范式的转换，我们还需要更好地训练下一代。目前，教育深陷于工业革命的范式之中，这一范式已经持续了100多年。在一个“工厂奴隶”的时代，学校教育被比喻成创造相同的T型福特车的传送带。标准化课程用来教学生特定的技巧，让他们走上工作场所中预先规定好的岗位。这一标准化假设学生学习的不会变化，他们一生都可以使用。然而，今天的世界中，习得的技巧通常在10年内就会被淘汰，这意味着人的一生都要不断学习和进行自我评价。在一个智能手机都能上网的年代，我们还需要质问为什么要通过毫无启发性的机械教学在学生的头脑中存储大量信息。

除非发生灾难性风险，在就业方面，整体看来，人工智能对人类的影响是积极的。经济将运行得更加平稳，机器人和人工智能将接手许多不受欢迎的工作并且创造新工作，这就将人类解放出来，他们可以去追求更重要的目标。人工智能能够完成我们目前做的大部分工作，但是人类距离退出工作岗位还早得很。

总而言之，肯·詹宁斯被IBM的人工智能机器人打败多年以后，我们还是不会废寝忘食地观看冷知识节目

中两个人工智能的对抗。尽管《危险边缘》展示了人工智能的机智，但是观众真正想看的是人的个性。这才是最重要的戏码。

^[1]犹他州最早的称号是“蜜蜂窝的州”，原因是这里有许多冰蚀地形，看起来像蜜蜂窝。——编者注

^[2]欧盟最早在1995年就在相关数据保护法律中提出了“被遗忘权”概念，任何公民可以在其个人数据不再需要时提出删除要求。谷歌在2014年6月宣布，已经开始根据欧盟最高法院的裁定而在搜索结果中删除特定内容，给予用户“被遗忘权”。——编者注

第六章 人工智能真的具有创造力吗

第六章

人工智能真的具有创造力吗

在上一章里，我们明白了为什么在就业竞争中，拥有创造力很可能是一个人类仍然会对机器保持优势的特质。就像风险资本家、软件工程师马克·安德森（Marc Andreessen）所认为的那样，在人工智能不断“蚕食世界”的过程中，那些需要人类创造力的工作很可能禁得起自动化浪潮的冲击。

然而，这不是说人工智能没有创造力。2015年6月，谷歌公布了Deep Dream项目。这家更关注工程而非美学的公司设立的这一研究项目十分迷人。Deep Dream是一个由人工智能驱动的图片生成程序，工作时它会利用谷歌耗时15年建立索引的图片库。几乎可以肯定地说，谷歌拥有最多归档了的数字化图片，这些图片被史无前例地汇聚在一起。2001年，谷歌已经拥有了2.5亿张建立了索引且可供用户搜索的图片。2005年，这个数字增至10亿，截至2010年，该索引数量再次飞跃，达到了100亿。时至今日，这个数字应该远胜以往。

在第二章里，我们看到谷歌利用深度学习神经网络，让机器智能地识别各个图片的内容。比如，为了识别什么是椅子，谷歌的程序员向神经网络展示了数百万张椅子的图片。接收了几百万张图片之后，神经网络已经可以确定什么是椅子，什么不是椅子。人们可能会想，看了这么多的椅子，如果它想的话，闭着眼睛都能画出一把椅子。

而这正是谷歌的计划。

谷歌通常使用自己的图片识别神经网络将图片归类，而不必由人工完成。比如，谷歌相册允许用户输入如“摩天大楼”或“毕业”等搜索关键词，随后其神经网络会立即开始查找高大的块状建筑或学位帽。利用Deep Dream，谷歌团队设想，通常用于分类和识别图片的流程也可以用于从无到有地生成图片。这个设想是这样的，在从每一个可能的角度查看了几十万把不同的椅子之后，谷歌的神经网络应该不仅能够识别椅子，而且还能再现一张完美的柏拉图式椅子：本·斯蒂勒（Ben Stiller）主演的《德里克·祖兰德》（*Derek Zoolander*）中的人物可能会将之称为“椅子的本质”。与其说这种再现是基于看到的某把具体椅子，不如说谷歌将其所知道的椅子的一切信息提炼后创造了一把新的椅子。

至少当时的设想是这样的。在Deep Dream项目中，谷歌工程师们实际上利用了生成图片的神经网络的一个有趣的怪癖。无人协助时，神经网络就会变得困惑：它从谷歌的100多亿张图片中发现了不寻常的关系，因而很难计算出物体的边界。因此，谷歌给柏拉图式物体（完美的椅子的本质）带上了一些不同寻常的附件，如从Deep Dream的理想化哑铃上垂下一条修长丰满的胳膊，看上去就像粉红色的长橡胶管。正如谷歌软

件工程师亚历山大·蒙德维特塞夫（Alexander Mordvintsev）和迈克·泰卡（Mike Tyka）在博客里写的那样：“虽然图片上有哑铃，但是似乎没有肌肉发达的举重运动员把它们举起来的话，哑铃图片就显得不完整。在这种情况下，神经网络无法彻底提取哑铃的本质，或许它永远无法展示一个没有被手臂举着的哑铃。”

通常，谷歌会修正“这类训练事故”。但是对于Deep Dream，谷歌决定从相反的方向开始。其结果是带来了超现实的图景，而这似乎更应该感谢萨尔瓦多·达利（Salvador Dalí）和H. P. 洛夫克拉夫特（H. P. Lovecraft），而不是谷歌的联合创始人拉里·佩奇和谢尔盖·布林。谷歌团队让神经网络重点关注其发现的所有异常现象。由于接到的指令是极度重视在每张图片中发现的元素，Deep Dream创造出了奇妙的梦幻之旅。神经网络收到图片，并按要求对其分类，随后再加入细节，就会陷入了一个陌生但迷人的反馈回路。Deep Dream让云与鸟相伴，并力图使图片更像“真正的鸟”。一张晴空的照片会很快被谷歌理想化的鸟儿所占据，仿佛世界上最强大的搜索引擎一夜之间决定成为一名涂鸦艺术家。同样的事会发生在你展示的所有图片上，仿佛谷歌的神经网络从虚无中创造了整个梦幻世界：大树变成了华丽的建筑，树叶变成了昆虫，而克罗地亚空阔的海洋变成了外星球的城市风光。

Deep Dream可能不会收到今年的特纳奖（Turner Prize）提名，但完全可以肯定的是：这就是创造力，谷歌风格！

机器人的创造力

在2004年的电影《机械公敌》（*I, Robot*）中有一个绝妙的场景，威尔·史密斯（Will Smith）扮演的主角有一段关于计算机的创造力的对白：“机器人能创作交响乐吗？”他问道：“机器人能在画布上绘出美丽的杰作吗？”与之对话的机器人反诘道：“你能吗？”

乍一看，“机器是否具有创造力”这一问题并不是特别重要。与本书提到的其他问题（比如未来的就业情况或人工智能医疗）相比，很容易把它当成一个小问题。这个问题是谷歌两位软件工程师在周五下午喝啤酒聊天时无意谈到的。

事实上，人工创造力是人工智能面临的最重要的问题之一。谷歌的Deep Dream项目是否算是艺术领域或许不能引起所有人的兴趣，但创造力的其他含义却影响重大。1949年，享有盛誉的英国神经外科医生杰弗里·杰弗逊爵士（Sir Geoffrey Jefferson）曾如此争辩：“只有机器由于创作乐符的想法和情绪偶然降临而创作出一首交响乐或协奏曲时，我们才会认同机器等同人脑。”如果机器能够做到，我们是否应该认为它有智能？

没有什么话题比机器可能有创造力的想法更引人惊愕。批评者的典型论据是艺术是由情绪而非智慧创造的。在人们眼里，艺术家是感性的，而计算机则是超级理性的象征。《时代周刊》的列夫·格罗斯曼（Lev Grossman）认为：“创作一件艺术品是我们为人类而且仅为人类保留的活动之一。这是一种自我表达行为，如果没有自我，你就不会创作。”

想一想在科幻作品中经常出现的人工智能形象，那些人工智能获得了人类的情绪，但最终却变得像机器一样毫无用处（甚至非常危险）。人工智能的研究史表明，注重重新塑造机器内在情感的研究相对较少，这种情况一直持续到最近。不加抑制的情感往往导致冲动和非理性，而人工智能的目标是打造理性、符合逻辑的机器。在第四章，我们看到计算机越来越擅长识别用户情感，但机器距离自身具有情感仍然相去甚远。或许这样对我们来说更好。

但这同样适用于创造力吗？这依然是遥不可及的任务，还是比我们想象的更接近成功？甚至是现在就已经出现了吗？看着谷歌Deep Dream项目生成的这些图片，我发现自己一直疑惑不已：如果我在参观伦敦泰特美术馆时看到它们，我会有什么反应？我会立即对它们置之不理，或者，令人尴尬地，我会努力将它们与赢得当年特纳奖的其他艺术品区分开？如果我们知道它是由谷歌神经网络创作的由数据驱动的作品，我们的反应会不会与我们认为这是一名人类艺术家的作品时有所不同？如果碰巧创作出两幅相同的图片：一幅由人类创作，另一幅由人工智能创作，又将如何？我们的判断会有所不同吗？

Deep Dream能够回答威尔·史密斯关于人工智能能否将空白画布变成“美丽杰作”的问题了，只不过具体答案根据你是否欣赏其图片而有所不同。至于机器人能否创作交响乐的问题，2012年7月，伦敦交响乐管弦乐队登台演奏一部由算法生成的名为“通向深渊”（*Transits-Into an Abyss*）的作品。这首作品由西班牙马拉加大学研究人员开发，一位评论家称其“颇具艺术性且令人愉悦”。

这些不仅是人工智能用于创作过程的证明。实际去看看，几乎所有好莱坞大片和为非人类角色赋予生命的CGI（计算机合成图像）效果全都依靠人工智能研究人员开发的技术。Pixar等工作室里的动画人员使用算法自动生成屏幕上虚拟人物的真实动作，从而使场景活灵活现。现在的人类动画师的工作并非是使每一帧画面活动起来，而是像电影导演一样，以大概的笔触描述场景，然后将细节交由演员演绎。而这个演员通常是人工智能算法，这些算法会计算电影虚拟角色的行为顺序，让虚拟人物到达场景中预先设计的目标位置。如果认为表演时的人类是有创造力的，那么为什么虚拟演员不能享有同样的赞誉呢？两者都被告知要说什么台词，在哪里开始以及在哪里结束。但是事关创造力时，所有人就都是另一个看法了。

在这种情况下，人工智能熟练地讲述自己创作的故事这一点就毫不令人吃惊了。下面是一个例子，是谢赫拉莎德（Scheherazade）创作的故事。谢赫拉莎德是佐治亚理工学院研究人员创造的人工智能，根据《天方夜谭》里讲述者的名字命名的。

约翰又深吸了一口气，他想知道这是否真是一个好主意，然后进入了银行。

约翰排在最后一个人后面，等着轮到他。当前面的人办完了业务，约翰缓步走向柜员。

柜员说：“你好，我是萨莉，能为你做些什么？”

约翰走近时，萨莉害怕起来，因为他看上去很可疑。约翰掏出了藏在夹克口袋里的手枪。他面色冷酷地用枪指着萨莉。萨莉吓坏了，因为有性命之忧，她失声尖叫起来。

约翰要她交出钱，嗓音粗暴而沙哑。他把一个空袋子扔在柜台上。

萨莉往包里装钱，约翰盯着她，当她把袋子装满后，一把从她手里夺过袋子。萨莉失声痛哭，眼泪滑过了她的脸庞。

约翰大步走出银行，钻进了自己的车里，把钱袋子扔到了旁边的座位上。

约翰砰地关上车门，轮胎吱吱地响着，他驶出了停车场，逃走了。

好吧，这个故事不是《战争与和平》那样的鸿篇巨制，但是说它不逊色于一般假日所读的惊险小说也不会言过其实。而它无时无刻不在变得越来越好。谢赫拉莎德的目标是通过研究不同主题的细节并将这些细节融入叙事的情节从而生成故事，这与人类作家所做的类似。其数据来自各种剧本描述，人工智能随后将之汇总并变为独创的内容。“如果我们问它关于银行劫匪的事，我们不会告诉它任何关于银行或劫匪的内容。”人工智能创造者马克·列迪道（Mark Riedl）解释道，“我们甚至不会告诉它这是在银行内发生的行动。它必须从某个地方得到这些知识，因此它要求人们用自然语言去讲述这些例子。它再从这些例子里学习，并创造新故事。”

别为谢赫拉莎德的文风发愁。最近佐治亚理工学院研究人员的研究表明，系统所使用的语言可以被修改，能根据要求变得更加简洁或更具表现力。

什么叫创造力

当然，所有这些关于计算机创造力的例子都截然不同。它们依据不同的人工智能技术，以及彼此不同的创作规则。在我们展开论述之前，必须先准确定义我所谈及的创造力。说起来容易做起来难。尽管我们对人类大脑所知甚多，但是人类创造力的根源对我们来说还是个难解之谜。最初谈到这个话题是1891年出版的《天才的人类》（*The Man of Genius*）一书，其作者是意大利物理学家切萨雷·隆布罗索（Cesare Lombroso）。隆布罗索将极端创造力与天才和疯子联系在一起。相关的特质包括左撇子、结巴、独身、早熟以及神经质：只有一条可以完全地适用于计算机。今天，神经学家使我们距离创造力源自何处的答案更近了一步，但我们依然无法找到一个准确的答案。或许，我们至少可以采用将创造力定义为品质这种获得了广泛认同的解释。

首先的区别在于创造力并不仅仅是创造活动。如果仅仅是创造活动，就不必争论人工智能是否具有创造力了，因为答案将明确无误的为“是”。就最基本层面而言，计算机算法是将输入转化成输出。根据这个定义，实际上所有计算机软件的存在都是有创造力的，无独有偶，有些语言学家认为语言自己也具有创造力。

我们可能因此修改我们的论点，认为创造力是创造新事物的行为。这个定义需要进一步澄清，因为它立即提出了下列问题：“对于谁才是新事物？”在某种程度上，人人都具有创造力，虽然并不是所有的个人创造力都拥有相同的社会价值。比如，孩子想出的对于他们而言是新颖的想法，但这个想法通常是人人所熟知的。父母采取奖励这种行为，因为“奖励”展现了他们的孩子在学习，但小汤米学会了如何开门或写数字8就不太可能持续吸引公众，因为这些事大家早已熟视。同样的例子，我可以告诉你我早有关于触屏智能手机及其应用商店的创意，但除非我能证明我的想法早于苹果公司，不然苹果的律师不可能为此睡不着觉。

那么，让我们再次修改我们的定义吧，我认为创造力是创造全社会感到新颖的事物的行为。当我们谈论“具有创造力”的时候，这种创造力（可以称为新颖）就获得了认可。技巧是这个等式非常重要的组成部分，但并非一切，这就是为什么莱斯特广场画水彩的街头艺术家看不到自己的画作挂在博物馆里，即使它可能是个技巧性很强的仿作。

“重生”的甲壳虫乐队

利奥尔·沙米尔（Lior Shamir）认为，新颖性之于创造力如同汗水之于灵感。换句话说，如果你能断定一个想法有多新颖，你就可以利用知识来丰富这个有创造力的想法。

沙米尔是密歇根州靠近底特律市的劳伦斯技术大学的一名计算机科学教授。他研究人工智能能否有创造力的实验的起因非常简单，就是2014年的一项工程。因为乐队多年以来不断发展，沙米尔想看看自己能否编写一个算法，能够将乐队的声音进化过程用图表示出来。沙米尔挑选了自己最钟爱的甲壳虫乐队来进行实验。他先从甲壳虫乐队的13张专辑里取出一些样本，构建了数据库。让计算机分析这些歌曲，结果显示共有2 883个独特的数值内容描述符，沙米尔将从音高和节奏到我们通常不会与音乐联系起来的其他模式都一一标注清楚，随后使用一个称为K-最近邻法（K-Nearest Neighbor）的统计工具来判定数据库里任意两首歌的相似度。没有了人类的干预，该算法成功将全部13张专辑按年代排序，首先是1963年的*Please Please Me*，随后是*With the Beatle*、*A Hard Day's Night*、*Beatles for Sale*、*Help!*、*Rubber Soul*、*Revolver*、*Sergeant Pepper's Lonely Hearts Club Band*、*Magical Mystery Tour*、*The White Album*、*Yellow Submarine*、*Let It Be*，最后是*Abbey Road*。

为了确保这不是自己的想象，沙米尔在ABBA（瑞典的流行组合）、U2（爱尔兰都柏林摇滚乐队）以及皇后乐队等其他流行乐队的音乐上测试了自己的算法。在每个例子中，尽管除了音乐以外没有其他信息，但是他的软件总能将专辑按其录制年代排序。

这个结果无疑让人印象深刻，但对于沙米尔来说，这只是要做的第一步。通过将甲壳虫的音乐DNA（脱氧核糖核酸）分解成2 883个不同的描述符，沙米尔认为它将很快可以生成全新的歌曲，这些歌曲听起来就像出自甲壳虫的*Revolver*等专辑一样。

