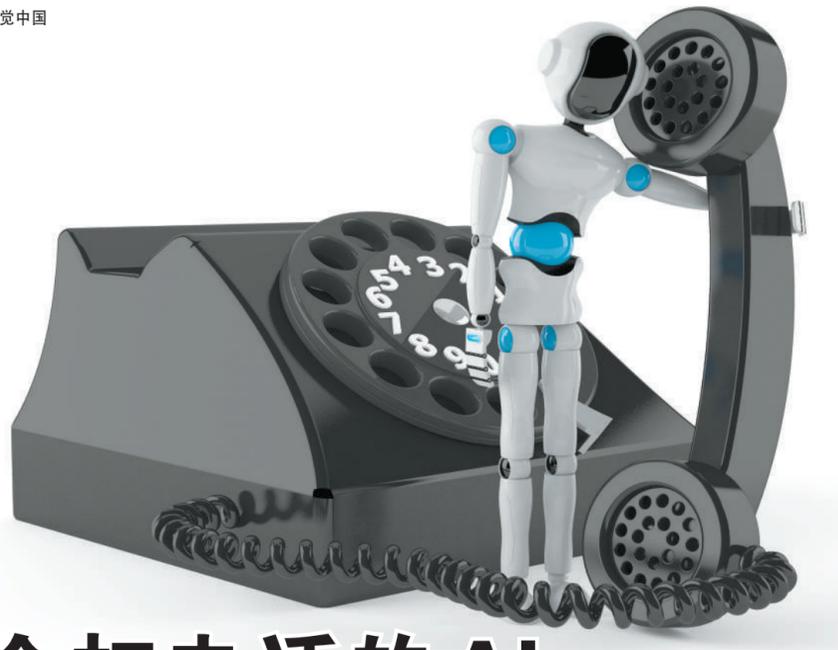


视觉中国



会打电话的AI： 即便为真，也不值得大惊小怪

实习记者 崔爽

近日，谷歌在I/O开发者大会上展示了备受争议的Duplex AI，它通过电话完成了订餐、发廊的预约，并没被对方“识破”。但过于干净的对话

环境，也引发了一些人对于视频真假的争论。

语音合成技术已经可以“以假乱真”？语音交互取得了突破性进展？人和机器可以对话交流？这段展示之后，围观者的惊喜、质疑、担忧纷至沓来。

灵隆科技CEO魏强也曾对媒体表示，目前消费者普遍反馈智能音箱存在远场识别差、误唤醒率高、连续对话功能不稳定和语义理解能力

力差、音质不好等问题，看似火爆的智能音箱行业正在面临创新能力弱、产品体验差、内容和技能不足、用户认知度低等诸多挑战。

技术加持 智能音箱爆发式增长

“相对于触摸、按钮等操纵，语音交互直接、方便、快捷。受惠于人工智能行业飞速发展，语音和语义的识别已经愈发准确。”智能设备厂商负责人赵明对科技日报记者表示。技术进步为智能音箱的产品化提供了支撑。

智能音箱的基本功能是与人类语言互动、播放媒体音频和智能家居提供统一入口，其中语言交互是核心功能。对于用户来说，智能音箱可以用来查询天气、听音乐、进行一些简单问答等。对于厂商来说，智能音箱意义更大：在语音时代，它是重要的流量入口，生产大量数据，并把用户牢牢稳固在自己的

生态链上。

从2014年亚马逊的智能音箱Echo面世，这种智能家居产品在过去几年爆发式增长，巨头纷纷入场，智能音箱发展史上的重要产品不断出现：国外的谷歌Home、苹果HomePod、微软Invoke，国内的京东叮咚、联想音箱、喜马拉雅小雅、阿里巴巴天猫精灵、小米小爱同学等。据赵明介绍，全球智能音箱市场规模在过去几年迅速增长，从2015年的255万台到2017年的3000万台。去年中国智能音箱市场达到176万台的销量，销售额4.9亿元，市场潜力明显。

对标Echo 产品功能初级单一

据赵明介绍，作为智能音箱领域的头部厂商，亚马逊通过接入Alexa语音助手以连接其他硬件软件服务，在先天优势和技术壁垒下，已经形成一定业务体系。亚马逊具有深厚的电商基础，并在内容、硬件、平台、云计算、物流仓储等领域布局，通过智能音箱整合软件、硬件、平台及内容资源，进一步盘活电商资源。

