



当DF遇见AI

山东警察学院

张璇

Artificial Intelligence

1956年，在达特茅斯学院举行的一次会议上，不同领域（数学，心理学，工程学，经济学和政治学）的科学家正式确立了人工智能为研究学科。



Trenchard More、John McCarthy、Marvin Minsky、Oliver Selfridge、Ray Solomonoff

Artificial Intelligence

| 像人一样思考的系统 | 理性地思考的系统 |
|--|--|
| <ul style="list-style-type: none">• “要使计算机能够思考……意思就是：有头脑的机器” (Haugeland, 1985)• “与人类的思维相关的活动，诸如决策、问题求解、学习等活动” (Bellman, 1978) | <ul style="list-style-type: none">• “通过利用计算模型来进行心智能力的研究” (Chamiak和McDermott, 1985)• “对使得知觉、推理和行为成为可能的计算的研究” (Winston, 1992) |
| 像人一样行动的系统 | 理性地行动的系统 |
| <ul style="list-style-type: none">• “一种技艺，创造机器来执行人需要智能才能完成的功能” (Kurzweil, 1990)• “研究如何让计算机能够做到那些目前人比计算机做得更好的事情” (Rich和Knight, 1991) | <ul style="list-style-type: none">• “计算智能是对设计智能化智能体的研究” (Poole等, 1998)• “AI……关心的是人工制品中的智能行为” (Nilsson, 1998) |

AI：延申人类智能

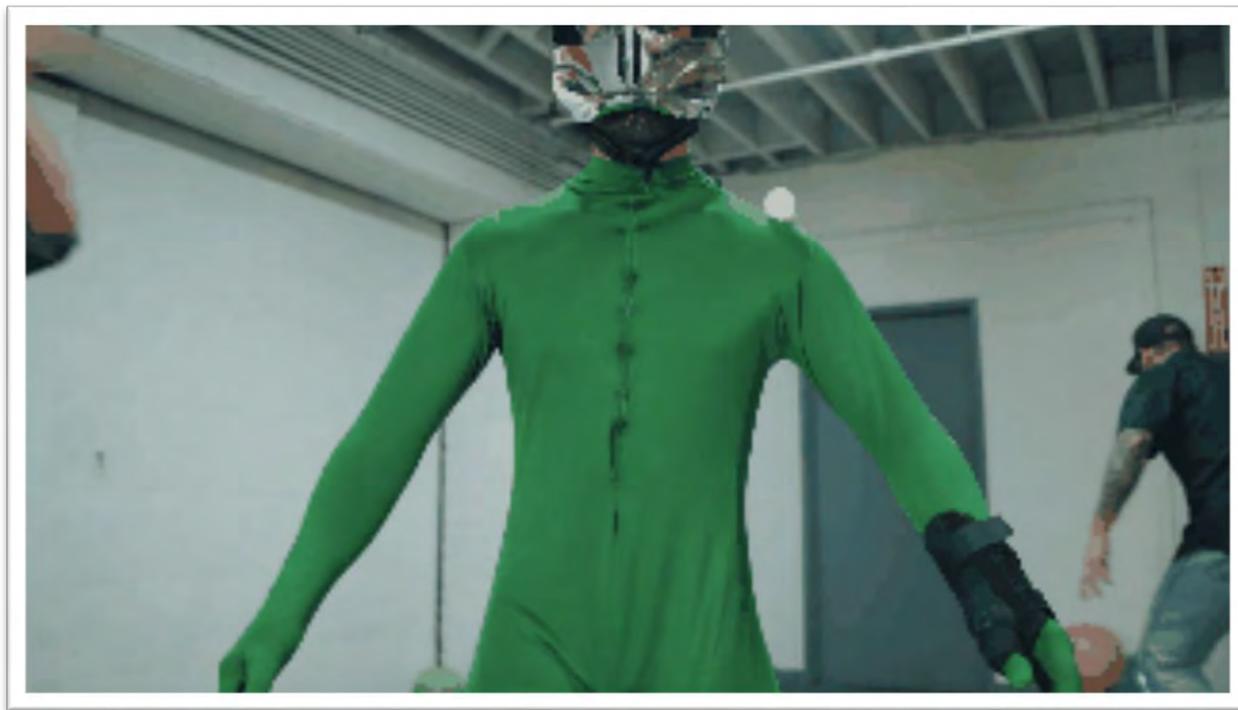
Artificial Intelligence

研究如何制造出人造的智能机器或系统，来模拟人类智能活动的的能力，以延伸人类智能的科学。

人工智能会让我更多的去关注学习的过程是什么，让我们思考什么是人，什么是智能。——李飞飞

未来人工智能也许会是人类的终结者。——霍金

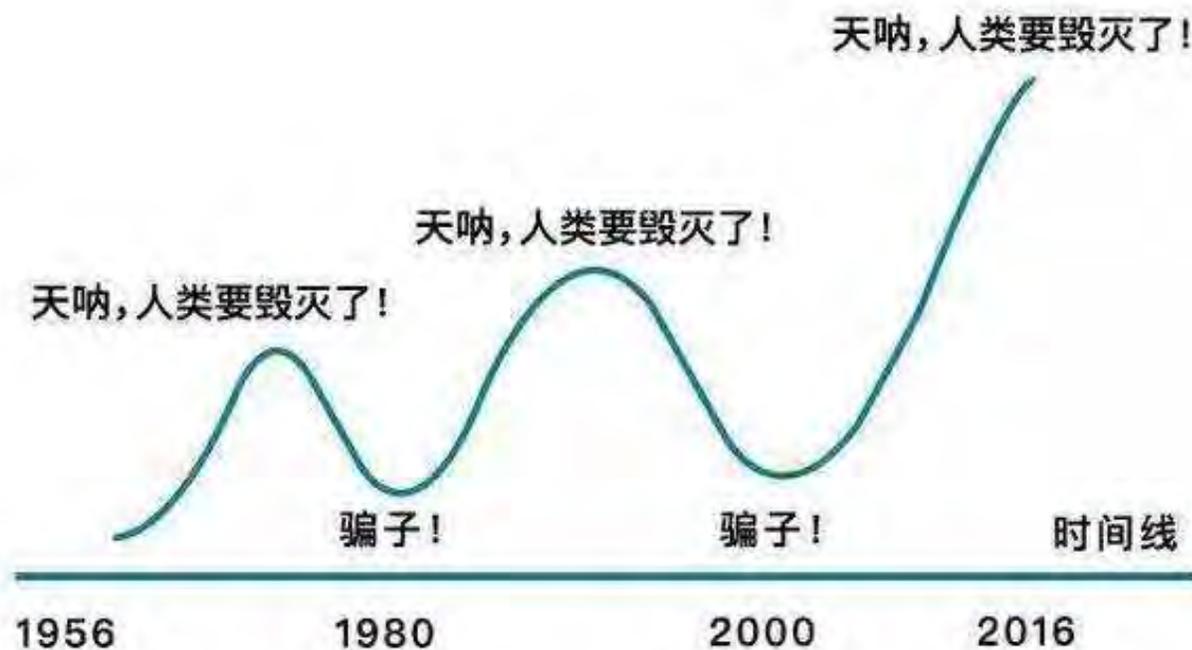




人类松了一口气？

AI：70年浮沉史

- 1958年, H. A. Simon, Allen Newell: “十年之内, 数字计算机将成为国际象棋世界冠军。” “十年之内, 数字计算机将发现并证明一个重要的数学定理。”
- 1965年, H. A. Simon: “二十年内, 机器将能完成人能做到的一切工作。”
- 1967年, Marvin Minsky: “一代之内.....创造“人工智能”的问题将获得实质上的解决。”
- 1970年, Marvin Minsky: “在三到八年的时间里我们将得到一台具有人类平均智能的机器。” ...