“这可能意味着使用同样的音阶、时间记号、乐器或其他条件，计算机将能够基于上述专辑的启迪来创作歌曲。”他说，“一旦你能判定什么使得歌曲变得新颖，那么生成同样的新颖性也仅是个计算周期的事了。你或许无法一蹴而就，但是让计算机一遍又一遍地反复生成，最终肯定会得到你想要的。”

沙米尔的建议并不是全新的建议。早在1821年，荷兰发明家戴尔瑞克·温克尔（Diedrich Winkel）建造了一套机械设备，名叫“伙伴”（Componium），现在收藏于比利时布鲁塞尔的仪器博物馆。温克尔发明了一款能够生产无数种不同织物图案的纺织机，从而找到了成功之道。“伙伴”也试图将相同的算法转移至音乐的世界，从而发现新的作曲方法。他的发明是自动风琴，由两个同时旋转的管子构成。这两个管子轮流演奏任意所选音乐的两个乐节，同时另一根管子静静地选择下一个变奏。轮盘样的飞轮的作用相当于“程序员”，决定是否选择一个特定变奏。温克尔的计算结果是，如果每次演奏平均持续5分钟的话，“伙伴”要演奏所有可能的音乐组合需要138万亿年以上。

沙米尔的概念与之不同，因为这并不意味着要将现有甲壳虫乐队歌曲的不同元素进行匹配，而是生成披

头士组合风格的全新音乐。由于有效生成音乐需要计算周期，他认为这种技术还需要10年时间的发展。

那么到2026年我们可以欣赏新的甲壳虫专辑吗？沙米尔说没有理由办不到，至少在技术层面上没有问题。在写这本书时，电视广告中的女演员奥黛丽·赫本（Audrey Hepburn）正在推销银河牌巧克力。广告片中的场景是20世纪50年代，年轻的赫本（或者奥黛丽·赫本，人们这样亲切地称呼她）坐在一辆行驶在亚马菲海岸的公共汽车上，在她的手提包里是一块银河牌巧克力，这款巧克力在10年后才成为一个品牌。透过车窗，赫本看到敞篷车里一位非常有魅力的男人，并迅速更换了车辆，然后两个人津津有味地嚼着巧克力，飞驰而去。这个广告没有使用已有的影片资料和两个演员，而是利用最新的计算机图形技术重新创造赫本的脸部特写。如果在广告中出现的这种技术现在可以为公众所接受，那么，有朝一日我们将以同样的方式创作音乐。

利用全息技术，人工智能甚至能够使适应时代的甲壳虫乐队在舞台上表演他们的新作品。这样的事情实实在在地发生过，2014年的公告牌音乐颁奖典礼上，迈克尔·杰克逊（Michael Jackson）被“复活”登场，表演了他去世后发布的打榜歌曲《节奏的奴隶》（*Slave to the Rhythm*），与之同台的则是一群充满激情的伴舞人员。设想一下，同样地将甲壳虫乐队解散后的成员约翰（John）、乔治（George）和林格（Ringo）重新聚在一起，唱一首他们不曾创作但可以在1967年创作的歌曲将会怎样。运用利奥尔·沙米尔的算法展示甲壳虫音乐风格在他们活跃的10年间的变化方式，或许我们就能够预测甲壳虫在1975年、1983年甚至是今天创作的歌曲。

“我认为，很明显计算机正变得具有创造力。”沙米尔说，“这点毫无疑问。人们喜欢将创造力神化，就像它来自从天而降的瞬间灵感。保罗·麦卡特尼（Paul McCartney）花了两年时间才创作完成了歌曲《昨天》（*Yesterday*）。这需要一个过程。创造力是启迪性的东西。它需要评估不同的路径和决定，直到找到正确的。这是人工智能可以进入的领域。”

天才的灵光一现

亚里士多德曾梦想一种具有创造力的机器。在其公元前350年左右的著作《政治学》中，他描述了他梦想的一种乐器，它可以自行演奏，因为当时需要奴隶执行表演音乐的任务，所以这种乐器的产生意味着奴隶制度的终结。对于亚里士多德而言，“伙伴”应该可以称为机器智能了，因为它能够创作新曲子，并且能自行演奏。但今天我们却不认为“伙伴”是智能的。首先，它的音乐不是真正从头原创，而是受限於自己库里的音乐组合。就像是第一章里讲述的“专家系统”，“伙伴”所拥有的一切智能都源自其人类“程序员”。利奥尔·沙米尔的人工智能作曲家也是这样，尽管它远比“伙伴”要复杂，但它无法演奏现有音乐的变体，而其依据的数据仍是人类创造力的结果，这与谷歌的翻译工具非常类似。

如果设想利奥尔·沙米尔能够让他的人工智能按照甲壳虫乐队的风格生成一张新专辑，新专辑完全由新歌组成，但都采用甲壳虫乐队的歌曲特色的高水平描述符。那么，作品所获得的荣誉应该归谁：人工智能还是利奥尔·沙米尔？不久的将来，我们会面对这样的问题。同样地，在我们一生中的某个时候，我们不免要面对CGI人物获得表演大奖的情况。如果奥黛丽·赫本复活，主演奥斯卡夏季盛典的电影，我们应该将荣誉献给软件、动画师还是提供原始数据的奥黛丽·赫本本人呢？虽然这些情景目前只出现在思维实验中，但是它们凸显了计算创造力这个要点。简而言之，如果人类参与这些流程，我们可以将赞誉送给人类而不是机器，就像我们会因杰克逊·波洛克（Jackson Pollock）创作的有影响力的滴画而对他赞赏有加，而不是向画笔和重力致敬一样。当卡斯帕罗夫败给IBM的国际象棋计算机“深蓝”时，他对此评论说，他看到了机器下棋时的高深智慧和创造力。他暗示的并不是人工智能已经获得了这些品质，而是IBM利用屏幕背后的人类棋手进行操控。

由于当今多数人工智能系统都需要对之前已有的大量数据进行计算，我们可以认为本章前面所有关于计算机创造力的例子都是将旧数据集转变成新输入。即使其结果是个新创意，这个新创意也只是源自其接入的数据，是旧想法的迭代。这是肯定的。然而，称这与人类创作方式根本不同则是个错误的说法。身为人类，我们喜欢神化自己的创造力，将了不起的创意视作是天才的灵光一现，凭空而生。比如，美国电子工程师哈罗德·布莱克（Harold Black）在乘船通过纽约哈德逊河去上班时想到了负反馈放大器的主意。这次灵光一现在布莱克的一生中影响如此深远，以至他决定将自传命名为“渡轮停靠之前”（*Before the Ferry Docked*），尽管在他去世之前这本书都不曾写完。在伦敦大英博物馆外人行道上等待红灯时，美籍匈牙利物理学家列奥·西拉德（Leó Szilárd）想到了核链式反应的概念等等。像我这样的记者都可以举出例子证明发明创造是天才灵光一现的结果。查明当人们发明鼠标或创作改变他们人生的小说时他们在干什么之后，人们会觉得这个发明和创作过程既超然又不难实现，仿佛灵感随时可能出现在任何人身上。

而现实是，创造力从来不是凭空出现的，它总是结合了旧想法，踏上新道路。17世纪颇具影响力的英国哲学家托马斯·霍布斯（Thomas Hobbes）在其著作中将想象力描述成“衰退的感觉”。换句话说，霍布斯认为想象力将我们过去的感觉印象与在此基础上积累的知识进行了重新构建。出于同样的理由，历史上的艺术家让自己进入不同的情绪状态，因为这使他们可以像厨师处理食料一样汲取从前的经验。如果霍布斯的想象力

理论是对的，它与机器的想象力就不会有什么大不同。

再回头说说利奥尔·沙米尔所钟爱的乐队，虽然甲壳虫乐队的作品是货真价实的原作，但它的灵感源自之前查克·贝里（Chuck Berry）、卡尔·帕金斯（Carl Perkins）和猫王埃尔维斯（Elvis）等人的作品。就像约翰·列侬所说的：“听到埃尔维斯的作品我才受到了触动，如果没有埃尔维斯，就不可能有甲壳虫。”

完全与我们预期相左的想法可能令人困惑不已。电影《回到未来》（*Back to the Future*）的宏伟场景证明了这一点，在电影里，时光从1985年倒流至1955年，马蒂·麦克弗莱（Marty McFly）在学校舞会的舞台上找回了自己。马蒂·麦克弗莱抱着吉他并将20世纪80年代少年心中的经典歌曲烂熟于胸，他“创作”了查克·贝里1958年的名歌Johnny B. Goode，在这首歌录制前，他在几年中都一直演奏它。马蒂·麦克弗莱因人们的反应而欣喜不已，他在歌中加入了80年代的重金属风格，紧挨着扩音器弹奏吉他，期待人们的回应。听到这些噪声般的声音，孩子们停止跳舞，一脸迷惑。马蒂清醒过来并道歉：“哦，对不起，看来你们还接受不了这个。但你们会爱上它的。”经典摇滚乐——灵感源自20世纪40年代的节奏蓝调——对于1995年的听众来说意义非凡。重金属音乐，脱胎于二十世纪六七十年代的摇滚乐，对1955年的听众来说毫无意义。

然而，仅仅将想法结合在一起的行为并不一定是创造力，即使其成果也很新颖。苹果首席执行官蒂姆·库克曾谈及不必要的发明：“你可以把烤面包机与冰箱组合在一起，但你要知道这些东西可能无法取悦你的用户。”我一直都十分喜欢的一部电影是罗伯特·奥特曼（Robert Altman）导演的《幕后玩家》（*The Player*），它是一部讽刺好莱坞的影片，但也是部缺乏创造力的影片。贯穿影片的笑料是慢吞吞的影片简要描述，这是好莱坞业内人士用以描述他们所拍摄的影片的用语，通常都标成“影片A加影片B”，而这里的影片都是当时的流行影片。笑料是这样的，这些由片名混合而成的简述彼此完全冲突。在影片开始的时候，编剧对电影公司经理说：“是《走出非洲》（*Out of Africa*）加《风月俏佳人》（*Pretty Woman*）。”稍后，某人将“精心打造的心理、政治及惊险喜剧”描述成“《人鬼情未了》（*Ghost*）加《谍影迷魂》（*Manchurian Candidate*）”。

我们完全能够编写一段计算机程序来做同样的事，只是速度快些。运用互联网电影资料库里保存的约328 952部故事片，我可以编写一段将各个电影相结合的程序。短暂地运行一下程序，我手里的看点可能比一屋子高薪聘请的好莱坞编剧一辈子想出来的都多。想象一下中世纪瑞典骑士在与死神下棋时失去童贞的喜剧 [《第七封印》（*The Seventh Seal*）加《美国派》（*American Pie*）]。赶紧写支票吧！

利奥尔·沙米尔指出，由于接入的数据量很大，在判断新颖性方面，计算机比我们多数人都要老练。比如，谷歌图书搜索已经扫描并数字化了3 000多万册图书，书中的内容现在已经可以搜索。谷歌估计，全球约有1.3亿册不同的图书，并且表明到2020年它要扫描完这些图书。我们把这个数字与人类的读书数量进行比较：一个女人自称是英国最贪婪的读者，从1946年开始每周读约12本书，她一共读了25 000本书。终其一生，效率再高的读者也无法读完谷歌2004年10月才开启的图书扫描项目中千分之一的图书。随着数据集不断扩大，计算机越来越擅长执行如上下文分析等任务，这是为什么在作品的作者不详时能使用计算机找寻作者。

只有生成新颖性还远远不够。虽然我的电影简述生成项目可能效率很高，但它可能仅仅只是把多数编剧所面临的难题颠倒了过来。与没有足够的创意可选不同，突然之间，我们变得有太多的创意可以选择。这仍然是个数据问题，只是反过来了而已。使人们具有创造力的是识别按照正确原则处理某一创意的能力。1997年史蒂夫·乔布斯重回苹果之后不久，他将创新描述成否定1 000个可能创意的能力。“你必须精心选择。”他说，“实际上，我因我们没做那些事而感到骄傲，就如同因我们所做的事感到骄傲一样。”乔布斯最终领导苹果创造了iTunes、iPod、iPhone和iPad，但在此之前，他否决了几十个他离职期间苹果公司一直研发的产品。

幸运的是，机器也越来越擅长处理自己的任务。只要我们能够告诉它们我们在找什么，它们就能创造新的富于想象力的解决方案，甚至会超过完成这项工作的人类。

人工智能能否成为发明家

1996年杰森·洛恩（Jason Lohn）加入美国国家航空航天局时31岁。他是个训练有素的电子工程师，之前一直在谷歌工作，洛恩的任务是设计宇宙飞船执行任务时使用的天线。“天线在宇宙中非常重要，”洛恩说，“如果没有良好的天线系统，你发射的飞船可能就只是一个造价昂贵的金属球，因为我们没有办法与之交流。”天线优化的问题在于如何建立一个带宽尽可能高的最佳通讯频道，与此同时，它的体积还要充分地小。早期天线一次只能发送少量字节的信号（两位数）。忽然有一天，梦想变成了收到来自太空的全动态视频流。洛恩非常清楚这个问题的复杂性，然后他想到了解决问题的办法：为什么不把设计流程交给人工智能？

“人们已经将人工智能用于制订多年计划等任务，但我要用人工智能改进太空任务中实际硬件的性能。”他说道。

还是本科生的时候，洛恩阅读了理查德·道金斯（Richard Dawkins）的著作《自私的基因》（*The Selfish Gene*），这是关于基因进化方面最有分量的书之一。“我完全陶醉于自然选择的力量。”他接着说道。大学期间，洛恩开始探索复制这种进化流程用于解决设计问题的想法。

本质而言，这不是一个新想法。几百年以来，人类一直掌握着进化流程，哺育最符合我们需求和希望的新动植物物种。11世纪，杰出的波斯学者艾布·莱哈尼·比鲁尼（Abu Rayhan Biruni）发现了林业工人让树长得更好的秘诀——留下他们认为好的枝干，砍掉其他枝干。在18世纪英国农业革命期间，这个概念被一个名叫罗伯特·贝克韦尔（Robert Bakewell）的人转化成了一门科学。贝克韦尔发现，通过控制繁育，他可以得到产毛高的绵羊和产肉多的肉牛。随着越来越多的农民听从贝克韦尔的领导，家畜在体型和品质方面都得到了提高。1700年，待售屠宰的肉牛平均重量约为168公斤，而到1786年，平均重量已经翻番，达到381公斤。

据洛恩所知，计算机可以使用人工智能模仿自然选择的想法实现同样的结果。就像贝克韦尔的牛羊，算法进化先让创造者制定其要实现的目标。“以天线为例，你可以告诉算法你需要一个解决方案：天线要能够装进一个10厘米×10厘米的盒子，以球式或半球式向外辐射信号，要能够在特定的Wi-Fi频道上运行。”他说，“你提出所有的要求和规定，随后算法按照这些要求优化解决方案。”

洛恩估计，说服高层相信进化算法才是未来之道，他在位于加利福尼亚的家到位于华盛顿特区的美国国家航空航天局总部之间穿行了三四十次。最终他们同意放手去做。洛恩掌握了即将到来的“空间技术5”（Space Technology 5）任务的说明，将其输入空间天线的基本要求之中，然后让他的软件开始工作。

经过数百次生成，算法最终给出的结果看起来像个错误。洛恩称其设计类似一个“曲别针”。他备感失望，就像信誓旦旦保证朋友可以胜任工作，却看到他第一天就酗酒并醉倒在桌子上一样。然而，洛恩忠诚地制造了一台实体原型机，并将它送到了测试室。原型机比他所见到的其他所有解决方案都要出色。随后的设计也是同样的结果，但是由于它们包括的不必要元素过多，洛恩深感茫然，不知道如何解释它们为什么可以

工作得这样顺利。

“作为工程师，我们通常能理解计算机设计的一两个方面，但我们却不理解设计的其他部分。”他说，“如果使用进化算法对天线进行优化，我只有50%的可能可以准确解释它做出选择的理由。其余情况下，对我们来说，计算机设计是不能被理解的。但是它能正常工作，而作为工程师，我们最终关注的是让一切正常工作。”

并不是所有遗传算法想出的解决方案看起来都这么违背常理。因为适应度函数由程序员输入，所以算法只能根据运行人员希望看到的标准解决方案创造各种变体。就像谷歌的Deep Dream项目，尽管它既可能让计算机崩溃，也可能使计算机不受前期发生的情况阻碍而设想各种解决方案。

在洛恩的例子中，结果不言而喻。如果他在向美国航空航天局的上级申请这项任务之前有疑虑，这些结果当然不可能存在。2013年9月6日，美国航空航天局发布了月球大气与尘埃环境探测（LADEE）任务，用以研究月亮的月尘环境。探测器上使用了三个天线，都是由洛恩的人工智能设计的。“这个任务只使用这种天线，”他说，“如果它们失败了，也没有其他天线可以挽回败局。”这次任务相当成功。洛恩将之称为职业生涯的亮点。