“亚马逊Alexa平台及云计算平台为打通多个服务链条奠定了基础，在硬件操作、软件应用及内容观看收听中的海量数据被有效收集并整理，实现内容与服务的个性化。”赵明表示，亚马逊在中文语音识别方面缺少布局，目前没有进入中国市场，但其发展为中国市场提供了借鉴，多数厂商以Echo为对标，布局智能音箱。

但据赵明介绍，目前智能家居行业还处于整体发展的中前期，行业标准尚未形成，巨头之间难以统一的平台标准为智能家居的普及造成了巨大困难。“目前的AI水平较为有限，很多智能家居实现了基本的AI功能，但都比较初级单一，用户体验并不理想。”

此外，中科院自动化所研究员王金桥表示，中文语音交互存在更多特殊障碍，中文的方言和口音种类过多，交流非常困难，语义鸿沟难以跨越。“市面上的智能音箱都‘傻乎乎’的，只能用来搜索和回答一些条条框框的问题，无法理解逻辑、上下文关系等，大家买回家玩几天，新鲜劲儿一过就闲置了。全国三百多家智能音箱厂商，没有一家赚钱的。”王金桥说。

水平有限 仍是简单场景的人机交互

谷歌演示Duplex AI的过程引发了真实与否的争议：人们注意到，这些通话中的两位女性员工都没有在电话上提到自己公司的名称，另外这两个录音中都没有环境噪音。不过谷歌并没有对是否编辑过电话录音的问题发表评论。

对于Duplex AI的表现，王金桥表示：“这只是展示了一个效果，不能说机器通过了图灵测试，也不代表理论上的突破。况且真实性没有验证。如果是真实的，那确实是先进的自然语言理解，这说明大家都在往自然语言处理的方向努力。”

思必驰副总裁、语音交互科学家初敏表达

了类似的观点。“这个预约的过程难度不大，是一个简单场景的人机交互。只是角色置换了一下，不是我们常见的‘人问机器’，而是‘机器问人’，引发了大家的热议。”初敏说。

初敏表示，谷歌展示的“机器预约餐厅、发廊”一定是可以实现的，没有被对方听出来说明语音合成的效果不错，但整体来看，全程“没有展现有挑战性的问题”——如多任务切换、跨场景、噪音、上下文逻辑理解等。她同样表达了对视频真实性的怀疑：“餐厅听起来应该是乱七八糟的，但视频中却没有背景噪音，所以不排除表演的可能。”

要想质变 需要自然语言处理技术突破

自然语言处理是语音交互的核心技术，也是三十年来几乎没有突破的痛点。通俗来说，这项技术就是“让机器理解人”。据王金桥介绍，自然语言处理的核心是理解，希望机器能理解人的语言和目的。但每个人都是个性化的，说话的轻重缓急、重音位置，看到夕阳西下时或开心或伤感的个人感受，这些都是目前的机器无法理解的。尤其是在中文语境下，常见的一语双关、倒装、句式简化等，人有时候都理解不了，对机器来说更是难上加难。

初敏表示，在语音交互背后，是语音识别、自然语言理解、对话管理、自然语言生成等诸多环节，各环节充分配合才能实现良好的交互效果。“相比于谷歌展示的预约，我们已经可以处理更复杂的交互，只是效果没那么好。”初敏说。“机器的认知推理和理解是一个系统的处

理工程，是各种复杂算法的结合，不是单项技术能解决的。”王金桥表示。目前比较可行的限定场景下的自然语言处理，比如儿童陪伴、学科细分的语料库，可以实现智能化交互。中国科学院院士谭铁牛也在两院院士大会期间谈到人工智能正处在从“不能用”到“可以用”的技术拐点，距离“很好用”还有诸多瓶颈。“像‘那辆白车是黑车’‘能穿多少是多少’一类的表述，机器无法准确理解和翻译。”谭铁牛说。