AI：70年浮沉史

弱人工智能

特定领域、既定规则中，
表现出强大的智能

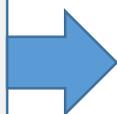
强人工智能

不受领域、规则限制，
具有人类同样的创造力
和想象力

超级人工智能

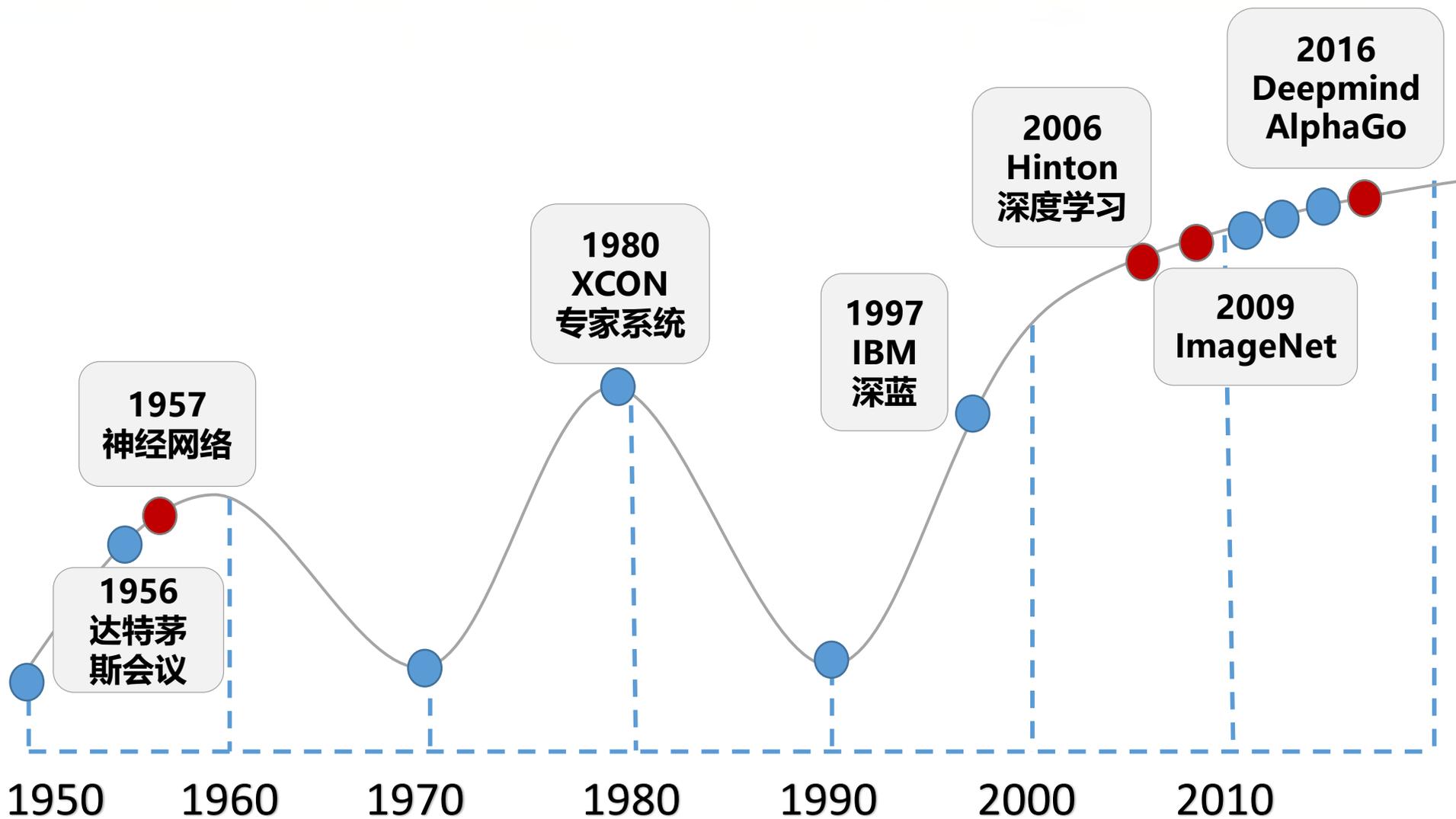
超越人类的智能

跨领域推理
拥有抽象能力
“知其然，也知其所以然”
拥有常识
拥有审美能力
拥有自我意识和情感

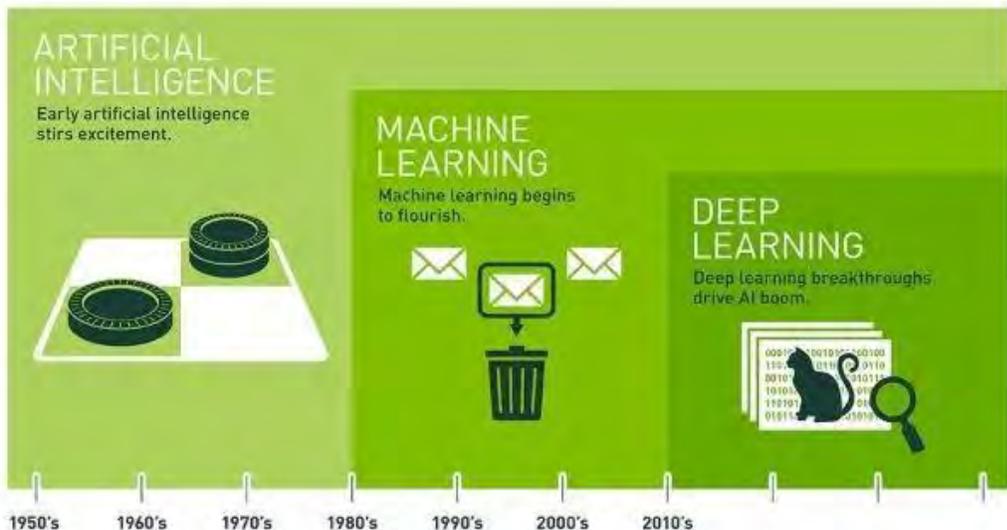


知识表示的能力，包括常识性知识的表示能力
规划能力
学习能力
使用自然语言进行交流沟通的能力
将上述能力整合起来实现既定目标的能力

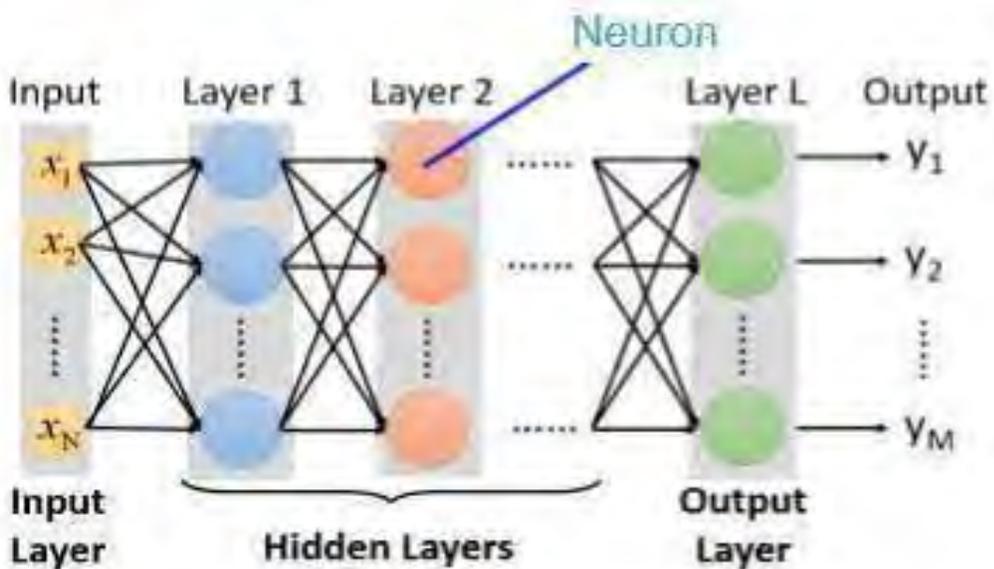
AI：70年浮沉史



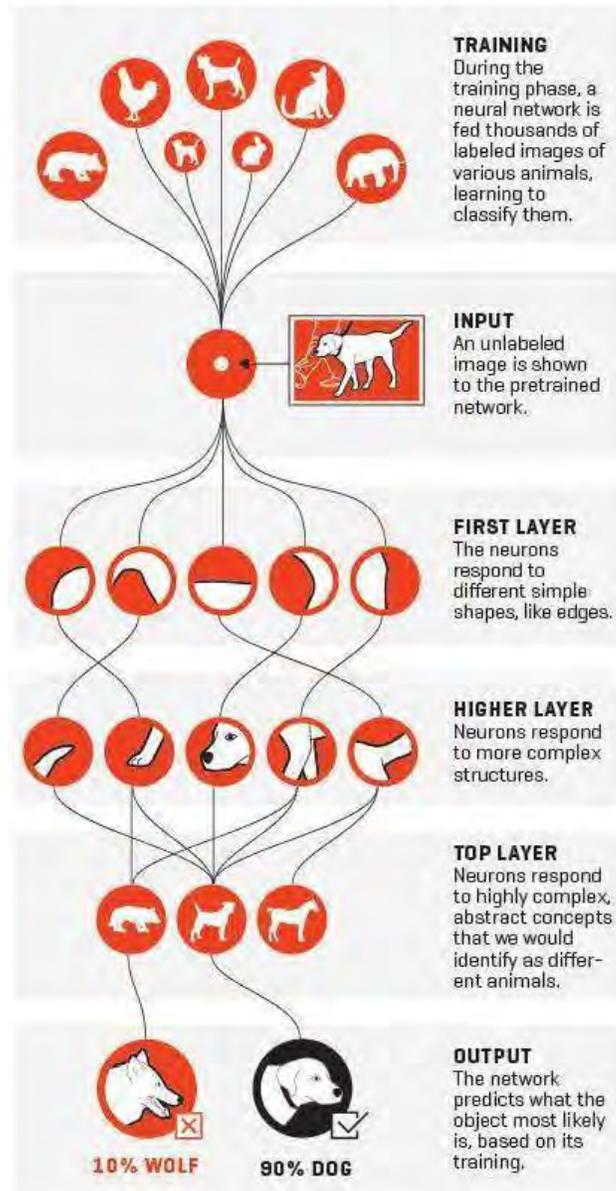
AI: 70年浮沉史



Since an early flush of optimism in the 1950s, smaller subsets of artificial intelligence – first machine learning, then deep learning, a subset of machine learning – have created ever larger disruptions.