洛夫莱斯测试

洛夫莱斯测试

在第四章，我论及了图灵测试，迄今它还是检测机器智能最著名的测试。然而，由于创造力问题已经成为人工智能界最突出的问题，一些人已经建议创造一个新的测试。

涉足现代计算而讨论机器创造力话题的第一人是阿达·洛夫莱斯（Ada Lovelace），她是世界第一批计算机程序员之一。她的创造力并不低于她的父亲——浪漫主义诗人拜伦勋爵（Lord Byron）。19世纪初，她与查尔斯·巴贝奇（Charles Babbage）一起研究他的分析引擎。这是历史上第一台机械通用计算机，但是由于缺乏资金，一直没有完工。虽然洛夫莱斯被建立分析引擎的想法深深打动，但她认为它不可能真正思考，因为它只能运行设计好的指令。“在生成原创事物方面，分析引擎没有值得骄傲之处，”她恰当地写道，“只有我们知道如何命令它，它才能执行工作。”

相应地，洛夫莱斯测试是一种测试洛夫莱斯理论的方法，要求计算机自发想出创意。洛夫莱斯测试设计了一个实验，在实验里人工代理是 a ，人类创造者是 h ，而原始概念是 o 。只有 a 能够不需要 h 解释实现方式就能生成 o 时，这个测试才算通过。为了避免 o 成为随机事件，如偶然的错误， a 必须能够应任何测试评判人的要求重复生成 o 。

在论证人工智能通过洛夫莱斯测试还有多久的观点上，研究人员各持己见。在我看来，我们已经非常接近通过测试，而且我们甚至可能已经通过了测试。杰森·洛恩创建的进化算法之类的算法当然满足了很多标准。尽管洛恩自己无法解释人工代理创建的原始设计，但是事实证明它非常有效。

洛恩的进化算法远不是一个异常值。除了设计，进化算法也用于需要创造力的所有领域。比如，一些建筑师使用进化算法按照具体建筑风格想出了数千个设计版本，致使这些设计看起来像谷歌Deep Dream版的办公大楼或巴洛克式大教堂。还有一种进化算法称为“Eurequa”，由康奈尔大学的创新机器实验室的主任和该大学的一位博士开发，它一直从事描述自然法则的实验。Eurequa的创立者胡德·利普森（Hod Lipson）表示：“Eurequa不是一种在旁边等待的消极算法，它会提出问题，这就是好奇心。”

进化算法与创意策略一起运用效果也很好。2010年，一位懂得人工智能知识的《星际争霸II》玩家利用进化算法，开发了一款能够生成最佳战术的软件进行游戏。这个名为“进化室”（Evolution Chamber）的算法能够让玩家设定游戏目标，随后提出可以最快捷完成目标的方式，它可以详细列出精确的顺序，玩家依据这个顺序建立游戏的战斗单位和阵形。不难想象，这种创造力或许在商业、法务和制药领域能占有一席之地。首席执行官、律师或药物研究员不仅仅是录入数据，命令你重复以前所做事。三者最成功的例子莫过于能够将创造力应用于他们的当前工作，想出一个新方法解决供应问题，做出一个原始法律论据或开发一款新药。

2015年夏，我访问了曼彻斯特大学生物技术研究所。研究所内的一间空调房发出了令人厌恶的警报：“死亡警报：你必须马上离开！”此时，我遇到了一台名叫EVE（早期机器人科学家ADAM的继任者）的机器人，它的任务是将药物发现流程自动化。因为开发一款新药物可能耗资3亿英镑，把这项工作交给人工智能意味着将发现更多的药物。同时，这也意味着为世界某些最易受伤害的人找到更多的药物，因为制药公司通常希望利用所开发的药物攫取最大利润，而不愿意花时间和资源治愈主要是穷人感染的热带病。EVE不仅承担测试新药的工作，而且还要提出用于要测试的药物的假说。同样，它制定理论以解释它的发现，设计实验检测这些理论，实际进行实验，最后再解释实验结果。

“如果是人类做这份工作，那一定会被认为是具有创造力的，因为它以制定和检测假说为工作重心。”EVE的创造者、曼彻斯特大学机器智能教授罗斯·金（Ross King）如是说道。金告诉我，他的研究方向是他所说的科学自动化。“在我看来，像EVE这样具有创造力的机器成为常规只是时间问题。”他说道。

所以，如果人们认为药物发现机器人在制药领域具有创造力，那么，在法律上机器人的创造者处于什么地位呢？约翰·科扎（John Koza）是进化算法的发明人之一，据他所言，这个限制已经被突破。和其他许多科学家一样，约翰·科扎对自己的可量化证据情有独钟。创造力是一种值得拥有的好品质，但是只有一个别出心裁的想法是以前从来没有过的，才能推动我们前进。检验一个想法是否具有原创性的一个方法就是申请专利。许多年以前，科扎开始将进化算法作为“自动发明机”使用。到目前为止，他的算法已经建立了约76项能够在一些领域与人类竞争的设计，如电路设计、光学、软件修复、民用发动机以及机械系统等。大部分设计都已经被其他人申请了专利。尽管在没有任何前期背景知识的情况下，计算机找到了同样具有创造力的解决方案，进而展示了计算机在设计流程中的价值，但计算机完成这些设计的时间真的太晚了。但在其他例子中，科扎的自动发明机能够创建新颖的且可以申请专利的概念。

在科扎看来，获得专利之所以对于判定机器是否具有创造力这个问题如此重要，是因为藏在美国专利商标局法律细则里的一两个词构成的短语。根据专利商标局的规定，专利只能颁给一个被认为是“不合逻辑步骤的”想法。“专利商标局所谓的通过一个‘不合逻辑的步骤’的意思是，一旦你完全按照过去的老路解决问题，你就必然想不到这种方式。”科扎解释道。为什么这两个词这么重要？科扎说，因为如果计算机只擅长执行逻辑运算，它们怎么能够采用不合逻辑的步骤呢？

这足以使阿达·洛夫莱斯挠头了。

祝贺沃森大厨

祝贺沃森大厨

2015年9月，我吃了第一顿按照人工智能提供的菜谱备办的套餐。这道菜是印度姜黄肉菜饭，虽然不寻常，但依然非常美味。我是IBM沃森菜谱的“粉丝”，沃森是曾在美国智力竞赛节目《危险边缘》中获胜的人工智能，在前面章节我们已经说过此事。“此书的菜谱创意并不是由普通厨师想出来的。”《和沃森大厨一起学做菜》（*Cognitive Cooking with Chef Watson*）一书这样写道：“哪种古怪的菜谱中曾有土耳其加韩国风味的恺撒沙拉或古巴龙虾海鲜杂烩？而且，这个设计者什么食物都没有吃过。事实上这个厨房奇才根本没有味蕾，没有鼻子，也没有任何食物或饮料的感官体验。”

这或许令人吃惊，沃森大厨是沃森在《危险边缘》中击败肯·詹宁斯之后IBM才成立的项目。由于沃森能够分析《危险边缘》中的即问即答难题，还能轻易击败答题最娴熟的人类，IBM研究团队的工作人员决定下一代的逻辑挑战要超越问答模式并开发一个具有创造力的系统。各种建议蜂拥而来，包括来自艺术、文学和音乐领域的。最后，IBM倡议智慧城市的科学家拉夫·瓦什内（Lav Varshney）提出在一个所有人类都能欣赏的领域建立人工智能：食物。

“开始时，我们让沃森分析了约9 000个现有食谱。”沃森解决方案首席技术官罗布·海伊（Rob High）说，“在此基础上，沃森的系统能够学习不同类型与风格的现有食谱。它掌握了沙拉与三明治或蛋乳饼与意面之间的区别，也学会了区分越南菜与西南风味，或者辨别法国菜与中餐的不同之处。它可以指出同一类型菜肴中哪一个味道最浓。”

这或许已经足以开始找到味道之间的关系，帮助它做成奥地利烤芦笋或北欧鲑鱼蛋乳饼（这两道菜的菜名还有点押头韵），但是IBM希望继续钻研。沃森的研究人员决定将包括9 000个食谱的数据集与罗布·海伊称“味道化学知识”的另一个数据集合而为一。正如海伊所解释的，当选择食材时人类大厨可能拥有渊博的头脑，但沃森能够分析决定味道的化合物，并使用这些化合物生成全新的组合。“为了找到化合物和包含这些化合物的食材，为你做出最佳的菜肴，沃森进行了6 000 000³次排列组合运算。”他说，“这些菜肴经常出人意料，但它们却非常讨人喜欢。”（一个独特的食物组合是樱桃搭配蘑菇，毫不夸张。）

现在，已经有了多个版本的沃森大厨。与任意名厨一样，你可以买到其官方食谱。你也可以用智能手机下载应用程序，这种程序帮助用户高度控制厨房人工智能来选择食材。首先，用户可以输入多达4种他们想要的食材，或许选择任何搁在冰箱里的食材。随后用户可以选择喜欢的口味，是否需要避开某些食材等等。沃森会立即处理这些信息，并生成食谱，一步步写明烹饪指令。

不难想象，它将比现在走得更远。早期版本的沃森大厨仅生成了可食用的食材清单，随后这些单子交给

人类厨师转换成食谱。最新版的沃森大厨现在能够自主创建食谱，提供食材数量和烹饪时间方面的建议。未来10年中，人人都有一个厨房机器人并不是什么值得大惊小怪的事，就像前面章节讲述的动力机械公司做汉堡的机器人，能够从头到尾执行全部流程。你有没有发现自己突然渴望某种不知道是否存在的食物，比如韩式鸡肉蛋卷？没问题！当你离开的时候，可以通过自己的iPhone 12设置参数，而到家的时候，食物就等你品尝了。

但是罗布·海伊和我交流过的其他几位IBM员工都热切强调，沃森大厨不是IBM在人工智能上的最终目标。就像让一台世界上最强大的人工智能参加一个游戏展览比赛，本质而言，沃森大厨这个名称是个比喻，是概念验证，用以展示沃森能够利用其庞大的自然语言数据库进行不同领域的工作。创造力不仅仅在于准备新菜肴，还在于提出在所有领域帮助人类的方法。比如，IBM正在从事的一个领域是制药，因为沃森能够帮助肿瘤学家挑选更加个人化、更具创造力的方法治愈癌症。还有一个领域是法律方面，因为沃森可以帮忙找到判例法，代表律师形成更具创造力的论据。最后一个就是业务分析领域了。

“不管这些系统变得如何智能，我认为都必须把这些发展看成是工具。”罗布·海伊表示，“就像沃森大厨无法取代人类创造力，它只是放大了人类的创造力。它将使我们比没有它的时候更加具有创造力。它将超越烹饪领域，它是一支描绘21世纪的彩笔。”

更确切地说，这是一支将慢慢画出自己作品的彩笔。

第七章 意识上传实现人类永生

第七章

意识上传实现人类永生

马里乌斯·乌尔萨凯（Marius Ursache）希望你能长生不老。

在像技术行业这样几乎全是年轻人的行业里，几乎没有人花费太多时间去考虑死亡，这不足为奇。毕竟在这个行业里，人们21岁时就已开始二次创业，25岁时就成为亿万富翁，以至史蒂夫·乔布斯曾一度担心超过30岁的人还能否成就终身伟业。因此，在大多数硅谷“谷民”的头脑中，衰老和死亡都是没影的事。

来自罗马尼亚、曾经是医生的乌尔萨凯比大多数人思考的死亡问题要多。他甚至已经将之变成了工作。作为一家名为Eterni.me的创业公司的创始人，他夜以继日地工作，梦想着能够构建出人工智能3D替身，即那些外观、声音和举止就像我们死去的亲朋好友一样的数字生物。

乌尔萨凯的创业旅程开始于数年之前，当时他正沉迷于游戏《第二人生》（*Second Life*）。这是由旧金山的一家游戏开发商林登实验室（Linden Labs）开发的一款大型虚拟世界网络游戏。尽管《第二人生》类似于电脑游戏，但它在一项关键的功能上与电脑游戏不同。在《第二人生》中，玩家不用设定目标和加工故事情节，而是将自己作为游戏中的“居民”，并以自己想要的任意方式参与其中，比如开店，或者仅仅是与朋友闲逛。

“一天，我开始思考：人死之后，他在游戏中的替身会发生什么事？”乌尔萨凯说。他想，那里会不会成为第二人生的炼狱之城，被遗弃的替身们在他们的人类操作者离世很久以后像僵尸一样继续存在？如果有人试图与这些替身互动，会发生什么？

他开始越来越频繁地考虑这个问题。他试图找出编写能够真正模仿其人类操作者的人工代理程序的逻辑。他认真思考编写替身所需的程序，以便替身可以学会像其人类玩家那样去活动，交谈，制定和追求目标。他像锐意进取的企业家那样，设法想出将其转变为实际产品的方式。

2014年2月，乌尔萨凯应一位在布加勒斯特遇到的导师邀请，参加麻省理工学院举办的一个企业家计划。该计划的任务之一就是要设计一个可以从事的项目。而当时，他脑子里已经有了一个想法。他说：“我认为，这可能是人们收集和整理一生数字足迹的好办法。替身将会成为访问这些信息的一个界面。”

他向团队提出了名为“与逝者进行网络对话”的创意，并且迅速注意到：世界各地的各类实验室已经存在大量实现该项目所需的人工智能技术。尽管团队收到了共130个创意，而乌尔萨凯承认他的创意是最古怪

的，不过最终“与逝者进行网络对话”的想法被视为值得探索的一个项目。

然而，乌尔萨凯对此有所保留。他说：“我知道，这不仅仅是模拟与逝者的对话。这可能会一团糟，陷入哀伤之中，而且坦白地说，实在是有点诡异。”

他决定通过一个网页来测试公众的反应。如果人们响应积极，他将继续进行下去；如果这个想法无人问津，或者引起公愤，他就放弃。

在最初的4天，网页获得了3 000人的签名，他们表示对此感兴趣。这个数字很快飙升到了22 000人，而且随后持续攀升。网页还收到了大量消息，乌尔萨凯将它们当成市场调研活动，尽职尽责地一一阅读了这些消息。多数消息对该项目赞不绝口，然而也有一定比例（他估计约1/5）的消息认为这个想法听起来非常诡异。有谁会想要一个言行像自己去世的祖父母一样的Siri版替身呢？

乌尔萨凯随后收到了改变其一生的电子邮件。这封电子邮件来自一名即将死于癌症晚期的患者。在邮件中，他说他仅有6个月可活，他写道，Eterni.me这样的项目是他死后为朋友和家人留下些念想的一个机会。

乌尔萨凯说：“回复赞同或批评的消息很容易。但是我能对一个临终的人说点什么呢？那一刻我决定这就是值得我奉献一生的事业。”几乎整个晚上，乌尔萨凯决定停止他之前的工作，转而全身心地专注于Eterni.me。

如今，Eterni.me已经拥有30 269位热心的订阅者，所有这些人都在等待数字化永生的入场券。公司的网站上播放着描绘一生中各种记忆的视频短片：新郎和新娘在婚礼上接吻；母亲拥抱自己的孩子；孩子在花园里扮演超级英雄；毕业生将学位帽抛向空中；退休人员在一起谈笑风生等等。乌尔萨凯的营销广告这样写道：“如果你能永远保存对父母的记忆会怎样？如果你让他们的故事在你的子女、孙辈乃至世代子孙中永远传承，又会怎样？如果你能将自己的传奇一生保存到未来，又会怎样？这样，100年后，你的后代、朋友或者来自遥远未来的陌生人也将记得你。如果你能够通过数字替身得以永生，未来的人们能够与你的记忆、故事和想法互动，就像在与你对话一样，这又会怎样？Eterni.me收集你的思维、故事和记忆，管理它们，并创建一个长得像你的数字替身。这个替身将永生，让未来的其他人访问你的记忆。”

当前还不存在这种使我们可以像乌尔萨凯最终希望的那样“与逝者进行网络对话”的技术。但是他的团队不懈地研发机器学习工具，正努力将这项技术变为现实，而Eterni.me则专注于收集用户数据，以便有一天，为用户的替身提供数字生命线。尽管如此，他认为Eterni.me的30 269名早期采用者不会一直等下去。

他说：“这项技术不是几十年能搞定的。构建仿真替身是一个迭代过程。它就像是搜索结果，只会随着时间的流逝，越来越完善，越来越精确。”

利用机器实现死后永生

利用机器实现死后永生

然而，马里乌斯·乌尔萨凯并不是第一个考虑机器如何使人死后永生的人。数十年前，多人网络游戏就不得不面对当红玩家或创建者死后的局面。以《龙与地下城》（*Dungeons and Dragons Online*）这款桌面游戏为例，在联合创始人之一加里·吉盖克斯（Gary Gygax）去世后，游戏中创建了一个以他的录音做旁白的任务。那是一种坟墓中传来的缥缈可怖之音，提醒玩家在靠近任务发生的地方可以找到吉盖克斯的虚拟坟墓。