语言交互是智能音箱的核心功能。智能音箱的销量剧增，大量的语音交互会带来多样化的数据样本，“目前研究界都在试图突破自然语言处理的难关，虽然从事这项研究的人比以前多很多，研究条件也很好，但理论的瓶颈、语言理解的语义鸿沟问题，不是一时半会能解决的。”王金桥说。

数字孪生、无感支付，让未来城市更智慧

产业界

本报记者 郭科

城市各类信息的共享、城市大数据的挖掘利用等，是当前城市发展建设的重要议题。人工智能技术如何更好地利用多元化的数据资源，为城市管理和产业发展提供服务？在近日举办的2018中国国际大数据产业博览会上，两个热词走进了人们的视野——数字孪生与无感支付，为这个问题提供了两种具体应用场景。

数字孪生：建一座“虚拟城市”

“数字孪生”将成为未来智慧城市的一种发展方向。软通智慧总裁冯皓告诉记者，“数字孪生”是从虚拟制造、数字样机等技术发展而来，是以数字化方式为物理对象创建的虚拟模型，来模拟其在现实环境中的行为，最早被用于工业制造领域。而将其引入智慧城市建设，是基于创新的数字孪生技术与全域数字化融合，将物理城市精准映射到虚拟环境，形成数字城市；再

通过城市科学模拟与仿真系统对城市进行科学预测，为城市管理者与企业提供有效的决策辅助，助力城市的发展规划和运营管理、产业转型与普惠惠民。

概括而言，“孪生城市”就是将城市管理的各种数据映射到网络虚拟空间的数字模型，是数字化万物与AI技术的具体体现。

AI算法为实现数字孪生提供了技术支持。它以大数据为基础，通过智能搜索分析，进行自我识别、自我判断，形成人们可见的自动驾驶、自动监控、自动生产等。一方面，AI算法方面的突破，让自动识别速度又快又准；另一方面，5G网络的数据传输速度的极大提升，让人工智能和城市的各种场景结合在了一起。比如运用到自动驾驶领域，原本是汽车有一套独立的智能驾驶系统，道路有一套独立的智能网联道路，运用数字孪生技术将他们整合在一起，就形成了解决出行问题的整体智慧方案。

冯皓表示，目前软通智慧已经将大数据和AI技术运用到了城市的科学规划、产业分析、舆情监测、精准扶贫、公共安全、交通出行、节能环保、文化旅游、医疗养老、社区服务等诸多领域。他们以

城市为平台，提供综合的大数据应用和运营服务，逐步覆盖了政府、产业、民生的方方面面，使城市更加宜居和可持续发展。

与以往相比，孪生城市AI技术的运用将极大提升城市各部门协调管理的自动化。比如，生态环境网格化监测云平台护航智慧环保应用，在降低政府投资风险、规范监管执法、实现部门协调等方面实现了“全空间、全区域、全过程”的动态化智能管控。通过这些AI技术的应用，为政府优化产业结构提供决策依据和大数据支撑。

无感支付：量身定制个性化金融服务

如今移动支付、智能信贷、智能投顾已经走进我们的生活。一个明显的趋势，就是整个社会在朝着数字化、定制化和智能化的方向发展。一切智能的应用一定是先有数据，数据足够大，智能才能更好地发挥作用。随着数据量越来越大，数据的维度越来越多，如何把包括大数据和机器学习的AI技术应用到金融服务等垂直领域的实践之

中，渗透到业务发展的每一个环节？

凡普金科创始合伙人、爱钱进总裁杨帆表示，未来金融将是“无感支付”。当你身处某一金融需求场景之中，一套完善、合理、定制化的金融服务方案便能够自动生成，大家只需简单的操作，就能轻松获得最佳的金融服务方案，真正实现“当你需要时，它就在那里”的无感体验。

在技术上，需要将人工智能、大数据等用于风控、建模等领域。杨帆告诉记者，他们的数据科学家研制了一个RNN神经网络模型，通过机器自动挖掘数据，进而找出一些特定的特征，然后再提取这些特征为用户提供互联网金融的服务。在风险识别的过程中，他们自主研发了FinUp云图风控系统，以及自动建模机器人——水滴和春风，并将这些AI技术用于多项业务中。