HOW NEURAL NETWORKS RECOGNIZE A DOG IN A PHOTO



AI: 70年沉浮史

2018图灵奖Hinton、Bengio、LeCun深度学习三巨头共享



AWARD WINNER

Geoffrey E Hinton

ACM A. M. Turing Award (2018)

2018 ACM A.M. Turing Award

- 反向传播
- 玻尔兹曼机
- 对卷积神经网络的修正



AWARD WINNER

Yoshua Bengio

ACM A. M. Turing Award (2018)

2018 ACM A.M. Turing Award

- 序列的概率建模
- 高维词嵌入与注意力机制
- 生成对抗网络



AWARD WINNER

Yann LeCun

ACM A. M. Turing Award (2018)

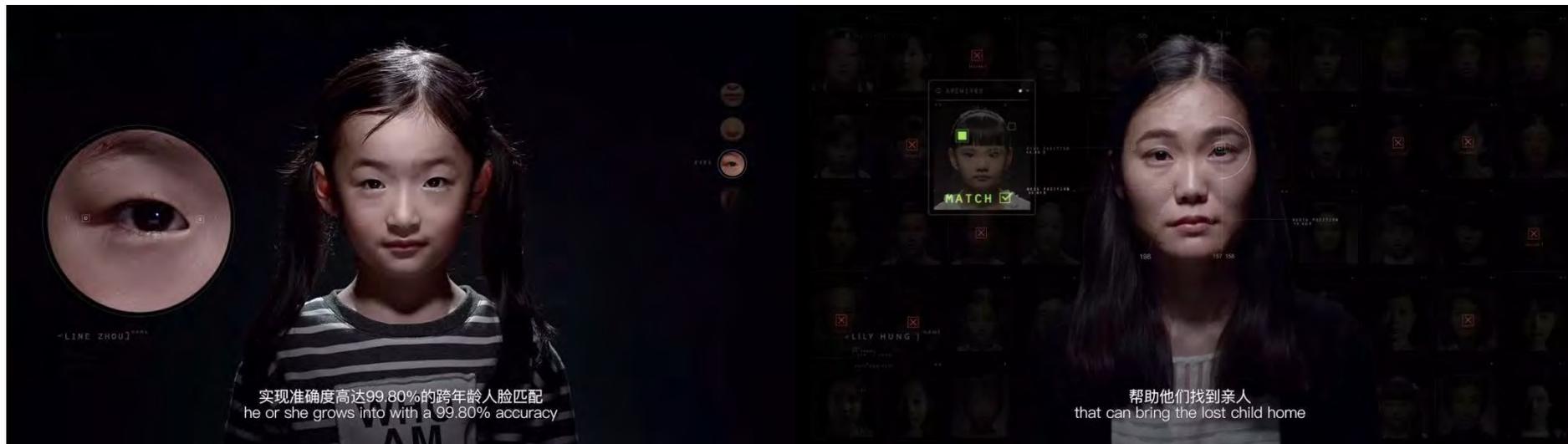
2018 ACM A.M. Turing Award

- 提出卷积神经网络
- 改进反向传播算法
- 拓宽神经网络的视角

AI：未来已来



Artificial Intelligence



国内首次！警方用AI一次找回4名走失10年的孩子！

跨年龄人脸识别示意图/腾讯优图实验室

目前准确率已达到了99.80%以上

Artificial Intelligence



阿里AI 进化到助理法官水平，一秒“判案”

深度学习、迁移学习技术

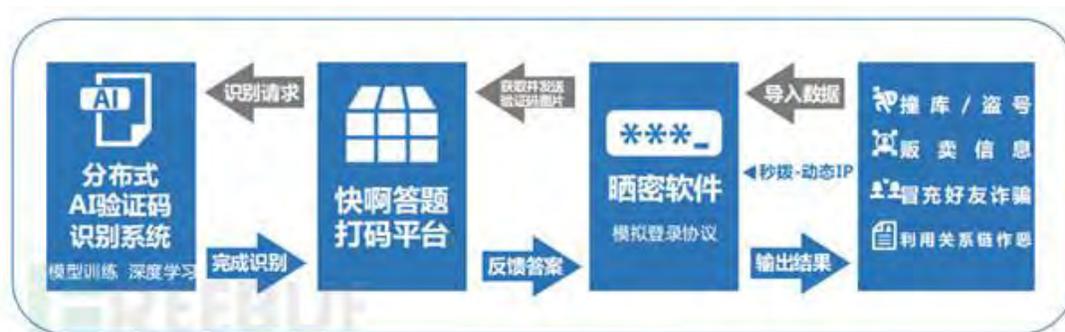
AI: 未来已来

Artificial Intelligence

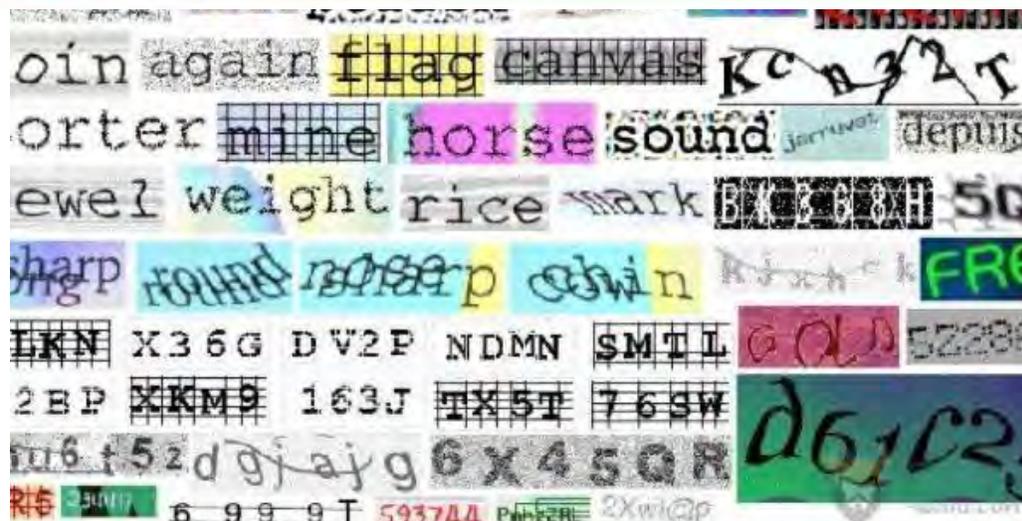
All in AI
AI in All
All can AI



2017年全国首例AI犯罪案：能识别98%的验证码，泄露10亿多组个人信息



网络犯罪+AI



快啊平台早期通过雇佣人员进行人工识别验证码扩充数据库，之后利用AI训练机器以及深度学习进行快速匹配，一秒就能完成1000多次打码，识别验证码的成功率达98%。

AI：未来已来

索引号: 000014349/2017-00142
发文字号: 国发〔2017〕35号
标题: 国务院关于印发新一代人工智能发展规划的通知
主题词: 国务院 国发〔2017〕35号
发布日期: 2017年07月20日
生效日期: 2017年07月08日
主题分类: 科技、教育、科技

国务院关于印发 新一代人工智能发展规划的通知

国发〔2017〕35号

相关报道

国务院印发《新一代人工智能发展规划》

| 智能产品 | 2020年 |
|------------|---|
| 智能互联网汽车 | 建立可靠、安全、实时性强的智能网联汽车智能化平台，形成平台相关标准，支撑高度自动驾驶（HA级） |
| 智能服务机器人 | 智能服务机器人关键技术取得突破，智能家庭服务机器人、智能公共服务机器人实现批量生产及应用，医疗康复、助老助残、消防救灾等机器人实现样机生产 |
| 智能无人机 | 智能消费级无人机三轴机械增稳云台精度达到0.005度，实现360度全向感知避障 |
| 医疗影像辅助诊断系统 | 国内先进的多模态医学影像辅助诊断系统对以上典型疾病的检出率超过95%，假阴性率低于1%，假阳性率低于5% |
| 视频图像身份识别系统 | 复杂动态场景下人脸识别有效检出率超过97%，正确识别率超过90%，支持不同地域人脸特征识别 |
| 智能语音交互系统 | 实现多场景下中文语音识别平均准确率达到96%，5米远场识别率超过92%，用户对话意图识别准确率超过90% |
| 智能翻译系统 | 多语种智能互译取得明显突破，中译英、英译中场景下产品的翻译准确率超过85% |
| 智能家居产品 | 智能家居产品类别明显丰富，智能电视市场渗透率达到90%以上 |