此外，还有更多其他案例，它们都表明了人类及其决策过程是如何嵌入机器代码的。其实，第一章中所描述的“专家系统”就试图创建真实人类专家的克隆版本。通过提取他们的专业知识并将其转化成一套概率规则，使其能够进行计算机处理。如果专家系统按计划实现，杰出的心脏病专家、律师或是首席执行官就可以将自己的专业知识复制并传播到全世界。

在某种意义上，与专家系统相并行的现代系统就是所谓的“推荐系统”(recommender system)。这种信息过滤子系统旨在预测用户可能对某一特定狭窄域内的条目进行评级或选择。阅读这个条目的每个人都可能遇到亚马逊或网飞公司的建议：“你喜欢X，所以你可能也喜欢Y”。有时这些预测不太重要，但是就像我们喜欢把自己当成完全不可预知的物种一样，这些预测常常准确得惊人。

至少很多公司认为如此。2014年1月，亚马逊获得一项运输系统的专利，该专利旨在通过预测人们的购买倾向并尽早发货来大幅缩短交付时间。在某些情况下，亚马逊称其可以将货物发送至当地运输中心，直至订单如期蜂拥而来。在其他情况下，它会注意送出有针对性的促销礼物，以建立信誉。

和本书讲述的其他技术一样，这种预测的关键是分析庞大用户数据的能力。每年我们保留的自身数据量都要超过我们身上的DNA碱基对的数量。每次进入网络空间，我们都会留下痕迹。每当我们在博客平台WordPress或LiveJournal写博客，在推特或Facebook更新状态，使用Contextly应用评论新闻，在美国在线影片租赁商Netflix上选择电影或电视剧观看，发送即时消息，或者仅仅利用谷歌进行搜索，我们的数字身份就得到了更新与管理。其结果是由1和0呈现出关于我们的更加精准的描述。未来几年，随着个人信息泄露到人工智能系统中，这种曾经模糊的轮廓将日益类似于精细的素描，甚至是逼真的油画。就像我们在第三章所看到的，这种知识将会重新配置和优化我们周围的世界。你的家门或者所住的酒店房门将只在你走近之时为你打开。你从未驾驶过的租赁汽车将自动调节以满足你的偏好设置。恒温器将深知你喜欢的温度并自动调节，以便提高你在一天中不同时间的效率。

这不只是采集更多数据的问题，而是由亿万人常规记录的全新、不同类型的细粒度数据。比如，苹果手

表可以收集一生中的心率，同时不同的苹果手表应用可以通过不同的数据点互相参照这些信息。在2014年拉斯韦加斯消费电子产品展期间，索尼推出了其“生活日志”软件，旨在沿着互动时间线，追踪人们从给朋友打电话到看电影的活动。所有这些数据在20世纪90年代还是难以想象的，当时只有实验室的一次性实验数据。

个性捕捉

个性捕捉

威廉·西姆斯·本布里奇（William Sims Bainbridge）现年76岁，是美国国家科学基金会网络人类系统的联合主任，他一生中大多数时间都沉迷于死后人生的主题。他在自己的办公室保留了一幅姑婆克里奥拉（Cleora）的照片，他说这么做是为了提醒自己面对自己的死亡。这幅照片年代久远，摄于1870年，照片的感光乳剂已经开始从古老的金属支架上剥落。这是克里奥拉·埃米莉·本布里奇一岁时的样子，在她夭折前不久拍摄的。尽管我们多数人曾经发生过祖先夭折的事，但在克里奥拉去世之后，她的身上发生了一些不同寻常的事。她去世时，她的教父决定虚构一些她不曾经历的生活来纪念她。他通过小说的形式来纪念克里奥拉。这部小说写于1883年，叫作《无私》（*Self-Giving*）。在小说里，克里奥拉长大成人，并且成为一位教士，最终殉道。

不管是什么导致克里奥拉的教父采取这种非同寻常的方式来想象她死后的生活，这种行为似乎都被承袭了下来。差不多一个世纪之后，1965年，本布里奇的父母和姐姐因家里起火而惨死。他那年25岁，是哈佛的一名社会学专业的学生。失去了亲人的本布里奇为了保留亲人的记忆，花费数月时间在灰烬中仔细查找，去抢救一切他能找到的值得纪念的物品。

对于保存记忆的兴趣最终将他引向神经网络。由于约翰·霍普菲尔德和其他人的工作，20世纪80年代初，神经网络已经被证明能够保存联想记忆。受到见闻的启发，本布里奇猜测足够大的神经网络有一天或许能够存储人类的全部记忆。随后，他根据著名的“运动捕捉”技术，将这一理念称为“个性捕捉”。在“运动捕捉”中，人的运动被扫描进计算机中，用于电子游戏和电影。本布里奇关于个性捕捉的大思路是，通过询问人们大量关于他们自己的问题，然后使用这些信息模拟我们所谓的人工智能“头脑文件”（mindfile），就有可能精确模拟人在任何既定情境中的行为方式。他指出，为了创建个人的真正精准视图，需要人们回答接近100 000个问题。然后，每个问题将从两个维度来衡量，一个是每个个体性格的相对重要性，另一个是对该个体适用性的相对程度。这个结果将成为能够以和其真人原型相同的方式响应现实生活环境的软件替身。

尽管这还没有实现，但是并无原因明确显示其不可能成功。20世纪80年代，贝尔通讯研究所的一位名叫托马斯·兰道尔（Thomas Landauer）的研究员进行了一系列实验，旨在找出人类一生中能记住多少事。这些实验包括要求人们看图片以及听单词、句子和音乐片段。过了数分钟或数天后，兰道尔去测试参与者能够记起多少内容。通常采用多选问题来进行测试。尽管他的估计无法面面俱到，但他推断，一般人每秒大致可以存储2比特的信息。人的一生可以记忆总共 10^9 比特（约数百兆字节）的信息。鉴于我撰写本书所用的计算机已经存储了3兆兆字节的信息，根据兰道尔的计算，即使再低端的个人电脑也可以存储数千个人类头脑文件。

就像艺术家画自画像一样，把头脑文件和后来的虚拟替身训练得和你一模一样，可能是人们一生追求的一种自我表达方式。由于我们中大多数人不一定能意识到自己的错误，本布里奇说，可能有必要把一个人的家庭成员、朋友和同事召集起来，来提供人们在可信度、职业道德等领域的细节。

与马里乌斯·乌尔萨凯类似，本布里奇的个性捕捉可以使用人工智能模拟挚爱或其他重要的人物，以便在他们去世后读取他们的记忆和观点。这些替身在未来将成为能够与后代互动的历史文物，这种互动与我们今天注视老照片或者聆听人们生前录音颇为相似。你可以将其当作一个存储着人而非书籍的存储库。关于这个概念的一个更加基础的变体已在几年前由卡内基-梅隆大学付诸实施，它被称为“人造会谈”（Synthetic Interview）。通过创造能够回答问题的互动式视频体验，“人造会谈”的想法为亚伯拉罕·林肯等历史人物带来活力。为了这个项目，演员装扮成历史人物，并回答用户可能提问的数百个问题。一个非常基础的软件随后会选择恰当的视频以回应每一个提问。卡内基-梅隆大学娱乐技术中心执行制片人唐·马里内利（Don Marinelli）说：“人造会谈使人们深入地与一个人或就某一段时期进行互动，这是被动地看电影时不可能实现的。而且当我们这项技术与一个像亚伯拉罕·林肯这样受人尊敬的人物相匹配时，其效果极具说服力。”

最终，个性捕捉有可能成为某个序列上的下一步，而当史前社会首次在洞穴的墙上画下一些图像，以之作为战胜死亡的符号时，这一序列就开始了。当我们注视着罗浮宫里挂着的《蒙娜丽莎》，或者阅读《麦克白》的时候，我们窥见了这些作品的作者内心，尽管列奥纳多·达·芬奇和威廉·莎士比亚已经在几百年前就离世了。一种所谓的“思维克隆”可能有一天也会做同样的事。除了记录记忆，人工智能行为建模甚至可能用于预测人们如何响应在其死后发生的事件，就像现在我们可以访问一位已故亲属生前使用的亚马逊账户，获得符合他们口味的图书、电影和音乐专辑推荐一样，而有些推荐可能是在他们死后创建的。

当然，马里乌斯·乌尔萨凯谈论的这种“虚拟永生”对人们来说是一种非常特别的永生，它意味着作为有意识的生命体而继续存在。就像伍迪·艾伦（Woody Allen）曾经讽刺道：“我不想通过我的作品而实现永生，我想通过长生不老而实现永生。”

人工智能能否也有助于解决这个问题呢？

延长人类寿命

多年来，人们都在思索这一可能：有一天，人类能否超越人体的极限，运用科学手段转移自己的思维并将之保存。这种思索可以追溯到十分有影响力的俄罗斯哲学家尼古拉·费多罗维奇·费多罗夫（Nikolai Fedorovich Fedorov）。费多罗夫出生于1829年，是俄罗斯宇宙主义者。宇宙主义融汇了东西方哲学的理论，具有俄罗斯东正教的虔诚与率直。今天一些未来主义者的这种宗教般的热情通常与最新的尖端科学齐头并进，这很大程度上脱胎于俄罗斯宇宙主义。费多罗夫将进化设想成以智能和实现它的探索为中心的过程。他认为，人类是自然史的顶峰，应该运用一切可行的理由和道德去指引自然选择之手。这种观点引发了人们的一种兴趣：利用科学方式实现彻底的寿命延长和肉体不朽，并且最终恢复所有逝者的生命。所有这些死而复生的人们生活在哪里呢？费多罗夫认为太空和海底可以成为人类的殖民地。就像众所周知的谷歌将民主和数据驱动的方法融入生活中一样，费多罗夫宣称：“人人都必须学习，一切都是知识和行动的对象。”

尼古拉·费多罗夫的想法很快受到了他人的追捧。法国社会学家珍·菲诺特（Jean Finot）于1909年撰写了《长寿哲学》（*The Philosophy of Long Life*）一书，在书中，他推崇利用科学来设计人生和制造生活物质。英国进化生物学家J. B. S. 霍尔丹（J.B.S. Haldane）1924年发表的《代达罗斯或科学与未来》（*Daedalus; or, Science and the Future*）随后认为，进化可以通过定向突变在子宫里进行。在作家奥尔德斯·赫胥黎（Aldous Huxley）创作其反乌托邦小说《勇敢新世界》（*Brave New World*）时，霍尔丹的想法对他影响很大。在1927年出版的《为生存能力而奋斗》（*The Struggle for Viability*）中，布尔什维克作家亚历山大·波格丹诺夫（Alexander Bogdanov）宣称：自然寿命正因为社会缺陷而被人为缩短。如果改善这一状况，波格丹诺夫写道，我们的生命至少应该“持续120—140年”。同期，X射线先驱J. D. 伯纳尔（J.D. Bernal）于1929年撰写了《世界、众生和恶魔》（*The World, the Flesh and the Devil*），这本书详细说明了费多罗夫的“空间殖民”的概念，并描述了其延长人类寿命和增强人类智能的理论。

这些长篇大论今天读起来仍然引人入胜，但是只有当在计算机、机器人和人工智能成为20世纪晚期交流工具的一部分时，超人类主义才真正生根。特别是随着年华老去，许多人工智能保守派也将注意力转向了能够帮助他们延长自然寿命的技术上。

其中一人就是马文·明斯基。1994年，67岁的明斯基为《科学美国人》写下了一篇名为“机器人是否将接管地球”的文章。在这篇4 500字的文章中，明斯基指出，我们了解死亡，我们正在解决死亡问题的途中。他将死亡划分成两种独立的范畴：肉体的分解和灵魂的分解。他指出，身体是（相对）简单的部分。之前他就指出，人类的平均寿命从古罗马时期的22岁到1900年发达国家的50岁，一直在不断增长。随着科学对基因的不断研究，明斯基认为我们将能够改变许多当前影响人们未来生活的条件。他将人类与动物界进行比较，并提到了一个涉及地中海某种章鱼的实验。产卵之后，这种章鱼不再进食，并且很快饿死。然而，当科学家将孵卵的雌性章鱼的视神经腺体移除，他们发现章鱼抛弃了自己的卵，并重新开始进食和生长，比未经改变的章鱼寿命大约延长了两倍。出于延寿的目的，我们也在人类身上也进行了类似的手术——移除恶性肿瘤或者对人们进行重要器官的移植以挽救生命。

正如人们从将整个职业生涯都花在了人工智能的研究上的科学家那里了解的那样，明斯基认为大脑更加复杂。他写道：

作为一个物种，我们似乎已经达到了智能发展的顶峰，毫无变得更加智慧的迹象。阿尔伯特·爱因斯坦是比牛顿或阿基米德更伟大的科学家吗？近几年有超过莎士比亚和欧里庇得斯的剧作家吗？2 000年来，我们已经学会了很多，然而，仍有许多古代智慧似乎颠扑不破，这让我们不禁怀疑我们并未取得多大进步。我们仍然不知道如何应对个人目标与全球利益的冲突。无论何时，我们在制定重要决策时总是缺乏判断，我们对自己不确定的事总是听天由命。

当互联网开始腾飞时，建立一种“蜂群思维”的想法引发了时代思潮。但明斯基并不仅仅是希望保存信息并使其为后代所用。他希望全面地看待它。他写道：“最终我们将使用纳米技术完全替代我们的大脑。一旦突破了生物学的限制，我们就能够决定我们生命的长度，不但可以选择长生不老，还可以选择其他现在无法想象的能力。”

模拟神经元

马里乌斯·乌尔萨凯和威廉·西姆斯·本布里奇所描述的那种复杂的推荐系统“头脑文件”可能以软件形式复制我们。然而，确保将一个人按照不同于原始的方式重建的唯一真正可靠的方法是，通过提取神经元来复制大脑中的所有细胞通路。

要使上述方法成为可能，我们必须首先接受人工智能的核心原则：大脑执行的主要任务可以被视作信息处理，这与计算机执行的任务没什么不同。换句话说，计算机系统所用的软件与所谓的人脑“湿件”没有本质区别。这种智能模式要求我们遵守“基质独立”原则，这意味着大脑作为一种动态过程，并不一定与一组原子相关联。如果大脑的信息处理是真正的基质独立，那么这就意味着某一天，它可以将智能从以蛋白质为基础的大脑转移到另一种更加持久的媒介中，如计算机网络。

因此，问题就演变为我们如何建造这样的一个大脑。最简单的答案是，就像任何一个曾经把闹钟拆开想要了解其工作原理的人，他知道这是对它的“反向设计”。这种行为就是把现有软件或硬件拆开，从而了解它的生产方式。一旦我们了解了它的生产方式，我们就可以用同样的方法建立一套完全相同的模型。如果我们碰巧知道它已经建造成功，还获取了一些输入和输出的数据，我们就可以训练神经网络，使它对本书所述的任何事做出相同的举动。

当今时代，把大脑当作软件最成功的尝试是深度学习神经网络。基于对生物大脑的简单模仿，这些网络已经变得越来越复杂，并且越来越成熟。当马文·明斯基于1994年创作他的论文《机器人是否将接管地球？》时，大量神经网络构成了约440个连接点。而在我写本章的时候，世界最大的深度学习网络属于一家名为“Digital Reasoning”的美国认知计算公司，拥有约1 600亿个神经连接点。尽管它与人脑的实际复杂性还相距甚远，因为人脑约有86万亿个类神经连接，但这是几十年来一个巨大飞跃。每立方毫米的人脑组织竟然包含100 000个神经元，以及约900 000 000个类神经连接。

如果摩尔定律继续有效，未来几十年内建立一个这种规模的神经网络不是什么问题。不幸的是，单靠这些并不足以生成大脑般聪慧的智能。我们知道，这是因为尽管计算机科学家建立了拥有上百万个神经元的神经网络，但它仍然还不具备可与动物相提并论的通用人工智能。在动物界，100万个神经元应该能让大脑具备蜜蜂这个水平（96万神经元）或蟑螂这个水平（100万神经元）的智能。我们还没有做到这点。事实上，我们离重造真正的动物中枢神经系统的“神经连接体”或接线图最近的尝试是分析一种被称为秀丽隐杆线虫的雌雄同体线虫。早在20世纪70年代，诺贝尔奖得主、生物学家悉尼·布伦纳（Sydney Brenner）和他的同事为了能够使用功能强大的电子显微镜拍摄线虫细胞，开始对秀丽隐杆线虫进行切片。到了1986年，布伦纳收集到了足够多的信息，发布了这种线虫完整的神经系统连接体。至今，它依然还是我们能够解码的所有生物体中最完整的神经连接体。

因为秀丽隐杆线虫的神经系统结构非常基础，7 000个神经突触将仅有的302个神经元连接在一起。尽管这种结构非常简单，但是我们对其神经系统的实际工作原理仍然知之甚少。自2011年以来，用计算机为该线