“通常的建模过程，是由建模师针对这些数据和本地的结果，去进行调优，非常费时，可能一个模型从有数据，到最终模型建成，要消耗几个月的时间。而通过自动建模机器人，可以把中间调优的过程、特征搜索的过程，完全由机器来完成。随着技术的不断迭代，我们建模的速度还在加快。”杨帆说。

好机友

国内首款下肢软体外骨骼问世 造福腿脚不便者

近日，位于上海张江的司羿智能对外发布消息，推出国内首款下肢软体外骨骼Easy Walk TMX1，该产品重量不到3千克，将使佩戴者告别笨重和不舒适两大难题。

记者了解到，TMX1没有刚性骨架，而是通过纺织物固定，穿戴舒适，可以帮助卒中患者、脑外伤患者解决行走障碍问题，还能帮助中风患者、阿尔茨海默氏症、多重硬化症、帕金森症患者。

司羿智能总经理尹刚刚介绍，目前国内外的外骨骼产品都为硬性外骨骼，其定价较高，重量都在12千克以上。外骨骼机器人强制人行走机械僵化运动形式，一般只能用于康复中心以固定式的步态带动人体进行康复训练，而不能用于患者日常生活辅助。“TMX1则定位于帮助运动能力欠佳者，为其提供运动辅助。此外，人工智能算法通过穿戴传感器数据，可识别出步态周期中需要助力的时间点，同时能自适应各种行走模式，帮助患者实现坐、站、行走、上下楼梯等基本功能，未来正常人都能携带该产品进行日常活动，或是郊游、登山。”尹刚刚说。

据悉，司羿智能研发的另一款仿生软性康复机器人手套，已与上海华山医院康复科开展密切临床应用合作。该康复机器人手套通过仿生气动力人工肌肉驱使手指关节活动，可使手指痉挛、麻痹、瘫痪等症状得到极大的改善，有效帮助手的各项功能的康复，同时帮助作用于脑部神经，促进运动康复。（记者王春）

新鲜事

日科学家研发全新交互技术 用人工智能给图像配乐



据遇见人工智能报道，近日，日本研究人员通过一款叫“Imaginary Soundscape”的网络应用程序，使用机器学习技术，为用户上传的图片和最符合的音频配乐。

在该项目页面中，用户上传一张图片，然后进入另外一个页面，显示上传的图片的同时，就会播放出经过AI算法匹配后的背景音乐。例如，上传日本的渔船木刻，系统会匹配波浪和水声；加载一幅抽象的夜莺画作，人工智能会为它搭配一首有风铃和鸟叫的背景音乐。更有意思的是，当用户上传的图片是一些稀奇古怪的元素时，系统反馈的配音也会带来惊喜，比如将擎天柱与拖拉机的声音配对。

该系统是一个相对简单的AI算法的产物。它使用物体识别技术识别出图片中的元素，然后将这些元素与超过52000个声音文件的数据集进行匹配。该系统的研究初衷是让音频与Google街景视图中的图像匹配。不过，该系统并不完美，比如模糊识别会造成错误，而声音数据库尚不完整。但它仍是AI支持新型创意表达的绝佳示例。

五角大楼又一军事项目曝光 美军用AI预测核导弹发射



据DeepTech深科技报道，有匿名的美国官员向媒体透露，美国军方十分看好AI技术在军事领域的应用，已启动多个AI军事项目。其中一个鲜为人知的项目是，利用AI寻找核导弹发射器的位置，并预测核导弹的发射。系统的侦测范围覆盖众多国家和地区，目的是更好地保护美国免受潜在的核导弹袭击。

研发中的AI系统将以远超人类的速度和准确度搜寻大量数据，之后根据这些数据寻找导弹准备发射的迹象。一旦有所发现，美国军方将收到早期预警，这增加了摧毁或拦截导弹的成功率，同时政府也可以更早地开展外交手段，从根本上避免导弹发射。

目前这一项目还在起步阶段，不过导弹发射器追踪系统的早期原型已经投入试验。除了军方手中的数据，研发人员还与商业公司合作，以获得更多的数据和行为模式，其中包括看穿风暴和树林的雷达数据。据熟悉计划的消息人士透露，该项目不过是AI军事计划的冰山一角。

(本版图片除标注外来源于网络)

扫一扫
欢迎关注
AI瞭望站
微信公众号