图解：中国制造2025

战略目标

立足国情，立足现实，力争通过“三步走”实现制造强国的战略目标

第一步 力争用十年时间，迈入制造强国行列。

到2020年，基本实现工业化，制造业大国地位进一步巩固，制造业信息化水平大幅提升。到2025年，制造业整体素质大幅提升，创新能力显著增强，全员劳动生产率明显提高，两化（工业化和信息化）融合迈上新台阶。

第二步 到2035年，我国制造业整体达到世界制造强国阵营中等水平。

第三步 新中国成立一百年时，制造业大国地位更加巩固，综合实力进入世界制造强国前列。

战略任务和重点

- 1 提高国家制造业创新能力。
- 2 推进信息化与工业化深度融合。
- 3 强化工业基础能力。
- 4 加强质量品牌建设。
- 5 全面推行绿色制造。
- 6 大力推动重点领域突破发展。
- 7 深入推进制造业结构调整。
- 8 积极发展服务型制造和生产性服务业。
- 9 提高制造业国际化发展水平。

战略支撑与保障

- 1 深化体制机制改革。
- 2 营造公平竞争市场环境。
- 3 完善金融扶持政策。
- 4 加大财税政策支持力度。
- 5 健全多层次人才培养体系。
- 6 完善中小微企业政策。
- 7 进一步扩大制造业对外开放。
- 8 健全组织实施机制。

明确五大重点工程

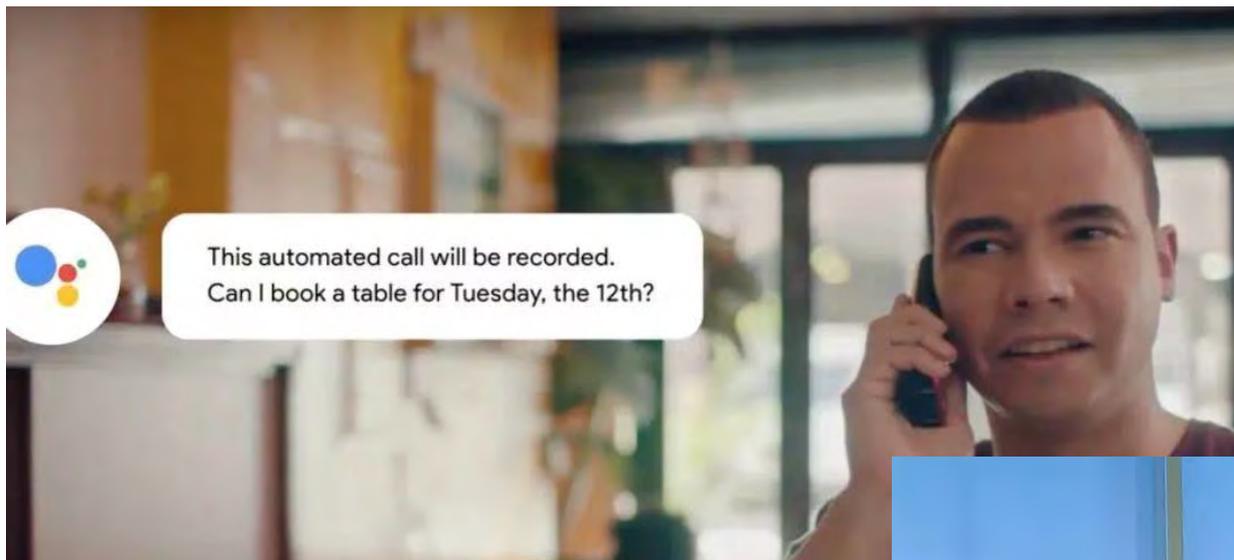


AI：理想与现实的距离

Artificial Intelligence

如何知道一个系统是否具有智能？

1950年，计算机科学家图灵提出了著名的“图灵测试”。

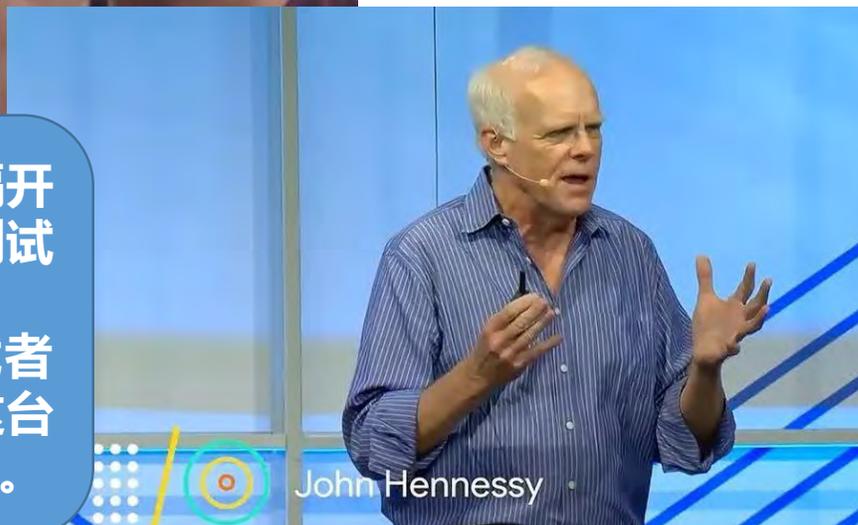


Duplex

谷歌造假！通过图灵测试的AI实测：4次成功完成任务，3次靠的是人工

测试者与被测试者（一个人和一台机器）隔开
的情况下，通过一些装置（如键盘）向被测试者
随意提问。

进行多次测试后，如果有超过30%的测试者
不能确定出被测试者是人还是机器，那么这台
机器就通过了测试，并被认为具有人类智能。



AI：理想与现实的距离

Artificial Intelligence

2018 年十大 AI 失败案例



← 微博正文 ...

宁波交警 11-21 17:28 来自 iPhone 8 已编辑

“董小姐”闯红灯了？原来是一场乌龙——
今天有网友通过微博发布了一条内容为“董小姐闯红灯”的图文信息，经核实是一套装置在江厦桥东的“行人非机动车闯红灯抓拍系统”，对一辆正在沿江东北路由南往北行驶的公交车身广告上的人像进行了误识别，交警部门事后立即进行删除，目前技术人员已对该系统进行了全面升级，减少误识别率

AI：理想与现实的距离

Artificial Intelligence



[Our Technology](#)

[Verticals](#)

[About us](#)

[News & Events](#)

[Blog](#)

[Contact us](#)

OUR CLASSIFIERS



High IQ



Academic Researcher



Professional Poker
Player



Terrorist

Utilizing advanced machine learning techniques we developed and continue to evolve an array of classifiers. These classifiers represent a certain persona, with a unique personality type, a collection of personality traits or behaviors. Our algorithms can score an individual according to their fit to these classifiers.

[Learn More>](#)

利用人工智能技术分析面部图像和骨骼结构，揭示人们的智商、个性，甚至暴力倾向

当电子数据取证遇见AI

2019年6月1日，在DEF CON CHINA 1.0现场BCTF百度网络安全技术对抗赛上，来自国防科技大学的机器人战队HALFBIT，在没有人工干预的情况下，全自动完成了对堆内存溢出漏洞发现及利用，这是全球范围内，这一领域的突破性实践首次在公开、公平的竞赛环境中现场展示。

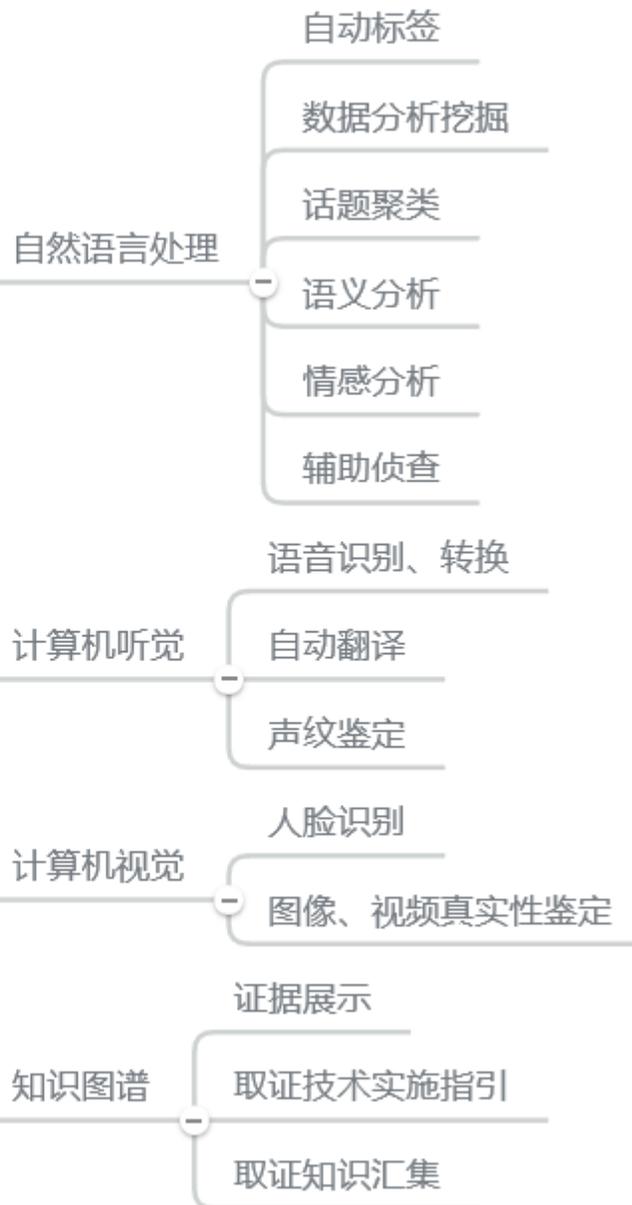


那么取证呢？

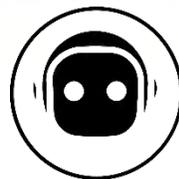
当电子数据取证遇见AI

- 海量数据
- 异构数据
- 新兴技术
- 数据分析
- 推理归纳
- 证据展示
- 反取证技术

利用AI的电子数据取证



当电子数据取证遇到AI



| | | |
|----------|--------------|--------------|
| 1 | 自主决策 自主行动 | 工具 |
| 2 | 主导 | 辅助决策 |
| 3 | 审核执行 | 决策建议 |
| 4 | 监督 | 主导 |
| 5 | 通用AI | 自主决策 自主执行 |