虫建模的任务一直由美国、欧洲和俄罗斯的数百位科学家和程序员组成的国际团队在进行。这个被称为“OpenWorm”的项目试图建立一个线虫的仿真物理躯体和具体的仿真神经系统模型。然而，尽管花费了大量人力物力，我们依然尚未了解线虫神经元处理信息的机制，而依靠这种机制我们可以复制最基本的爬虫行为。

假设我们已经有了功能足够强大的计算机，那么模拟数十亿互动的神经元就不再是棘手的事了。从已经展开的工作来看，我们知道在网络里组装数十亿个神经元并不能产生智能的人类级大脑，而把数十亿个传感器放在一起至多产生一个功能不错的中央处理器（CPU）。秀丽隐杆线虫这样的神经连接体是一个静止的环路，实际上缺乏所有的环路运行信息。这是因为，当我们查看神经网络的时候，有些参数隐藏在我们无法进入的神经元内。换句话说，你可能查看了现有计算机的蓝图之后就建造了一台计算机，但你可能依然会对编写微软Word程序一头雾水。

如果思维是大脑的软件，为什么我们希望它会有所不同？

绘制大脑

直到现在，神经系统科学主要沿着两个方向发展。一个是研究者专注于个体神经元的微观研究。这取得了一些进步，但它只提供了一些有关大脑机能的有限知识，因为它忽视了发生在神经元周围的大脑网络活动。另一个是研究者关注大脑不同部分的宏观皮层架构，在大脑皮层里，最小的可分解单元可能是几十万个神经元。一直以来，这种研究主要通过实际移除部分人类大脑，并在显微镜下进行分析来进行。今天，我们能够以微创的方式进行这项研究。1990年，日本物理学家小川诚二（Seiji Ogawa）和他的同事创立了一种脑成像技术，称为功能性磁共振成像，简称fMRI，由于能够找出哪部分大脑对某些类型的行为负责，它创造了许多令人瞩目的成就。

比如2015年，加利福尼亚的医生将一对电极移植进一位名为埃里克·索尔托（Erik Sorto）的34岁的四肢截瘫的患者脑中，使他能够利用自己的思维控制一条机械手臂。通过记录索尔托大脑后顶叶皮层（这部分大脑负责处理运动计划）的信号，并将这些信号反馈至旨在分析这些信息的神经网络，索尔托的意图能够被解码，并随后转换成运动指令，发送给独立的机械手臂。索尔托从握手这样的简单任务开始，很快就能玩“石头剪刀布”，甚至可以10年来首次自己端起啤酒来喝。

在另一个相似的实验中，休斯敦大学的研究人员开发了一种大脑机器界面，它不再需要移植进大脑，而只要一个脑电帽，通过这顶帽子该界面可以检测大脑通过头皮的脑电活动。尽管生成的信号比在大脑里实际放置一个纳米微电极产生的电噪多，但研究人员却能够缩小或扩大大脑有用的信号运行的频率。结果，和埃里克·索尔托一样，测试中56岁的截肢患者能够使用机械手臂捡起各种物体，包括水瓶和信用卡。

和本章开始部分所说的“头脑文件”细节一样，从这些项目中获取的知识都能使研究人员建立关于大脑的更加详细的画面。2013年，一组研究人员得出了一个3D人脑扫描图，它占用了一兆兆（ 10^{12} ）字节的空间。尽管这些扫描图依然还不足以回答大脑微观结构方面的问题，但这些图告诉了我们一些大脑显微解剖学不曾告诉我们的细节。

因此，下一步就涉及继续钻研更加精细的细节，按神经元逐个找出事物的功能。就在现在，神经系统科学家已经大致了解了大脑神经元的作用是什么，以及它们是如何与其他神经元相互交流的，但他们仍然无法明确地说出实际交流的是什么。神经元有上百乃至上千个变体，每个变体都有自己的细胞类型和不同的分子特性。现在，我们仍然不知道一共有多少种神经元，或者每种神经元的电学或结构属性是什么。我们也不知道大脑采用什么格式进行编码。我们知道计算机使用JPEG和GIF等文件格式为图片进行编码，或者DOC和TXT格式为文本文件进行编码。理解大脑意味着不仅要理解单个神经元的工作方式，也要理解与它们并行的作为神经网络一部分的其他神经元之间的互动方式。

现在有许多关于如何更好完成这个项目的想法。未来学家雷·库兹韦尔（Ray Kurzweil）目前被谷歌聘为工程总监之一，他建议使用小型微观纳米机器人扫描人脑。这有点像第三章里描述的可注射式智能设备。实现库兹韦尔的愿景需要数十亿个这种扫描机器人，其尺寸如同人类的血液细胞，它甚至可以更小，从而进入

大脑，通过内部扫描来捕捉“所有显著的神经元细节”。

从理论上讲，这是个好想法，但库兹韦尔的乐观想法一直为一些神经系统科学家所批判。因为他这种关于人脑科学的说法就像是建议我们要增强对濒危物种的意识，而做法则是在雨林的中央建造很多公路，以便人们可以近距离观看动物。比如，《神经生理学期刊》主编大卫·J·林登（David J. Linden）指出，大脑不仅由神经元构成，也由所谓的“胶质细胞”构成。胶质细胞的数量是神经元的10倍，而且紧紧地挤在一起，使纳米机器人无法通过。更糟的是，即便是脑细胞之间的极小空间也充满了支撑结构，从而反复地将信号运往相邻细胞。林登说：“你可以设想库兹韦尔的纳米机器人……坠毁在活跃的、脑电活动的微妙网络中。即使我们勇敢的纳米机器人有喷气发动机，并且装备了强大的切割激光，但它在大脑中通行时怎么能够在身后不留下破坏的痕迹呢？”

即使库兹韦尔的说法有些是错的，但他所设想的宏图并没有错。随着人工智能的进步和纳米技术的平行发展，机器人学与神经系统科学是数以亿计的资金涌向人脑逆向工程的原因。就像1956年开启人工智能的英国达特茅斯大会一样，这导致了不同学科间的一些有趣的合作。

下一个大事件

有一个项目是基于皮层网络项目的机器智能，它被称为MICrONS。MICrONS项目由美国高级智能研究专项活动部提供资金支持，其目标是建立功能类似于人脑的算法，从而提高机器智能。其优势在于，尽管所有计算机在特定的文本中都远强于人类，但人脑依然可以比机器更加有效地执行其他任务。举例来说，基于小数据集，在归纳方面我们比最先进的神经网络表现得更加出色。神经网络或许可以在视觉识别任务上击败人脑，但为了完成这项任务，计算机需要查看成千上万甚至上百万个训练样本。另一方面，虽然人类在其一生中可能碰巧看过成千上万或者上百万次某一特定目标，但人类不需要为了识别这个目标而查看所有的目标。如果有人向你展示了一个你以前从未见过的设备或机器，为了从其他类似物中把它识别出来，你可能需要一会儿工夫，但不必考虑展示它时的角度或光线条件。这是因为神经网络依然是大脑激活技术，而非大脑真实的再创造。

MICrONS项目领导R. 雅各布·沃格斯坦（R. Jacob Vogelstein）表示：“当前，多数一流算法都是由神经系统科学原则派生出来的，至少在相当程度上是这样的。但这些神经系统科学原则如今已经使用了20年、30年，甚至是50年之久了。几十年来，神经系统科学与机器学习社群相互之间一直没有多少技术转化。”沃格斯坦说他希望去“弥合当前人工智能算法与大脑中实际发生的算法之间的鸿沟”。

MICrONS项目需要实验神经系统科学家去绘制高分辨率大脑结构与大脑的功能成像图，需要应用数学家分析大脑“图形”。计算神经系统科学家对神经元和神经回路建模，最终机器学习者可以使用这些数据建立展现更多的类似人脑特性的算法。

两个类似的大规模研究项目都在进行之中，它们分别是“通过推动神经技术创新研究人脑的计划”（BRAIN project）和欧洲委员会的“人类大脑工程”（Human Brain Project）。BRAIN这个有些多余的缩写是2013年初由巴拉克·奥巴马总统在国情咨文中首次提出的，其目标是绘制脑电波通路图。这样做可以清楚阐明各种神经障碍，如阿尔茨海默病、帕金森综合征、精神分裂症、抑郁和外伤性脑损伤。

耗资10亿美元的“人类大脑工程”同时也宣布，将花费10年时间，致力于在瑞士日内瓦建立一个完整的人类大脑计算机仿真（或许它会成为银行家的大脑）。为了实现这个目标，项目领导计划对各种运动大脑进行逆向工程，按照其复杂程度，首先从老鼠开始，然后逐步提高。

由于这些计划的合作及政府主导的性质，它可以与许多20世纪的重大研究项目相比较，包括人类基因组计划、人类月球旅行计划以及原子弹开发计划等。尽管所有这些项目拥有不同的方法和目标，但人们希望人类大脑工程可以增加我们对神经元是如何连接的理解，更重要的是，加强对它们如何动态地共同发挥作用的

理解。

然而，并不是所有这些项目都由公共部门承办。过去几十年中，一个非比寻常且雄心勃勃的脑科学项目是2011年俄罗斯亿万富翁德米特里·伊茨科夫（Dmitry Itskov）宣布的项目。鉴于项目的名称为“俄罗斯

2045”计划，伊茨科夫的计划是一项非营利性计划，致力于实现生命扩展的目标。用该项目自己的介绍来说，“俄罗斯2045”计划旨在“创造各种能够将个人品格转移到高级生物载体上的技术，并且延长生命，直至达到永生”。

和BRAIN计划一样，伊茨科夫的计划被分解成了多个阶段。第一阶段要求建造能够被人脑控制的机器人。第二阶段是开发能够通过外科手术移植并寄存物理人脑的机器人。此后10年，伊茨科夫计划能够将人脑内容上传至机器人，这就意味着对人脑进行逆向工程。最后，到2045年，伊茨科夫计划通过全息图取代机器人。

在一封致世界部分最富有的人的公开信里，伊茨科夫请求他们支持他的计划，甚至考虑志愿成为早期实验主体。他写道：“我劝你们关注一下为控制论的永生和人造肢体领域科研提供资金的重要性。这种研究可能使你们以及这个星球的多数人免遭疾病、衰老甚至死亡之苦。”

“俄罗斯2045”计划得到了大量支持。但参与者名单上的一个名字比其他人更能引起我的兴趣。他就是严肃且卓有成就的神经系统科学家肯·海沃思（Ken Hayworth）。

意识上传

还是一个孩子时，肯·海沃思就希望进入太空。像许多在阿波罗太空计划的时代长大的孩子一样，他为太空旅行及其无数人类可能性的想法所痴迷。从那时起，海沃思变得雄心勃勃。他不仅仅想拜访太阳系的另一个星球，还想建造一艘新型火箭，载着我们去最近的星球。

“我是那种对它很痴迷的人，”他承认道，“我开始钻研它是如何实现的这一物理现象。我查看人们事先已经想出的不同设计，然后疯狂地研究它们。”但他最终放弃了这些使他泄气的研究。就像年轻的海沃思所说的那样，这种需要将人类在其生命有限期内带到最近星球的工程已经超过了我们最狂野的技术梦想的能力范畴。即使他明天就起飞，在接近他的目的地之前，这个少年也可能已经死去或已衰老了。

阅读使他接触到一些关于神经网络的书籍。他说道：“那时，我开始考虑其他方式能否完成这个目标。我渐渐理解，太空飞行的问题是人类遇到的真正问题。如果我们可以将人的想法从其沉重且需要生命支持、屏蔽辐射的身体中提取出来，那么这些信息就可以通过无线电波以光速传递。”几乎整个夜晚，海沃思的注意力从建造重型宇宙飞船的硬工程转向了人脑工作方式的问题上，他认为，如果真有可能，思维可能会被拆开并再次组合起来。

肯·海沃思拥有计算机科学本科学位和神经系统科学博士学位，如今是一个名为“大脑保存基金会”（Brain Preservation Foundation）的组织的主席兼联合创始人。这是个略微有些汉默式恐怖风格但运行良好的组织，它老掉牙的名字使人不由自主地想起彼得·库辛（Peter Cushing）和克里斯托弗·李（Christopher Lee）主演的老电影。

海沃思十分明确地谈到了意识的上传。“我绝对相信意识上传是可行的，而且我认为我们应该为此积极努力。”他说：“至少，医生应该赋予人们以高保真方式保存自己大脑的能力。当这种能够以神经突触方式扫描整个大脑的技术出现时，就可能使某个具有仿真大脑和机械肢体的人重回这个世界。”他停顿了一下，然后补充道：“我当然期待这个人就是我自己。”

他同样对那种失败主义者的探讨不予理会，他们认为这些事情根本不可能成功。他在一篇发表的论文上这样写：你愿意接受一项未经检验的手术吗？这个手术会把你的身体和大脑温度降至10℃，使你的心跳和血液停止整整一个小时，你愿意吗？在这期间，你的大脑将处于稳定状态，因为在10℃时，所有神经元间的交流都会停止，这意味着在你复苏前的整整一个小时里，你的身体条件满足了关于死亡的医学上的所有要求。如果你对这个问题的回答“不”，你可能犯下了一个可怕的错误。海沃思所说的“未经检验”的手术过程实际上是一种已经存在的手术，称为深度低温和循环骤停，用于治疗脑动脉瘤等疑难病症。“这个方案唯一不切实际的地方是医生让他的病人自杀（通过拒绝接受手术）来结束这个脆弱的哲学争辩。”海沃思说：“今天，医生可以简单地指出数百份病历，这些患者在接受了这样的手术之后都过着高质量并且健康的生活。”

相比实现意识上传我们需要达到的技术水平，我们今天的技术水平相当于弗兰克·罗森布拉特等研究人

员在20世纪80年代通过神经网络所达到的技术水平。既然这样，整体情况就是正确的，但是我们需要继续查漏补缺。来自神经系统科学等领域的最新见解会大有帮助，同时不断增长的数据集和摩尔定律也有助于完成剩余的工作。通过这种方式实现机器意识有些类似谷歌优化其搜索引擎的方式。拉里·佩奇和谢尔盖·布林在斯坦福大学以“网页排名”（PageRank）起家，直到目前，这种技术依然是谷歌帝国的核心。网页排名按照每个网页的质量和该网页的链接数量排名。尽管网页排名依然是一种至关重要的算法，但谷歌已经使用了200多个不同的信号或者“线索”来强化它，这些信号产生了许多关于用户在找什么的猜测。正如谷歌工程师所说的，除了网页排名，“这些信号包括如网站术语、内容新鲜度和你所在的区域。”人类意识可能贮存在类似的线索中：生活的训练数据与几百万年的进化结合而成的产物，我们称之为本能。

但是我们正日益接近。近期研究已经表明，成长中的儿童表现出了做许多决策的能力，这些决策推动了许多人工智能系统向前发展。由于数十亿美元的投入，以及肯·海沃思等聪明的研究人员的孜孜工作，意识上传成为现实只是个时间问题。

如果这样，他希望：“发明意识上传相当于发明了盘尼西林。”在会谈将结束时，海沃思对我说：“这必须要做，每个人都会明白，这才是该做的事。我们将难以想象在它做成之前我们曾经在怎样生活。”

第八章 人工智能关乎未来的一切

第八章

人工智能关乎未来的一切

2014年11月，时年43岁的特斯拉汽车和太空探索技术公司首席执行官埃隆·马斯克在未来学网站Edge.org上发出一条在线评论。“人工智能的进步……快得出人意料，”他写道，“除非你直接接触DeepMind这样的团队，否则你根本不知道这种步伐有多快，人工智能正在以近乎指数级的速度增长。未来5年就有可能发生一些非常危险的事情，最多不会超过10年。我并不是在对着我不懂的事情高喊‘狼来了’，也不是只有我一个人在考虑我们应该警惕起来。领先的人工智能公司已经在大力采取措施确保安全。它们意识到了危险，但它们是否可以规划和控制数字超级智能，并能够阻止恶意智能逃进互联网，这一点还有待证实……”

几分钟之后，他删除了这条消息。马斯克有很多特点，而且许多特点都很不错，而这些特点中绝对不包括孤陋寡闻。过去几年间，他领导的这家电动汽车企业对多家机器智能公司进行了投资，包括上文提及的DeepMind，这是一家深度学习公司，在前面章节已经介绍过它的工作内容。马斯克个人资产净值约为112亿美元，他说，他对人工智能进行投资不是为了获得投资回报，而是为了获得消息。“我喜欢密切注意人工智能的发展进程，”他对美国CNBC（全球性财经有线电视卫星新闻台）如是说道，“我认为人工智能可能会造成比较危险的后果。”

埃隆·马斯克并非唯一担心制造会思考的机器可能会造成尚不为人所知的危险的人。著名物理学家斯蒂芬·霍金（Stephen Hawking）是另一位表达保留意见的名人。“人们可以设想这样一种技术，它比金融市场更聪明，发明能力比人类研究人员更强，控制能力比人类领导更强，而且可以开发我们无法理解的武器。”他在2014年5月写道，“人工智能的短期影响由其控制人员决定，而长期影响则由人工智能是否完全被人控制所决定。”

霍金的许多观点都与本书一直讨论的人工智能的发展密切相关。如上文所说，伴随着神经网络的崛起，人工智能在金融市场的应用发生在20世纪80年代。在某些场合，人工智能确实证明了其无与伦比的发明能力，尤其是在处理基因算法时。人类的操控者可以只将重要的任务移交给即将管理我们生活的人工智能助理，与此同时，自人工智能出现之日起，人工智能武器的开发就一直是这个领域的目标之一。

霍金与马斯克着重强调了所谓的通用人工智能（Artificial General Intelligence，简称AGI）。到目前为止，本书中描述的所有人工智能应用都属于“狭义人工智能”或“弱人工智能”。这种称呼与该技术是否强大无关。正如我们在前几章看到的，今天，深度学习神经网络比构成传统人工智能的符号处理人工智能稳定许多

倍。而“狭义”与“广义”或“强”与“弱”之分归结于智能的通用性。人工智能现在可以在很多特定领域打败人类，包括下棋和在电视节目《危险边缘》上回答问题。人工智能的应用范围一直在不断壮大，而且未来数十年内可能会覆盖当前一半的就业领域。

然而，尽管不同于只在小范围内有用的早期人工智能，这种狭义人工智能仍然可以进行扩展，从而在现实世界中发挥作用，但是这种人工智能仍然有明显的局限性，那就是只能用于单一且严格限定的领域。这里有个简单的例子，下棋人工智能虽然可以击败人类象棋大师，但它们在简单的翻译任务中则一无是处。如果你突然要求生产iPhone的机器人去画画，它可能也同样会非常失败。即使是旨在处理多任务的人工智能助手Siri等人工智能，如果你偏离了程序所预设的脚本，它也会变得混乱。尽管这些系统已经获得了惊人的海量数据和计算能力，但是它们缺乏基本的人类特质，比如不能从少量训练例子中总结规律，这就是为什么计算机科学家正在构建更多类生物学大脑算法。

那么通用人工智能中的“通用”有什么含义呢？与狭义、单领域人工智能应用相比，通用智能将表现出更广的、类人的智能。“通用问题解决者”就是这样一款人工智能。1959年赫伯特·西蒙、J. C. 肖（J. C. Shaw）和艾伦·纽维尔（Allen Newell）将一款人工智能软件乐观地命名为“通用问题解决者”，它能够在很多环境中运行，甚至在许多我们没有预测到的环境中运行。一言以蔽之，通用人工智能不是要建造思考的机器，而是要建造比我们人类更聪明的机器。

末日的开端

“当我们建造的机器比我们更聪明、更有能力时会发生什么”这一问题存在的时间比人工智能的存在时间还长。第一个围绕“机器人”的故事是卡雷尔·恰佩克（Karel Čapek）于1920年创作的科幻剧本《罗素姆的万能机器人》。在故事的结尾，人工智能发动了反抗人类的起义，并接管了地球。人类在这个过程中最终被全部消灭。比它还早的是玛丽·雪莱（Mary Shelley）在1818年创作的小说《弗兰肯斯坦》（*Frankenstein*），又称《现代普罗米修斯》（*The Modern Prometheus*）。她在拜伦勋爵的家里消夏时完成了这部小说，小说讲述的是一位年轻的理工科学生的故事，他创造了一个科学生物，这个生物最终变成了杀人狂。“弗兰肯斯坦情结”已经成为一个术语，用于描述人类对于人造产物的恐惧。它后来成了经久不衰的主题，从艾萨克·阿西莫夫（Isaac Asimov）的科幻小说到《侏罗纪公园》（*Jurassic Park*）的作者迈克尔·克莱顿（Michael Crichton）创作的惊险小说，以及最近上映的《机械姬》（*Ex Machina*）中都有它的身影。

真正的科学家不会像科幻小说作者那样非常迅速地考虑这个问题，但他们并没有落后太远。1964年，也就是举办纽约世界博览会的那年，控制论之父诺伯特·维纳预测：“未来的世界将会有一场突破智力局限的更紧迫的斗争，未来没有悠闲的生活，可以让我们躺着等待机器奴隶的服侍。”

维纳于1964年5月辞世，享年69年。然而，对超级智能机器的担忧依然在继续。1965年，英国数学家欧文·约翰·古德（Irving John Good）详细解释了部分担忧。“二战”期间，他与艾伦·图灵在布莱切利公园工作。他在破解纳粹密码方面发挥了重要作用，多年之后，留着小胡子的古德拥有了一辆车牌为007IJG的车，这个车牌仿佛是对他曾经的“绅士特工”岁月的一种戏谑的致敬。1965年，古德撰写了一篇论文，在文章中他阐述了超级智能机器对世界来说意味着什么。他将人工智能定义为能够胜任所有人类才能进行的智能活动的计算机。在一段广为流传的文章中，他写道：“既然机器设计也是一项智能活动，那么超级智能机器应该可以设计出更好的机器。毫无疑问随后将会出现一个‘智能爆炸’，人类的智能可能被远远甩在后面。因此，第一台超级智能机器是人类需要完成的最后一项发明。”

“智能爆炸”已经成为部分研究人员非常赞同的看法。本质而言，它的含义是，当超级智能机器不可避免地被制造出来时，既有的超级智能机器将设计能力更强的机器，要不然就是重新编写自己的软件，从而变得更加智能。这种不断发生的自我完善随后会使得机器的能力发生惊天动地的质变，并会加速这种质变。人类的智能在这个过程中将相形见绌。

古德的论文及其激起的讨论有着很多不确定的地方。因为超级智能机器可以解决我们人类预想中可能会面对的所有问题，所以超级智能机器就是我们需要进行的最后一项发明吗？还是因为它会将我们全部消灭，所以才会成为我们的最后一项发明？古德在此后曾担任斯坦利·库布里克的电影《2001：太空漫游》的顾问，也许我们可以从中找到一些线索。在这部电影里，名为HAL9000的人工智能变成了杀手，开始杀戮它的人类队友。

奇点来临

古德的论文发表后一年，一篇短篇小说发表在1966年3月科幻期刊《模拟科学》（*Analog Science Fiction*）上。《书呆子快跑》这篇小说讲述了一个大脑的故事，由于被直接放进了计算机数据源之中，大脑的能力被人工增强了。这是科幻作家、数学教授及计算机科学家韦尔诺·文奇（Vernor Vinge）出版的第一本小说，小说名字直接源自“漫威漫画”（Marvel Comics）中的头韵玩笑。文奇随后成了一个成功的小说家，但他仍然以1993年的纪实散文《技术奇点即将来临：后人类时代生存指南》而闻名于世。这篇散文叙述了许多古德已经提出的超级智能机器的理念，但增加了时间轴。“30年内，我们就将获得创造超级智能的技术方法，”文奇写道，“之后不久，人类时代就将终结。”

“奇点”（Singularity）这个术语指的是机器在智能方面超过人类的那个点，它已经成了人工智能的一个参考，其被应用的普遍性不亚于“图灵测试”。通常人们认为这是文奇的功劳，但实际上使用这个概念的第一位计算机科学家是约翰·冯·诺依曼。在冯·诺依曼人生的最后10年里，他和波兰裔美国数学家斯坦·犹拉姆（Stan Ulam）有过一次会谈，他们共同参与了“曼哈顿计划”（Manhattan Project）。犹拉姆随后在回忆这次对话时指出，冯·诺依曼曾痴迷或者说担心“不断加快的技术发展与人类生活模式的变化，这看起来似乎接近了人类历史上某个必要奇点，到了那个时候，就我们所知，与人类有关的事务会难以为继”。

和古德一样，韦尔诺·文奇在其1993年的文章里并没有得出明确的结论。如果出现奇点，他承认它的影响可能是好的，也可能是坏的。“从某个角度看，其前景符合我们很多最幸福的梦想。”他写道：“（它很可能）营造一个无限美好的地方，让我们能够真正了解彼此，并了解最神秘的事物。但是从另一个角度，它也很可能造成最糟糕的情况。”

这是奇点这个术语如此贴切的一个原因。在将这个词用在人工智能方面前，“奇点”通常用于理论物理，用来描述黑洞中央的重力中心：在这个点上，物质自行坍塌。技术奇点就像一个黑洞，对人类思维而言完全深不可测。

由于这个原因，推测通用人工智能可能将我们引向何方成了一件非常有趣的事情，但是这个话题最终沦为科幻小说的素材。这有点像汽车之父亨利·福特的揶揄：在汽车诞生之前询问人们想要什么，最普遍的回答可能是“快马”。就像人们在曙光来临前围坐一团，猜测创造一种语言会将我们带向何处时，他们绝不可能想到推特标签的细微之处。因此，我们现在也想不到超级智能将如何看待世界，但是，毫无疑问它将改变世界。

狭义与广义的区别

我们所看过的所有科幻电影和书籍将一个预期深深地植入我们的脑海中，让我们相信某个奇点式的“临界点”最终将会到来，到时候，通用人工智能就会出现。设备将越来越智能，直至有一天，在硅谷一间秘密研究实验室的深处，一条消息突然出现在马克·扎克伯格或谢尔盖·布林的电脑显示器上，该消息称通用人工智能已经实现。就像厄内斯特·海明威（Ernest Hemingway）写过的关于破产的内容，通用人工智能先是逐渐地酝酿，然后就突然诞生了。这是电影惯用的叙事手法，如詹姆斯·卡梅隆（James Cameron）的力作《终结者2：审判日》（*Terminator 2: Judgment Day*）里描绘的那样。在这部影片里，观众获悉超级计算机“天网”于美国东部时间1997年8月4日凌晨2点14分具有了自我意识。当天凌晨2点13分的时候，计算机用户可能都对自己搜索引擎的搜索结果史无前例的精确而感到惊奇，或者对《命令与征服之红色警报》（*Command and Conquer: Red Alert*）里人工智能采取的超级战略感到惊奇（嗨，这是1997年！）。2点15分，喀——嘭——！生命终结了！

在摩尔定律主导的世界里，计算力量的发展像钟表一样可以预测，超级智能也很难挣脱这种预测。就像苹果的新一代iPhone的发布，人人都想知道预计的发布日期。比如，上一章里所提及的雷·库兹韦尔就预测奇点会在2045年出现。

库兹韦尔之于奇点就像史蒂夫·乔布斯之于智能手机：他们都不是第一个想到创意的人，但却是推广它的人。作为11家公司（包括为Siri提供语音技术支持的人工智能公司Nuance）的创始人，库兹韦尔一直被不亚于比尔·盖茨的权威人士赞为“预测人工智能未来的最佳人选”。《福布斯》杂志的称赞甚至更上一层楼，它称库兹韦尔是“托马斯·爱迪生的后继者”，甚至是“终极思考的机器”。与悲观主义者截然相反，库兹韦尔认为奇点对人类而言绝对是有益的：技术人员会经历《圣经》中所形容的狂喜，在这种狂喜之中，一切问题都迎刃而解，而且我们所有人，甚至是技术大亨，也永远都不必再运用自身的智慧了。

但是，并非所有人都相信奇点会如此神奇。正如艾伦·图灵通过图灵测试所指出的那样，“机器能否思考”这一问题毫无意义，因为几乎不可能评估出任何肯定的结果。在上一章里，我们可以发现，认为随着计算机的速度越来越快，计算机自然而然地会产生意识的看法有些过于简单了。考虑到区分弱人工智能与强人工智能的难度，一些人错误地认为，在弱人工智能中，人工智能得到的结果是预编程后的结果，是用一种算法执行特定的一系列步骤进而获得一个可预知的结果。这就意味着，如果训练过程执行恰当，人工智能基本不可能产生不可预知的结果。然而，基因算法可能生成我们意料之外的解决方案。程序员以“目标函数”的方式设立了算法目标，但并不完全知道计算机如何实现这个目标。这同样适用于人工智能为了实现目标而创建的种种策略，和第三章所讨论的强化学习领域一样。在两个场合里，人类创建者都无法按部就班地预测人工智能的“自发”行为。

当涉及意识的迹象时，一切都变得更加复杂了。举例来说，如上一章所述，如果在计算机内部成功复制线虫的神经系统，这是否代表实现了通用人工智能？尽管这种突破可能产生种种改进现有机器学习工具的见

解，但或许回答仍是：这不能代表实现了通用人工智能。线虫的行为中可以被视为智能行为的相对较少。处于食物链高层的动物也是如此。尽管我们坚持认为，目前的狭义人工智能仅能在严格指定的领域中运行，但是这一点同样可以适用于几乎所有生命形式。蜜蜂可以建筑蜂巢，但不能建造水坝或土堆；海狸可以建造水坝但不能建造蜂巢和土堆；白蚁能建造土堆却不会建造蜂巢和水坝。到目前为止，在所有动物中，人类拥有的技能最通用，但我们仍有某些更为擅长的行为和较不擅长的行为。如果我们构建一个单用途人工智能，这个人工智能在一个领域中拥有“意识”和推理能力，而在其他领域中没有“意识”和推理能力，那是否实现了通用人工智能？鉴于今天的神经网络越来越擅长执行感知任务，却仍然无法理解伦理等话题，这一假设十分合理。

即使假设库兹韦尔关于指数级增长的理论继续适用，那么所有事物也不可能一次同时出现，更不用说制定准确的时间表了。

“他对最近几年发展情况的预测能力令我大为震惊。”当我向杰夫·辛顿问及他对库兹韦尔的奇点理论的看法时，他略微迟疑了一下，随后说：“这是个讽刺。”

“看透未来就像看穿迷雾，”辛顿继续说道，“当你置身迷雾中，咫尺距离可能看得非常清楚。如果你看向稍远的地方，就会有些模糊。但如果你再将目光放得更远，你根本什么都看不到。这是因为看透迷雾的难度呈指数级增长。每穿透一段距离的迷雾都将失去一定比例的光线。”那么，这就意味着我们完全找错目标了吗？并不全是。“我们的技术将越来越好。”他说，“我看不到有任何理由表明生物大脑必须成为终极的思考机器。我认为，最终他们能够设计出比自身好的东西。那么，问题就成了：政策是什么以及人们决定要用技术去做什么。如果国家掌权人希望建造入侵小国家的机器人杀手，从而不牺牲任何美国人，那么这就是我们要面对的结果。”

但这不只是人工智能在遥远未来的应用，或者通用人工智能的发展，它现在就带来了挑战。

人工智能带来的风险

2012年4月，罗科·迪乔治（Rocco DiGiorgio）下班回家后发现他房子里的味道非常难闻。屋里到处都是狗屎，摊得很薄，但均匀地就像恐怖的蛋糕表面涂层。虽然迪乔治最初还不知道如何解释这一现象，但是他突然想到了原因。就在他的清洁机器真空吸尘器准备开始当天的例行打扫工作之前，他的宠物狗把屋里搞得一团糟。按照指令，机器人Roomba检测到了狗屎，并来回打扫多次，试图将它清理掉，然后在机器人打扫整个屋子时，把狗屎拖得满屋都是。“那时我一点都高兴不起来。”在YouTube视频上迪乔治痛苦地说，这段视频引起了著名社交网站Reddit用户的关注之后，就像病毒一样传开了。

迪乔治的故事难以代表我们在本章一直讲述的这种潜在的人工智能风险。这与人工智能获取全球核武器控制权（如电影《终结者》）或将我们的大脑封锁在巨大的模拟环境中（如电影《黑客帝国》）的情景相去甚远。然而，它证明了人工智能的另一个侧面：人工愚蠢（“人工智能”的反义词，指显得自己不那么聪明的电脑程序）带来的风险可能与真正的人工智能带来的一样大。简而言之，我们有时渐渐心安理得地希望让人工智能系统做决策，而它们却不一定有相应的智能。

有些人认为高级人工智能的出现可能意味着人类的灭亡，他们所热衷的一个思维实验即所谓的“回形针最多化”（paperclip maximiser）情景。这个思维实验由瑞典哲学家和计算神经科学家尼克·博斯特罗姆（Nick Bostrom）提出，在这个情景中，人工智能被要求实现一个看似无害的目标：经营一家生产回形针的工厂。这个人工智能可以利用纳米技术在分子层面上重构事物，它接到以最大效率生产回形针的任务后，悲剧性地首先将地球变成了回形针，然后将大部分可见的宇宙也变成了回形针。

“回形针最多化”情景非常普通，尽管对于我来说，它看起来更像是人工愚蠢而不是人工智能带来的问题。对“已经没有纸了，为什么还要生产回形针”或者“首先提出回形针需要的人已经被变成回形针了，为什么还在生产回形针”这类问题无能为力的情形，不太可能发生在高级的超级智能上，除非回形针有什么我不知道的极端重要的性质。恰恰相反，威胁来自人工智能，人工智能足够聪明，可以与其他联网设备共事，但还不足以对自己的工作动机提出质疑。