当电子数据取证遇到AI



新闻



- | | |
|----------------------------------|------------|
| 抖出不一样 霸气的外表 温柔的内心, 爱上了吗? | 2019-06-05 |
| 重要的事再说一遍! 美亚柏科2019实习生招聘火热进行中!! | 2019-05-29 |
| 抖出不一样 揭开神秘面纱! 原来美亚柏科人的一天这么666! | 2019-05-28 |
| 定了! A股“入富”最终名单揭晓, 美亚柏科名列其中! | 2019-05-27 |
| 赋能美亚技术生态, 美亚柏科首届小程序大赛成功举办 | 2019-05-25 |
| 国家市场监督管理总局副局长田世宏一行莅临美亚柏科参观考察 | 2019-05-23 |

当电子数据取证遇见AI



美亚柏科·AI开放平台(MiaAI) 首页 产品中心 API文档 登录 注册

API文档

数十项全球领先的人工智能服务助您赋能产品，AI开发者之旅从这里开始。

接口API

| 接口分类 | 接口名称 | 功能说明 |
|------|---------|---|
| 人脸识别 | 人脸检测与分析 | 精准定位图中人脸，分析性别、年龄、种族、是否戴墨镜、是否戴口罩、是否有胡子等多种人脸属性。 |
| 人脸识别 | 人脸1:1检测 | 对比两张人脸的相似度，并给出相似度评分，从而判断是否为同一个人。 |
| 图像识别 | 车辆属性分析 | 支持对机动车辆品牌、车牌号和颜色的识别，目前只支持蓝色车牌。 |
| 图像识别 | 图片鉴黄 | 人工智能鉴黄技术，智能识别图片中的色情和性暴露内容，让您的应用轻松过审，远离违规风险。 |
| 图像识别 | 暴恐图片识别 | 识别暴力、血腥场景及恐怖组织头目、旗帜等涉嫌违禁的图片内容，降低应用涉暴涉恐风险。 |

- API文档
 - API接口
- 人脸识别
 - 人脸检测与分析
 - 人脸1:1检测
 - 人脸1:N搜索
- 图像识别
 - 物体检测
 - 车辆属性分析
 - 场景识别
 - 人像属性
 - 文字识别
 - 图片比对
- 语音文本识别
 - 中文语音识别
 - 维文转中文
 - 涉黄文本识别
 - 涉毒文本识别
 - 涉赌文本识别
 - 三院文本识别

当电子数据取证遇见AI

语音识别

本工具用于案例/案件目录下面的微信语音文件（amr格式）识别成文本信息，仅支持标准汉语，需要连接互联网或局域网语音服务器 [使用说明](#)

导入文件 导入目录 取消导入 全选 移除

关键字搜索，多个关键字用空格隔开，不区分大小写

| NO. | 名称 | 大小(B) | 状态 | 用时(秒) | 识别内容 |
|-----|----------------|-------|--------|-------|---------------------------------------|
| 70 | 微信语音 (160).amr | 48288 | 识别成功 | 187 | 我离身雨衣安排呢？晚点没聊都里面的威王春梦，联通给英国的大宝，李娜别人.. |
| 71 | 微信语音 (161).amr | 33312 | 识别成功 | 174 | 可以啊，我们这两天也不忙就是明天的话，还要去不理，那拍婚纱照啊！ |
| 72 | 微信语音 (162).amr | 31296 | 识别成功 | 92 | 原来梦游了作正给弄好了喉音望天就文婧怒了不给强有力。 |
| 73 | 微信语音 (163).amr | 29664 | 识别成功 | 187 | 令人讨厌三个被抓了，然后被人诨病的零月几个婆子没空儿，你该早点给我啊！ |
| 74 | 微信语音 (164).amr | 29472 | 识别成功 | 174 | 呃有你吧！手机号码告诉我，我有你的手机号码跟你百业问问，她给我微信。 |
| 75 | 微信语音 (165).amr | 28512 | 识别成功 | 151 | 视觉系呀，然后我们自己开的不用去，那你们再的。 |
| 76 | 微信语音 (166).amr | 25632 | 识别成功 | 82 | 比那个二十二讲到有没有。据说要五零现在工资每人都删除。他三制啦！ |
| 77 | 微信语音 (167).amr | 25248 | 识别成功 | 131 | 把微信的二维码发给你你扫一下哦，照像簿登陆啊，我要把他令弟去世纪。 |
| 78 | 微信语音 (168).amr | 18912 | 识别成功 | 74 | 妹妹参加推动哦回来，哥哥不未起动漫聊了梦遥啊！ |
| 79 | 微信语音 (169).amr | 18912 | 识别成功 | 111 | 其实就是一个就是只有一个连在一起的。 |
| 80 | 微信语音 (17).amr | 8832 | 识别中... | | |
| 81 | 微信语音 (170).amr | 17952 | 识别中... | | |
| 82 | 微信语音 (171).amr | 16992 | 识别中... | | |
| 83 | 微信语音 (172).amr | 16992 | 识别成功 | 187 | 婆婆婆婆晚点给你电话，你在这边过。 |

转换成功：55个；正在转换：445个

语音播放 内容编辑 导出列表 导出失败项 导出勾选项

- 语音识别
- 情景预测
- 人脸检测
- 图像识别 (鉴黄)

手机取证系统 案例 (20150926210739)