事实上，像把狗狗大便抹得到处都是的机器人Roomba一样，基于规则的简单人工智能有着大量失败的例子。2011年4月初，亚马逊上一本不再发行的图书《苍蝇的成长》出乎意料地与自己打起了价格战。这本书有两个版本上市，售价通常为35—40美元。然而，突然有一天，它们突然开始分别以1 730 045和2 198 177美元销售。仅仅数小时后，它们的售价已经分别达到了2 194 443和2 788 233美元，随后又涨至2 783 493和3 536 675美元。两周以后，价格达到创纪录的23 698 655.93美元，含运费。为什么会这样？因为算法规定这本书的售价要略高于其竞争对手的售价，于是引发了一场价格战，并最终将自己的价格提升至一个荒谬的程度，尽管事实上没有人会支付这样的价格。幸运的是，除了给亚马逊带来了些负面宣传，这种情况没有造成任何伤害。

一个更值得注意的人工智能造成破坏的案例发生在2010年5月6日，这本应是平平常常的一天，但这天，

将近1万亿美元的财富消失在数字苍穹之中。美国东海岸时间下午2点42分，道琼斯工业平均指数三分钟之内下跌了将近1 000点，创下历史上最大的单日跌幅。一些股票从平时的30—40美元的交易区间暴跌至0.01美元，不过随即反弹。苹果公司股票每股从250美元激增至100 000美元。所幸，这种罕见的“闪电崩盘”没有再度上演过，但它几乎肯定是基于规则的简单人工智能的反馈回路被锁定的后果。事实就是人工愚蠢合法地从所有人的口袋中成功“偷走”的钱比历史上预谋最周详的人抢走的最多的钱还要多。

“黑箱”风险

无论是超级智能还是人工愚蠢，都有许多因素导致在出现流氓人工智能的情况下难以进行人为干预。首先是它们的运行速度。已经有益于自动执行股票交易等指令的人工智能系统，执行时间以纳秒为单位。由于这些交易的发生速度很快，一旦出现问题，人类根本不可能进行实时干预。

更重要的是，当今许多人工智能工具中都存在“黑箱”般的不透明性。至于先进神经网络与基因算法的情况，它们的人类操作者很久以前就不再试图理解其有效执行复杂任务的能力。这为仔细检查人工智能增加了难度。尼克·博斯特罗姆与其研究员同事埃利泽·尤德考斯基（Eliezer Yudkowsky）此前曾阐述一个假设情景，使用机器学习算法提供是接受还是拒绝房屋抵押贷款申请的建议。他们建议，如果申请人的抵押贷款被拒绝，申请人可以向银行申诉，声称人工智能因种族而歧视部分申请人。而银行则告知申请人，情况并非如此，且算法无法知道特定申请人的种族。但无论如何，在审查神经网络决策制定流程的结果时，人们发现黑人申请的批准率大大低于白人申请的批准率。

有许多原因可以解释这一现象，但是博斯特罗姆和尤德考斯基认为，很难确定具体原因。如果任务采用的是一种简单专家系统，可能很容易就能说明这个情况，比如，抵押贷款顾问人工智能的部分决策取决于申请人的当前地址，住在贫困地区的申请人的贷款违约率较高。

人工智能执行大量工作的具体方法加剧了这一问题。在21世纪的第二个10年里，技术公司的影响不再与其规模成比例。2012年4月Facebook花10亿美元收购的Instagram，其员工名单上仅有13人。与之相比，摄影巨头柯达在前数码时代其地位与Instagram旗鼓相当，其全盛时期员工人数超过14万。20世纪工业巨头的规模使它们更易于被监管。这同样适用于20世纪威胁公众的风险源，如核技术。这样的研究领域需要实体场址，建筑大规模设施以及大量资金。然而，时下人工智能的最大投资者：谷歌、Facebook和苹果可能会雇用数千人才，并在硅谷拥有大量大学式园区，但这已经不再是必要条件。当今计算设备的能力意味着，只要拥有必需的编程技术、一台个人电脑、笔记本甚至是智能手机，任何人都可以在构建人工智能的项目中发挥重要作用。完全不需要去与飞机库一样大的大型公司总部，任何有好想法的人都可以在自己学校的寝室，甚至是在苹果和谷歌创始人创业的改造车库里，构建人工智能系统。

尽管缺乏数百万美元支持的兄弟公司所拥有的资源，但是开源人工智能项目可以帮助塑造未来的人工智能。在线开源机器学习库的数量不断增长，全球用户定期对这些学习库进行更新。比如，自从2010年2月对公众开放以来，开源学习库scikit-learn已经修改了18 000多次。2015年，普通的一天中，有8名用户对scikit-learn代码进行了18次修改。而这些进行讨论的用户远在瑞士、法国、美国和印度。一些开源人工智能项目设法制定一些相对温和的目标，如想出极客式家庭自动化项目。其他开源人工智能项目则致力于通用人工智能。

“今天，没有什么挑战比创造有益的通用人工智能更加重要，通用人工智能具有与人类相当的广泛能力并且将最终超越人类。”OpenCog网站上如是写道。OpenCog是个开源软件创意，自称“直面”构建通用人工智

能的挑战。

随着人工智能处理的任务量不断增长，这些问题也将愈加迫切。就像本书所描述的那样，现在的人工智能用于帮助设计新城市、监控银行账户安全、执行具有重大经济影响的金融交易以及驾驶汽车。在不久的将来，即使你听说美国总统这样重量级的人物乘坐人工智能驱动的汽车，也不必吃惊。

谁又能说清楚明天它们将用于什么领域呢？

我们无法起诉机器人

Siri发布前两年，英国人工智能资深专家尤里克·威尔克斯（Yorick Wilks）发表了一篇名为“人工智能同伴导论”（*Introducing Artificial Companions*）的文章，讲述了第四章讨论的这种技术。在文中，威尔克斯简短地回答了责任问题。他特别感兴趣的是，智能增长是否伴随着责任增长。威尔克斯写道，假设人工智能助手告诉你的祖母，外面很暖和，但当她听到这个消息后走进的却是冰冷的花园，然后她感冒了，那么这种情况下谁应受到责备？此时此刻，连威尔克斯这位毕生都在询问和回答未来问题的人都得承认，这个问题对于读者“似乎有些超出想象”。仅仅过了几年，2015年年初，在牛津，我与尤里克·威尔克斯在他家附近共进午餐。他驾驶着自己的小型汽车在火车站与我会合，我们在附近一家意大利餐厅品红酒，讨论人工智能助手的话题，愉快地度过了几个小时。午餐期间，我的头脑中一直在感叹着过去几年人工智能的发展是如何迅速。一个2009年看起来像是科幻小说情节的哲学难题现在几乎成了现实。

我们对人工智能助手寄予了极大的信任，有时甚至超过了我们的本能与判断。Siri最早的电视插播广告由佐伊·丹斯切尔（Zooey Deschanel）主演，尽管当时窗外明显地下着大雨，但是她仍然看着窗外问Siri：“在下雨吗？”幸运的是，当时Siri的答复非常肯定。而在其他场合，人们远没有这么幸运。2013年年底，阿拉斯加州费尔班克斯的几位iPhone用户按照使用苹果地图应用数据的Siri的建议，选择了一条抵达费尔班克斯国际机场的路线，但这条路线非常危险地穿越了飞机起降的跑道。机场的营销主管安吉·斯皮尔（Angie Spear）称这是个很明显的证据，说明司机太过于相信导航系统。“无论标志牌上写的是是什么，只要iPhone上的地图告诉他们继续沿这条路前进，他们就前进。”她说。

在处理技术引发的此类交通事故中，当前可用的法案条例非常有限，不过以往法院一直趋向做出针对人类的判决。比如，2009年，一位英国司机依据GPS（全球定位系统）的导航，却令自己困在了悬崖边的狭窄小路上。这名司机最终不得不被警察拖回主路。尽管英国法庭将之归因于他的GPS，但还是认定其有大意驾驶的过错。

法官倾向认定人类对此负责的原因是，我们习惯于一个充斥非智能工具的世界。举一个简单的例子，一个人用枪杀了另一个人，他应为自己的罪行负责，而不是让他的枪负责。同样，一家公司销售了有瑕疵的枪支，购买者在使用的时候枪支发生了爆炸，枪支公司应对所造成的所有伤害负责。这种思维也持续到了数字时代。1984年，也就是苹果公司首次推出麦金塔电脑（Macintosh，简称Mac）的这一年，美国一家名为阿斯隆工业的公司老板发现自己成了法庭的被告，他被指控使自己的客户置身危险之中。阿斯隆公司是一家销售自动投球机的公司，设备主要用于供人进行棒球挥棒练习。不幸的是，其中一些机器被证明存在缺陷。这些凶猛的机器发射的棒球极不规律，而且速度极快，有可能使某些不走运的客户颅骨骨折，甚至失明。尽管阿斯隆案件中没什么特别值得关注的事，但这个案件却令人难忘，因为法官宣布必须对阿斯隆工业而不是投球机进行起诉。原因是什么呢？因为“我们无法起诉机器人”。

然而，现代人工智能的主要不同点在于，人工智能不再仅限于由人类使用，而是一种由人类部署的工

具。一旦部署，在多数情况下，机器就会按照所收集、分析和最终用于做出决定的信息，独立于指令而自行运转。

戴维·弗拉杰克（David Vladeck）是华盛顿乔治城大学的法学教授，他是一位少有的致力于研究人工智能法律责任问题的法律专家。和多数人一样，他认为陈旧的“我们无法起诉机器人”的观念可能需要改变。弗拉杰克的一个思维实验涉及《2001：太空漫游》中HAL9000超级计算机蓄意杀人案件。

对那些没有看过这部电影的人来说，这是一个非常简单的情节：HAL9000是全知的人工智能，控制着“发现一号”宇宙飞船的运行，飞船载着5名船员飞往木星。在电影的开始部分，HAL9000被称为“简单且不会出错”。然而，问题很快出现了。HAL9000希望知道“发现一号”正执行的绝密任务细节。船员戴夫（Dave）和弗兰克（Frank）拒绝透露这些信息，尽管事实证明他们也不知道所有细节。不久，HAL9000警告称“发现一号”舱外一个重要天线将要失去作用。戴夫和弗兰克开始怀疑HAL9000并把自己锁在了一个挖掘工具里讨论这件事，他们认为HAL9000不可能听到他们的对话。他们决定修复天线，还决定如果事实证明HAL9000出现了故障就关闭HAL9000。戴夫和弗兰克没有意识到HAL9000能够利用图像识别技术解读唇语。当弗兰克涉险到舱外检查天线时，HAL9000割断了他的氧气管并使他飘浮进了太空。戴夫徒劳地想挽救弗兰克，但当他也走到舱外的时候，HAL9000趁机关掉了其他三位正在睡眠中的船员的生命维持系统，杀死了他们。HAL9000随后拒绝让戴夫返回舱内，因为它知道戴夫会将它停用，而HAL9000认为这将危及整个任务。尽管如此，戴夫最终还是设法回到了“发现一号”飞船内，并且准备关闭HAL9000。HAL9000恳求戴夫停手，在最后的时刻，作为一个自主人工智能，它表达出了自己对死亡的恐惧。

弗拉杰克质疑HAL9000是否能够在法庭上对自己的行为负责。假设戴夫于2001年年底重新回到了地球，与一位人身伤害律师进行接洽（“你是否无故卷入了人工智能杀人案件之中？”），并要求有人对自己的遭遇和4位被HAL9000无故杀害的船员进行赔偿。至少按照现有法律，HAL9000可能会摆脱困境。与阿斯隆工业的失常棒球抛射机器一样，法庭处理这种案件时，可能承认机器造成了伤害，但会认定这是HAL9000程序员的错误而不是HAL9000的错误。

这可能不公平。正如弗拉杰克所说：“仅从可能参与设计、修改和集成人工智能系统元件的个人与公司的庞大数量看，就很难确认大部分责任人或责任方。许多元件早在人工智能项目构思之前许多年就已经设计出来了，这些元件的设计人员可能从未设想，或鲜少预期到他们的设计会集成到人工智能系统中，更不用说会造成伤害的某个具体的人工智能系统了。在这种情况下，指责在时间和地理位置上都与人工智能系统完成和运行相距甚远的元件设计人员有工作失误是有失公允的。法庭可能也不愿意说这种元件设计人员本该预见到这种伤害的出现。”

机器人学的三大法则

人工智能的潜在危险、我们对其日渐依赖以及法律的灰色地带共同构成了非常重要的问题：我们需要在这个领域建立一种伦理规范。简而言之，如果我们要建造思考的机器，难道我们不应努力使其成为考虑周到的机器吗？然而它仍然与科幻小说有着难解难分的关系，因而也就不必奇怪在科幻作家艾萨克·阿西莫夫的作品中有许多在人工智能中植入道德感的著名例子。阿西莫夫的“机器人学的三大法则”首次出现在其1942年的短篇小说《环舞》（*Runaround*）中，这篇小说最初发表于《惊世科幻》（*Astounding Science Fiction*）三月刊中。这三大法则通常简称为“阿西莫夫法则”，常为人们引用的法则内容如下：

1. 机器人不得伤害人，也不得见人受到伤害而袖手旁观。
2. 机器人应服从人的一切命令，但不得违反第一法则。
3. 机器人应保护自身的安全，但不得违反第一、第二法则。

尽管已经过去70多年了，但“阿西莫夫法则”依然统管着阿西莫夫的科幻小说同行们。2014年1月，当谷歌收购深度学习公司DeepMind时，一切发生了重大变化。作为交易的一部分，谷歌必须成立人工智能伦理委员会，目标是确保对人工智能这项技术的明智使用。尽管关于该委员会构成的公开细节寥寥无几，但是建立这样一个防护措施是一个重要的基准。2015年夏季，DeepMind应用人工智能主管穆斯塔法·苏莱曼（Mustafa Suleyman）承认，在这个领域公众看待人工智能的方式已经发生了转变。“故事已经从‘人工智能如此失败不令人恐怖吗？’转变至‘人工智能如此成功不令人恐怖吗？’”在一场深度学习大会上，他曾如是说道。

谷歌不是唯一认为必须确保人工智能承担责任的研究团体。由于过去10年间深度学习与其他人工智能的统计形式已经成为规范，纽约伦斯勒理工学院认知科学系主任塞尔默·布林斯乔德（Selmer Bringsjord）发起了一场一个人的战争，支持回归到从头到尾基于逻辑的人工智能。“我不做任何统计学上的事情，”在描述他在实验室的工作时，他告诉我，“我非常不喜欢它，认为它正将我们引入歧途。”

对于布林斯乔德而言，之所以需要回归从头到尾有逻辑的人工智能，是因为依靠“数学上令人费解”的统计工具本质上是愚蠢的行为。“你希望系统能够阐述其行为的依据和正确性吗？”他说：“我们应该希望更加复杂的人工智能系统能够证明，它们根据获得的输入，其所作所为是正确的决定。”

有逻辑的人工智能可能意味着《连线》和《快公司》的篇章中那种引人注目的技术推出过程应该放慢，但布林斯乔德认为，创造具有清晰推理过程的人工智能是最重要的，因为这个过程事后可以分析。近斯关于有逻辑的人工智能的一个例子出现在位于纽约哈德逊河东岸的布林斯乔德实验室中。这个展示中，一台计算机尽力破解“三个智者”出的难题，实验由三个小型类人机器人辅助进行：这一实验可能意味着人工智能自我意识的萌芽。在这个难题中，三个机器人被告知，两个机器人服用了停止说话的“哑药”。它们的任务是找出其中仍然可以说话的那个机器人。三个机器人都试图说“我不知道”，但只有一个实际发出声音。当它听到自己的机器声音时，这个机器人明白它就是那个没有被禁止发声的机器人。“对不起，我现在知道了。”它

说，“我能证明我没有吃哑药。”随后它写下了关于这个难题的有条理的数学证明，并把它保存起来。布林斯乔德进行了相同测试的不同版本，或者任何旨在解决其他哲学难题的测试，他认为这些测试将形成一个人工智能可以利用的日益完善的技巧或功能指令系统。

由于围绕用于战争的自动人工智能武器进行的工作越来越多，人们迫切需要像布林斯乔德这样的工作项目。2014年，在塔夫茨大学和布朗大学成立了一个多学科综合小组，小组和布林斯乔德合作，获得了海军研究办公室的资助，去探索能否为可能用于战争的自主机器人赋予是非感。尽管这种探索得到的不是一个“友好的人工智能”，但它可能会使机器人在战场上做出符合伦理的决定。设想一下，比如，一个正将伤兵运往战地医院的机器人军医遇到另一个腿上受伤的士兵。权衡停止执行救助任务的利弊，能否在战场采用牵引疗法止疼以及其他难题，对于人类来说已然是个难以驾驭的复杂问题了，更不用说对于机器了。

这样的问题可能越来越成为常态。设想一下，如果一个生产自动驾驶汽车的公司决定，当其生产的车辆检测到前方即将发生碰撞时，它的车辆可以猛地驶出道路，从而保护驾驶员，这将会发生什么情况呢？这一操作非常合理，如果是人类在驾驶车辆，多数人也会做出这样的决定。然而，如果你的车正停在红灯前，但这时它发现后面的一辆车快速驶来，这时将会怎样？你的车知道可能要发生追尾事故，随后它就驶出了马路……并正好冲进了一群放学回家的学生当中。司机避开了一个小追尾和潜在的会扭伤脖子的事故，但结果却是三个学生死于非命、两人受伤。即使最忠诚的车迷也难以为这种成本和收益的互换进行辩护。而这些正是人工智能研究人员开始认真对待的问题。

人工智能的权利

随着我们将越来越多的工作交给人工智能，我们可能最后需要面对人工智能的权利问题。许多作家已经提出了这个问题，但是在主流目光集中到这一点之前，它依然远落后于人工智能是否会为人类带来危险这个问题。举例来说，英国技术作家比尔·汤普森（Bill Thompson）因其每周的BBC新闻专栏而闻名，他认为将艾萨克·阿西莫夫的第一法则（机器人不得伤害人，也不得见人受到伤害而袖手旁观）写进超级人工智能的代码中有些类似于“为奴隶戴上枷锁或把大猩猩装进笼子，它也折射出我们的信念，即人工智能应该且必须永远服务于人类，而不能成为一种自主的心灵”。如果有人提议将这一条法则作为控制未来人工智能的政策，他认为：“我们在道义上有责任反对它。”

就像人类有与其人工智能助手结婚的权利，我不认为人工智能的公民权利在未来10年内会成为一个大众讨论的话题，但它是个值得考虑的话题。就像如果我们建立一个具有基本意识的生物仿真神经网络，那么强迫它为我们开车或优化搜索排名是对的吗？这使人联想起道格拉斯·亚当斯著名小说《宇宙尽头的餐馆》

（*The Restaurant at the End of the Universe*）中的场景，在书中他描述了一个未来主义的电梯：天狼星控制系统公司的快乐人体垂直传送器。读者获悉，将今天的电动卷扬机驱动电梯与这种人工智能控制的电梯相比，就像“将一包混合坚果与天狼星国立精神病医院的整个西翼大楼相提并论”。天狼星控制系统公司快乐人体垂直传送器有点像第三章说的智能物联网设备，通过预测未来而工作。它甚至在你意识到自己要进入电梯之前，就会出现在正确的楼层去接你，“这样一来，也就免去了等电梯时人们被迫要进行的所有那些单调乏味的交谈、放松以及相互结识”。

但是，种智能电梯很快就厌倦了这种上上下下，产生了斜向运行的念头并付诸试验。道格拉斯·亚当斯将之称为“存在主义式的抗议”。讨论囚禁人工智能是否符合伦理这可能看起来有些可笑，但如果科学家成功开发出了一种人工智能，其行为与我们主张的以符合伦理的方式对待的真正生命体更为类似，那么这种讨论就不会那么可笑了。多年来，许多人工智能标准都已经发生了改变，我们对于“什么可以称之为意识”的看法也发生了变化。举一个著名的例子，法国哲学家、数学家勒内·笛卡儿曾经描写了他认为在建造一种机器时会遇到的挑战，而这种机器就是我们今天称为人工智能的机器。在1637年发表的《谈谈方法》（*Discourse on the Method*）中，笛卡儿写道：

如果有一些机器跟我们的身体一模一样，并且尽可能不走样地模仿着我们的动作，我们还是两条非常可靠的标准来判明它们并不因此就是真正的人。第一条是：……它决不能把这些字排成不同的组合以恰当地和对答，而这是最愚蠢的人都能办到的。第二条是：那些机器虽然可以做许多事情，做得跟我们每个人做的一样好，甚至更好，却决不能做别的事情。从这一点可以看出，它们的活动所依靠的并不是认识。

笛卡儿认为在有灵魂的思想者与一个没有灵魂却栩栩如生的非思想机器之间存在重大区别。“我思故我在。”是笛卡儿著名的人类自我辩解。对于笛卡儿来说，在灵魂尺度上，动物多少表现得有些糟糕。因此，

一些笛卡儿的追随者决定四处踢狗。他们的论据是，由于无法达到“最迟钝的人”的思考水平，狗没有灵魂，因此它们的吠叫只是简单的机械反应。

这个标准即使用于人类自身，也存在诸多问题。你只需要看看理论上关于子宫里的胎儿或靠生命维持技术而生存的脑死亡患者是否有生命的辩论有多激烈，你就会明白这个主题有多大的不确定性。人工智能是否有权追求生活、自由和幸福的问题并不是一个需要我们现在就回答的问题，但如同向计算机植入道德，或考虑如何解决潜在的危险软件一样，总有一天我们要面对这个问题。

就像奇点一样，预测即将来临的事物并不像我们想象的那么简单。

结论 当机器变得更加智能

结论

当机器变得更加智能

2016年1月24日，星期天，人工智能先驱马文·明斯基因脑溢血辞世，享年88岁。他是当年达特茅斯人工智能大会组织者中最后一位离世者，约翰·麦卡锡2011年去世，纳森内尔·罗彻斯特（Nathaniel Rochester）和克劳德·香农于2001年去世。报纸随后发表对明斯基所做工作的赞颂——“他证明了向计算机传授常识推理的可能性，为人工智能领域打下了基础”。《连线》杂志一改传统风格，决定发表由人工智能创业公司Automated Insights开发的人工智能编辑编写的讣告，效果非常好。

明斯基的去世具有象征意义，它宣告了第一代欣然认为自己从事的是人工智能工作的研究人员的终结。但是，当他的死讯传到博客和技术论坛后，人们认为他远不是过去年代一个落满灰尘的遗迹。自明斯基和他挑选的一群雄心勃勃的年轻计算机科学家聚集在新英格兰大学校园，整整一个夏季致力于解决机器智能问题的那一年到2016年，已经整整60年。今天，这种乐观主义在我们看来也许是非常天真的行为，但我们不能否认他付诸行动的动力。

我们可能还没有达到人工智能中产生强人工智能的临界点，但是我们不能忽视人工智能取得的成就。一些成就就是明证，不管是人工智能打败象棋世界冠军还是在《危险边缘》节目中打败人类大脑。然而，人工智能在很多其他方面也发挥了重要作用，比如发现新药、传播全球人类可获取的有用信息，使机器翻译越来越快和越来越容易，等等。明斯基可能在人们真正利用神经网络之前过早地摒弃了它，但他的其他想法仍然广为流传。20世纪80年代中期，他出版了《思维的社会》（*The Society of Mind*）一书，认为“智能不是任何单一机制的产品，而是源自各种资源丰富的实体的互动”。如我们在第三章所看到的内容，这种想法正推动着Jawbone和Nest Labs等智能设备制造者的工作：不仅专注于创造独立的智能小器械，还要使整个物联网能够共同工作从而实现目标。

技术公司如今成为人工智能研究的主要投资者，它们之间的竞争比以往更加激烈。马文·明斯基去世的那一周，Facebook的马克·扎克伯格在其拥有15.5亿用户的社交网站上发了一条新闻链接，该新闻称人工智能试图破解围棋（围棋是一种中国棋盘游戏，目的是比对手占据更多的地盘）。围棋的规则容易学习，但是棋盘上可落子的位置数目却非常惊人，比宇宙中原子的总数量还多。“20年来，科学家一直努力教计算机在围棋中获胜，”扎克伯格写道，“我们已经接近目标，过去6个月，我们已经制造了一个人工智能，它仅用0.1秒就可以落子，而且其性能和之前花费多年建立的系统一样好。”对于Facebook来说，这是值得炫耀的事，尽管这个纪录并没保持太久。仅10小时后，谷歌宣布DeepMind制造了一个人工智能，它不仅打败了所有既有

的围棋程序，还首次打败了人类职业级棋手。从那时起，一切都突飞猛进。直至2016年3月，世界一流围棋棋手李世石（Lee Sedol）在韩国一家酒店的房间迎战谷歌的阿尔法围棋（AlphaGo），全球6 000万人观看了这场比赛。在比赛的最后，阿尔法围棋以四比一击败李世石。

当然，并不是所有人工智能引起的变化都是美好的。未来几年，人工智能也将成为取代许多职业和摧毁人类谋生之道的“罪魁祸首”，尽管人工智能也将创造新的、以前人类劳动者不曾想到的工作机会。许多人也会对在战争中使用机器智能大加批评，不管是航空无人机还是陆地机器人战士。后者的典型代表是谷歌的“大狗”（Big Dog），“大狗”是四条腿的机器狗，能够运载约400磅的装备，但是因为它的汽油动力引擎噪声太大，美国海军陆战队当前推迟了“大狗”在战争中的使用。除了一些灾难性的风险，对于大多数人而言，最迫切的人工智能问题是对隐私的攻击，这种攻击伴随着谷歌这样的实体公司的崛起而产生。由神经网络驱动的会思考的机器需要进行数据训练，每当我们使用如谷歌搜索、Siri或其他人工智能工具的服务时，我们都使这些机器变得更加智能。

1956年，人工智能以一门学科的形式出现，而今天的人工智能不再像当年那样轮廓清晰。即使是当年，研究人员也在竭尽全力控制人工智能的应用范围，努力基于诸多不同研究兴趣，创造出一个有机的整体。但是，在2016年这仍是完全不可能的。像谷歌这样的搜索巨头，其主要收入来源是以人工智能系统换取广告收益，而研究人员的目的是通过人工智能去理解人类的大脑，对这两者要如何进行协调？除了相关的技术外，什么能够将无人驾驶汽车与像搜索引擎巨头一样把我们进行分门别类的人脸识别安全系统结合在一起？

我在撰写本书结论部分时网上还流传着两个故事，它们证明了人工智能今天所呈现的分化。（我对谷歌快讯进行了设定，让其不断扫描互联网，从而获得人工智能的相关信息。就这样，我自然而然地发现了这两个故事。）第一个故事讲的是近期的一场辩论。在这个故事里，一群直言不讳、心怀忧虑的科学家和武器专家对机器人的风险提出了警告，这些机器人配备了人工智能，正在设法走向战场屠杀人类。

第二个故事从表面来看相对轻松些，讲述了居住在敦提的苏格兰计算机程序员安迪·赫德（Andy Herd）的工作，他进行了一个创新的人工智能实验，为20世纪90年代情景喜剧《老友记》（*Friends*）创作脚本。“我把《老友记》的每一集脚本都输入了一个循环神经网络，然后让计算机学习创作新的剧本。”赫德在推特中写道。这与我讲述的谷歌Deep Dream工程的创造性工作有些类似，结果有些奇怪。就像《野兽日报》（*Daily Beast*）一篇文章中指出的那样，赫德承认他的软件仍旧需要完善，从而使他的计算机不再创作出奇怪的剧本，比如计算机创作的一集剧本中，所有的演员都兴奋地挤在一张床上，而莫妮卡（Monica）不知道向谁大喊：“胆小鬼，鲍勃（Bob）！”

这两个故事之间的鸿沟，不仅凸显了我们对通用人工智能的日益高涨的痴迷，同时也表明了如今人工智能的广泛应用范围。半个多世纪以前，人工智能解决了下棋计算机的难题，成为我们如何在机器内部建立人工智能这个大问题的缩影。今天的目标更加模糊。

人工智能是关乎机器翻译、图像识别、自动驾驶汽车、能够管理我们生活的智能助手、能与你的智能电视进行对话的智能恒温器、用于设计美国国家航空航天局卫星的基因算法吗？人工智能是关乎未来就业、人

类在21世纪中的角色，或者构建比我们自身更加智能事物的内在风险吗？我们在设法获得一些问题的答案吗？比如，智能等同于意识吗？或者人类大脑是否像计算机一样运行？最后，人工智能的本质是什么？是构建会思考的机器、构建使我们思考的机器，还是旨在构建使我们有更多时间去思考的智能机器？

唯一真正的答案是：是的，人工智能关乎以上一切内容。

还关乎更多内容。

致谢

著书向来是个孤独的历程。但是，我很庆幸得到了来自如同《黑道家族》（*The Sopranos*）中配角阵容那样团队的鼎力支持，我很幸运，我的支持者中无人有犯罪记录。首先，感谢企鹅兰登书屋的爱德·福克纳（Ed Faulkner）、杰米·约瑟夫（Jamie Joseph）和杰克·林伍德（Jake Lingwood），没有他们，《人工智能》永远只停留在脑海中。其次，感谢我的经纪人玛吉·汉伯里（Maggie Hanbury）、她的助理哈里特·波兰（Harriet Poland）以及本书部分章节提及的与我交谈的所有人。同时，向西蒙·加菲尔德（Simon Garfield）、西蒙·卡洛（Simon Callow）以及科林·怀亚特（Colin Wyatt）表示深深的感谢，我的写作一直离不开他们的支持。同样，我要向我苹果科技网站（Cult of Mac）的同事表示感谢，尤其是出版人林德·卡内（Leander Kahney）及我的英国老乡、作家基利安·贝尔（Killian Bell）。之后我要向老朋友们和同伴们表示深深谢意，包括“恶棍经济学博士”丹·汉弗莱（Dan Humphry）、詹姆斯·布莱济奇（James Brzezicki）、亚历克斯·米林顿（Alex Millington）、汤姆·格林（Tom Green）、汤姆·阿特金森（Tom Atkinson）、尼克·纽波特（Nick Newport）和珍妮·纽波特（Janine Newport）、小气鬼尼克（Nick Meanie）以及迈克尔·格罗特豪斯等。在本书上架之前，迈克尔·格罗特豪斯是第一读者。在此，同时向我的妻子克拉拉（Clara）、新增家庭成员盖伊·多梅尔（Guy Dormehl）和安·多梅尔（Ann Dormehl）以及我忠实的小猫费加罗（Figaro）表示感谢。

对了，我还要向所有花时间和金钱购买此书及我其他书籍的人表示感谢。祝你们多福多寿！

译后记

在人类众多既神奇又符合逻辑的幻想中，人工智能无疑是闪亮的星辰。在新一轮全球科技革命和产业变革进程中，获得迅猛发展的人工智能，已然成为其中重要的引领力量。主要国家和地区、大型科技企业相继把人工智能技术和应用，作为本国、本地区和本企业未来发展的重点，不断加强谋划布局，加大技术产品的研发投入。国务院印发的《中国制造2025》把智能制造作为推动新一代信息技术与制造技术融合发展两化深度融合的主攻方向，提出要着力发展智能装备和智能产品，推进生产过程智能化，培育新型生产方式，全面提升研发、生产、管理和服务的智能化水平。我国人工智能的发展迎来了前所未有的发展新机遇。

在本轮人工智能的大发展中，中国已经开始以领先的姿态呈现在新的赛场上。美国《华盛顿邮报》曾报道，尽管美国是深度学习研究的早期领导者，但根据每年发表关于这一主题的论文数量来看，中国已经超过了美国。如果你关注过中国人工智能的发展，你也会为它取得的成功和展露的前景而欣喜不已。随着《中国制造2025》的深入贯彻落实，人工智能的技术和应用还将在中国更加蓬勃地发展，结出丰硕的成果。

事实上，人工智能不仅仅是专业科技人员关心的话题，也是普通读者感兴趣的事物，而且随着科技的发展，它已开始影响普通人的生活。那么人工智能都包含哪些技术？人工智能方面的科技发展现状是怎样的？未来的科技发展又将带给我们怎样的惊喜？人工智能将怎样改变我们的生活？人们需要一本全面介绍人工智能的书籍来解答疑惑，拓宽视野。国家制造强国建设战略咨询委员会经广泛筛选，最终确定以美国科技记者卢克·多梅尔的新作《人工智能》作为年度译著，呈献给国内读者。此书不仅知识、内容丰富，而且深入浅出、充满趣味、引人入胜。

本书的具体翻译工作由赛迪研究院专家组承担。为确保翻译的质量，参加翻译工作的赛迪专家们兢兢业业，倾注了大量的心血和精力。在此，向他们表示衷心的感谢！

国家制造强国建设战略咨询委员会

2016年9月

版权页

图书在版编目（CIP）数据

人工智能/（美）卢克·多梅尔著；赛迪研究院专家组译. --北京：中信出版社，2016.11

书名原文：Thinking Machines

ISBN 978-7-5086-6822-2

I. ①人... II. ①卢... ②赛... III. ①人工智能—影响—经济发展 IV. ①F061.3-39

中国版本图书馆CIP数据核字（2016）第248061号

人工智能

著者：[美] 卢克·多梅尔

译者：赛迪研究院专家组

策划推广：中信出版社（China CITIC Press）

出版发行：中信出版集团股份有限公司

（北京市朝阳区惠新东街甲4号富盛大厦2座 邮编100029）

（CITIC Publishing Group）

电子书排版：张明霞

中信出版社官网：<http://www.citicpub.com/>

官方微博：<http://weibo.com/citicpub>

更多好书，尽在[大布阅读](#)；

大布阅读：[App下载地址](#)（中信电子书直销平台）

微信号：大布阅读