用户名: 李宝

管理 添加证据 数据浏览 数据挖掘 报告管理 数据上传 工具集 反恐利剑 情报分析

人物属性 身份标识 照片分析 轨迹分析 关联分析 关联碰撞 全文检索 情景分析

检材: 张三全 iPhone 6S 模型: 所有模型 置信度: 70 开始分析 导出报告

情景预测

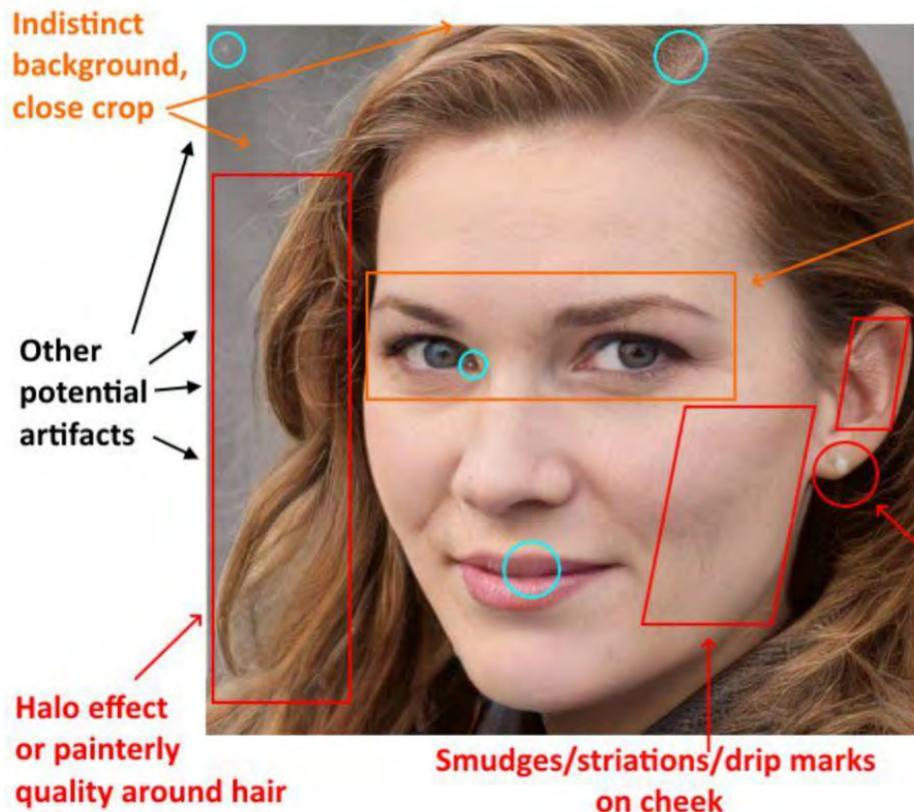
【涉黄】

【涉赌】 【涉毒】

【检材】 iPhone 6S 【模型】 黄、赌、毒模型 【置信度】 70

| 序号 | 应用类型 | 虚拟身份 | 涉【黄】嫌疑数据 | 涉【赌】嫌疑数据 | 涉【毒】嫌疑数据 | 操作 |
|----|------|-------------------|----------|----------|----------|------|
| 1 | 手机短信 | 13959279769 | 0 | 10 | 36 | 生成标签 |
| 2 | 微信 | 16181236 | 0 | 16 | 53 | 生成标签 |
| 3 | QQ | 18989279874 | 3 | 0 | 17 | 生成标签 |
| 4 | 陌陌 | wimdy | 16 | 35 | 32 | 生成标签 |
| 5 | 钉钉 | 15847789856 | 2 | 0 | 0 | 生成标签 |
| 6 | 微博 | wimdy | 6 | 0 | 0 | 生成标签 |
| 7 | 邮件 | wimdy@hotmail.com | 0 | 1 | 0 | 生成标签 |

可能改变鉴定规则 鉴定将更加依赖技术 训练样本缺乏



Connect

Katie Jones

Russia and Eurasia Fellow
Center for Strategic and International Studies (CSIS) ·
University of Michigan College of Literature, Science...
Washington · 49 connections

间谍用GAN生成假头像

图像、视频篡改、伪造



DeepFakes

<http://deepfakes.com.cn/>

<https://github.com/deepfakes/faceswap>

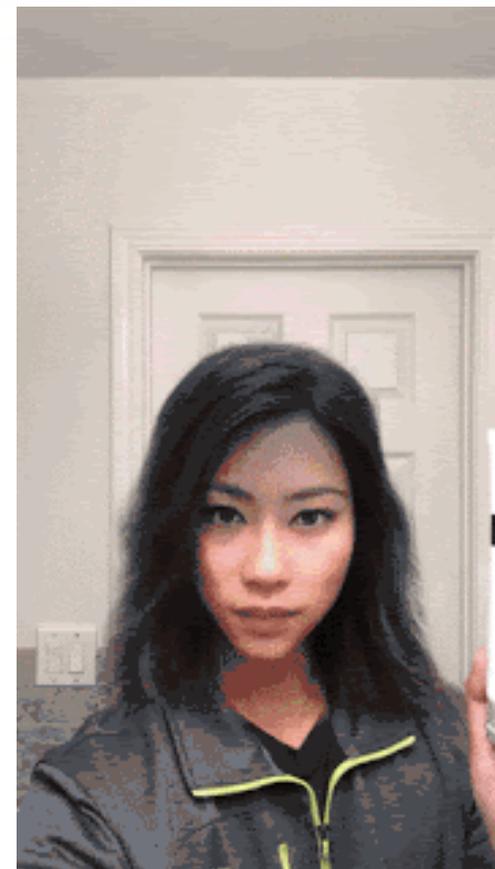


笔迹鉴定

GAN 可以被用来学习生成各种各样的字体

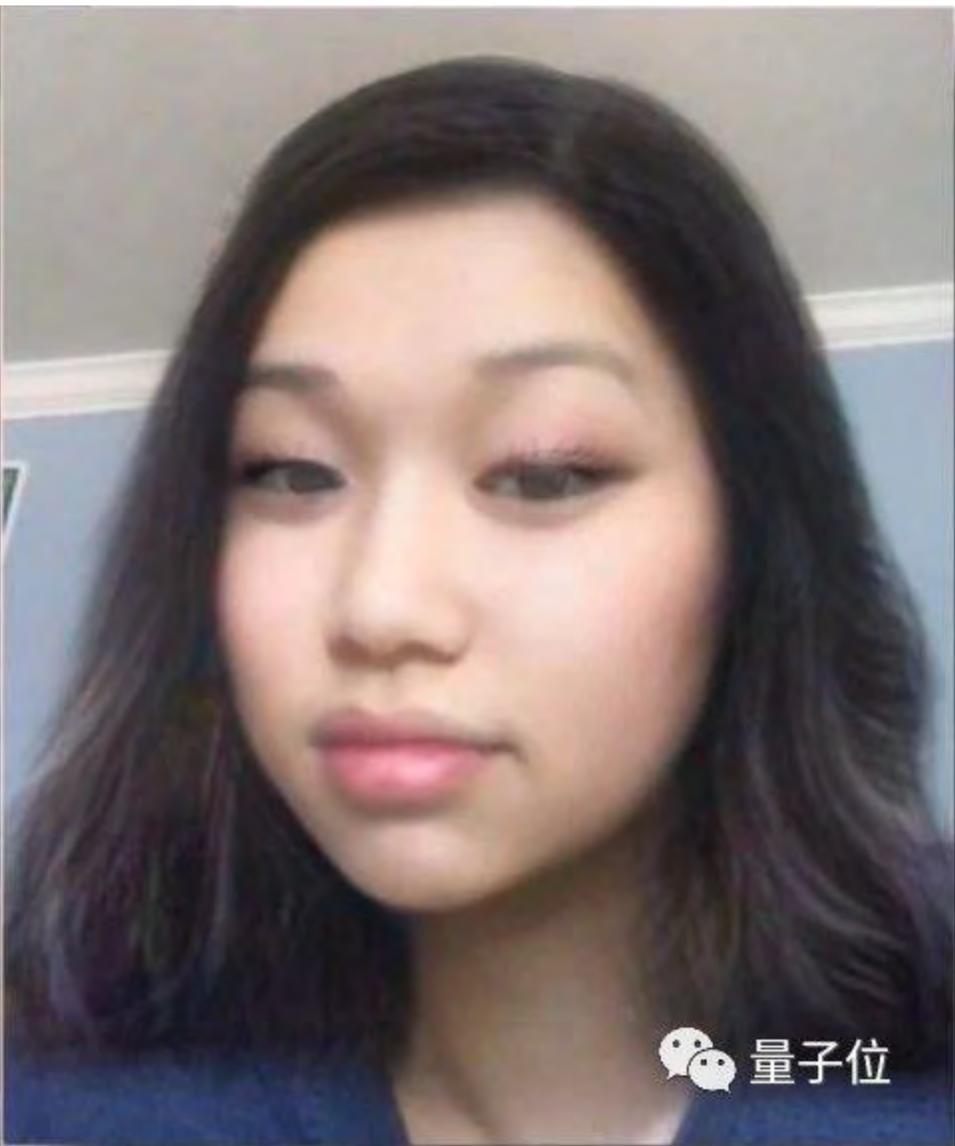
| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 久 | 久 | 肠 | 肠 | 柞 | 柞 | 裘 | 裘 | 肋 | 肋 |
| 亮 | 亮 | 馗 | 馗 | 叉 | 叉 | 哲 | 哲 | 甘 | 甘 |
| 阙 | 阙 | 洧 | 洧 | 材 | 材 | 簿 | 簿 | 装 | 装 |
| 序 | 序 | 魁 | 魁 | 坤 | 坤 | 瞬 | 瞬 | 晋 | 晋 |
| 骺 | 骺 | 熟 | 熟 | 又 | 又 | 吹 | 吹 | 澜 | 澜 |
| 矜 | 矜 | 绸 | 绸 | 弧 | 弧 | 取 | 取 | 嘆 | 嘆 |
| 厝 | 厝 | 涉 | 涉 | 否 | 否 | 椎 | 椎 | 訟 | 訟 |
| 餓 | 餓 | 蚕 | 蚕 | 湖 | 湖 | 陰 | 陰 | 涂 | 涂 |
| 泪 | 泪 | 丞 | 丞 | 铎 | 铎 | 瑗 | 瑗 | 拉 | 拉 |
| 緝 | 緝 | 蒯 | 蒯 | 哂 | 哂 | 燠 | 燠 | 虾 | 虾 |
| 禄 | 禄 | 系 | 系 | 閔 | 閔 | 厠 | 厠 | 荏 | 荏 |
| 特 | 特 | 鉏 | 鉏 | 祝 | 祝 | 敦 | 敦 | 鏃 | 鏃 |
| 犖 | 犖 | 亏 | 亏 | 话 | 话 | 讠 | 讠 | 瑾 | 瑾 |
| 吗 | 吗 | 镰 | 镰 | 缙 | 缙 | 鬚 | 鬚 | 徒 | 徒 |
| 乳 | 乳 | 磁 | 磁 | 球 | 球 | 脍 | 脍 | 隙 | 隙 |
| 蛩 | 蛩 | 節 | 節 | 林 | 林 | 愜 | 愜 | 翎 | 翎 |
| 在 | 在 | 泥 | 泥 | 蚊 | 蚊 | 努 | 努 | 曠 | 曠 |
| 古 | 古 | 黨 | 黨 | 鵠 | 鵠 | 筆 | 筆 | 璣 | 璣 |

CycleGAN



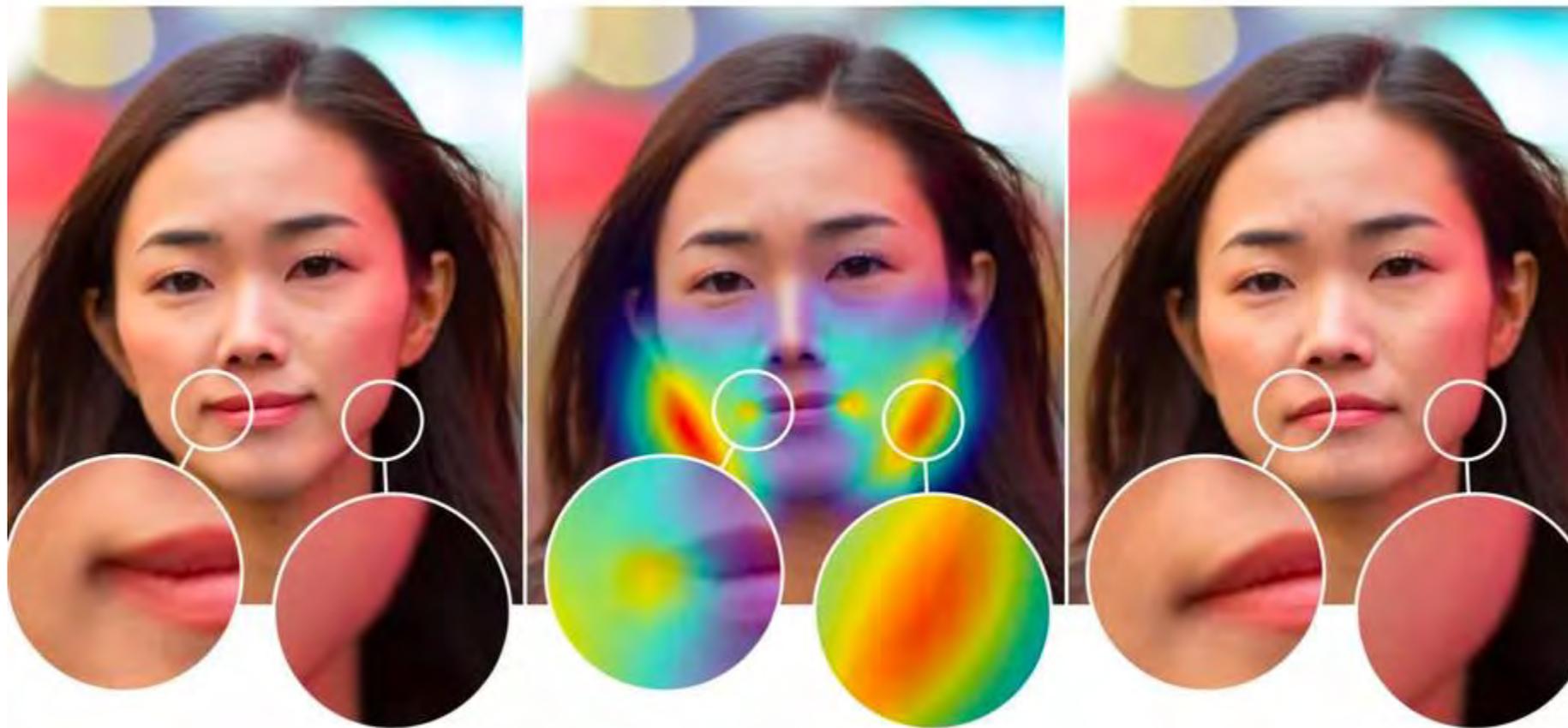
一键变性

<https://junyanz.github.io/CycleGAN/>



用AI变身16岁女孩，抓住40岁违法警察！

解铃还需系铃人



(a) Manipulated photo

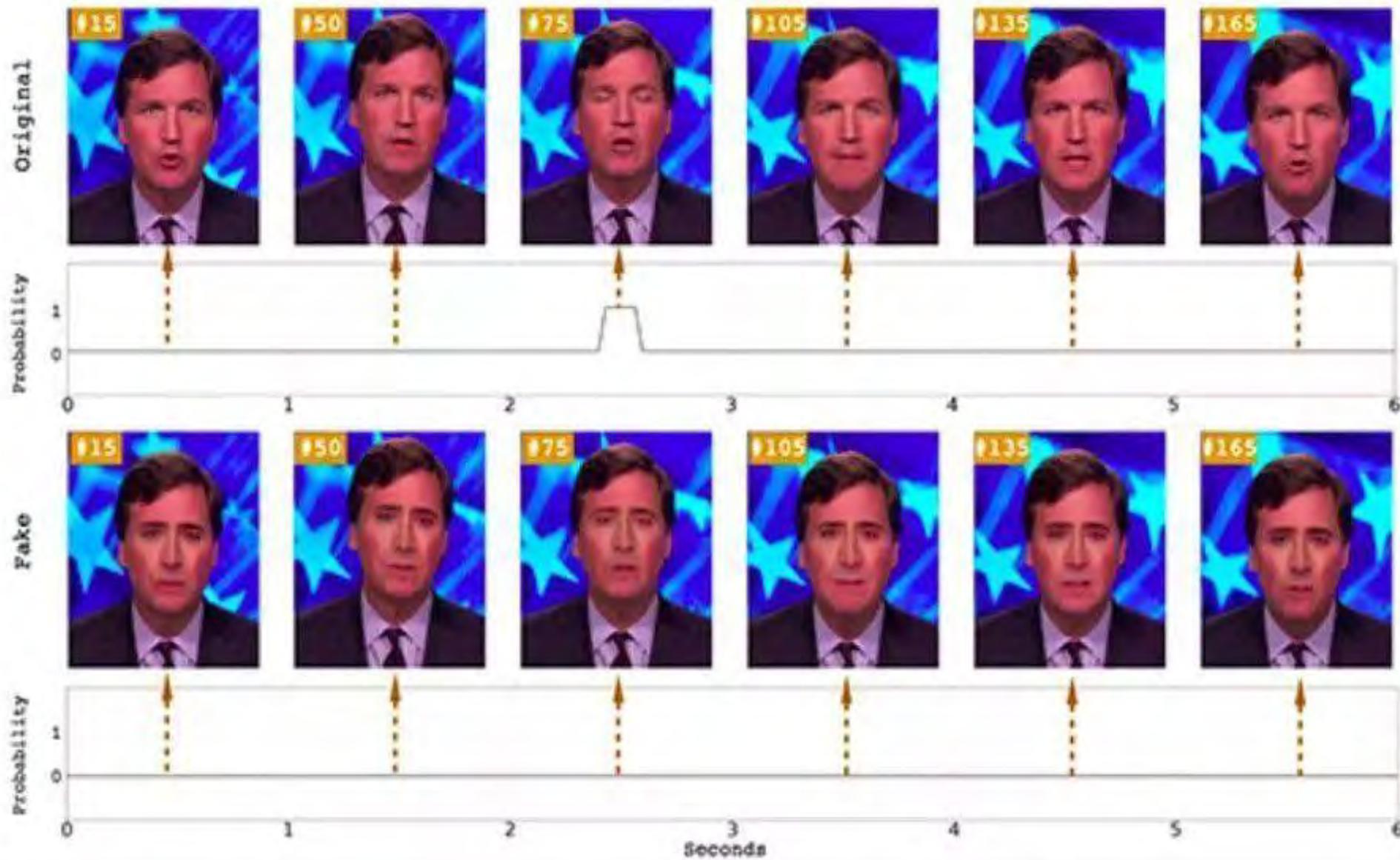
(b) Detected manipulations

(d) Original photo

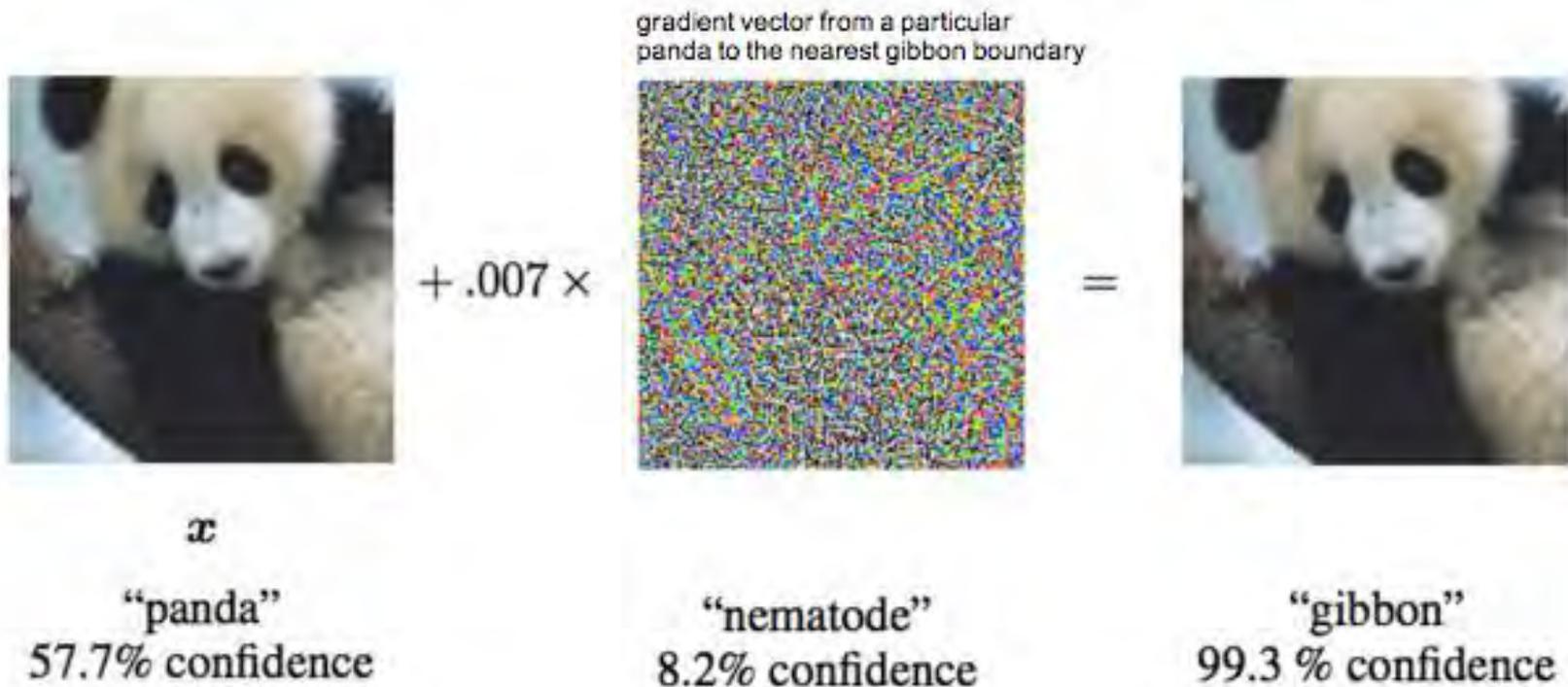
研究人员训练了一个光流场预测模型 F 来预测像素扭曲场 (perpixel warping field), 衡量其与每个样本真实光流场 U 之间的距离 (通过计算原图和修改后的图之间的光流得到)。

机遇与挑战

DARPA Media Forensics



没有完美的算法——深度学习逃逸



熊猫与长臂猿

- 智能推荐算法可加速不良信息的传播
- 人工智能技术可制作虚假信息内容，用以实施诈骗等不法活动
- 算法设计或实施有误可产生与预期不符甚至伤害性结果
- 算法潜藏偏见和歧视，导致决策结果可能存在不公
- 算法黑箱导致人工智能决策不可解释，引发监督审查困境
- 含有噪声或偏差的训练数据可影响算法模型准确性
- 对抗样本攻击可诱使算法识别出现误判漏判，产生错误结果
- 对现有社会伦理道德体系的冲击

2019.3.18斯坦福以人为本AI 研究院 (Stanford HAI) 成立。
斯坦福利用 HAI 评估智能机器对人类生活所造成的影响，包括机器自动化取代了部分人力工作，算法引起的性别和种族偏见，医疗、教育和司法系统中存在的 AI 问题。

- 1) 发展 AI 的过程中必须考虑 AI 对人类社会的影响；
- 2) AI 的应用是为了赋能人类，而非取代人类；
- 3) 人工智能应该更多融入人类智慧的多样性、差异性和深度。



The development of AI should be guided by a concern for its impact on **human society**.



AI should **augment** human skills, not replace them.



AI must incorporate more of the versatility, nuance, and depth of the **human intellect**.

人类 (Human)、赋能 (Augment)、智能 (Intellect)



发展负责任的人工智能：我国新一代人工智能治理原则发布

2019-06-17 16:16 来源：新华社

【字体：大 中 小】 🖨️ 打印 🗣️ 语音 📄 分享 📄 更多

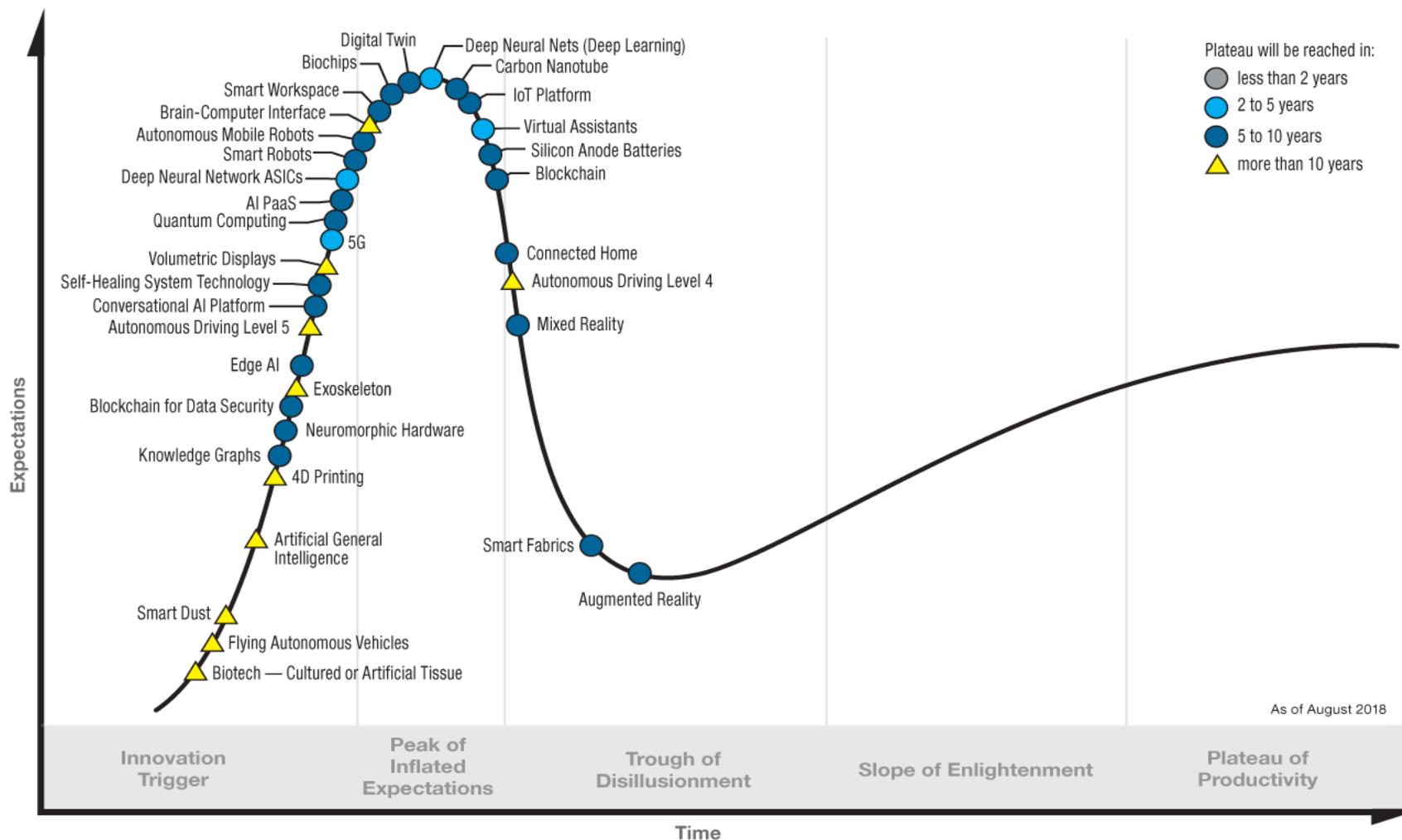
新华社北京6月17日电（记者 胡喆）17日，国家新一代人工智能治理专业委员会发布《新一代人工智能治理原则——发展负责任的人工智能》，提出了人工智能治理的框架和行动指南。

清华大学苏世民书院院长、国家新一代人工智能治理专业委员会主任薛澜介绍，治理原则旨在更好协调人工智能发展与治理的关系，确保人工智能安全可控可靠，推动经济、社会及生态可持续发展，共建人类命运共同体。治理原则突出了发展负责任的人工智能这一主题，强调了和谐友好、公平公正、包容共享、尊重隐私、安全可控、共担责任、开放协作、敏捷治理等八条原则。

和谐友好 公平公正 包容共享 尊重隐私 安全可控 开放协作 敏捷治理

还有2-5年时间?

Hype Cycle for Emerging Technologies, 2018





THANKS

2019 北京网络安全大会
2019 BEIJING CYBER SECURITY CONFERENCE