当DF遇见AI

山东警察学院

张璇

AI：延申人类智能

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

AI：延申人类智能

Artificial Intelligence

像人一样思考的系统	理性地思考的系统
• “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）	• “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**
• “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）	• “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

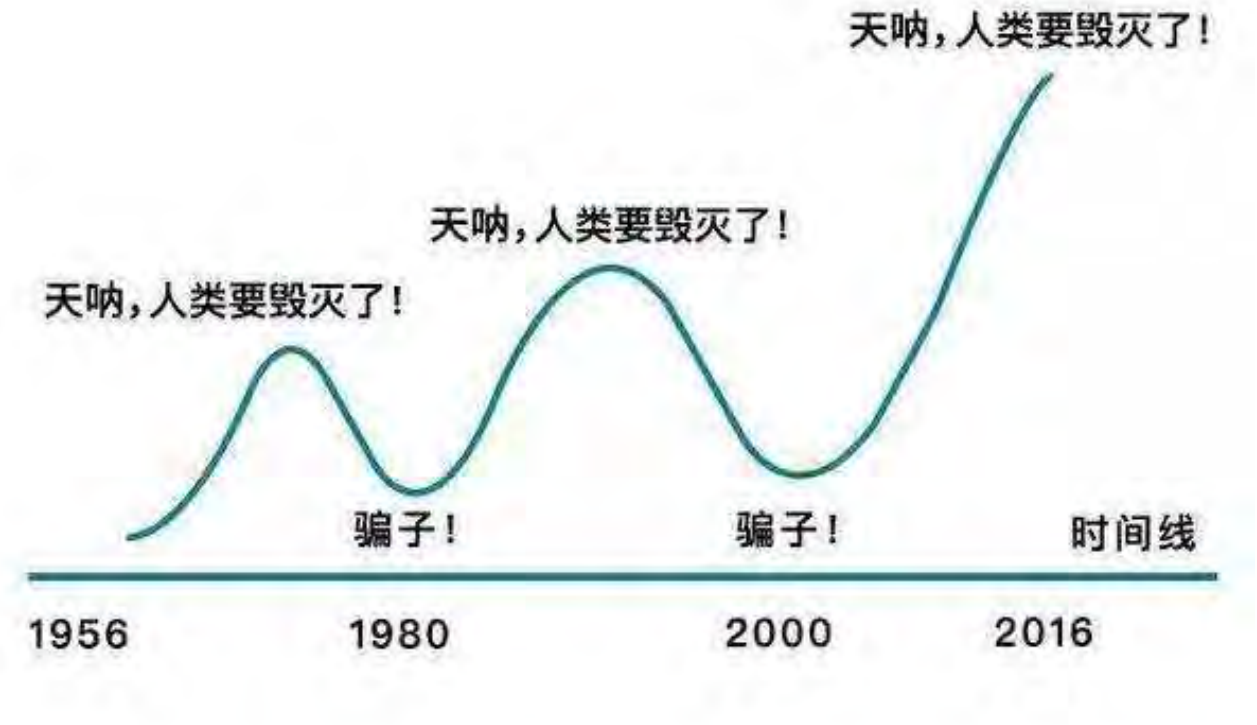
未来人工智能也许会是人类的终结者。——霍金

AI：延申人类智能

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

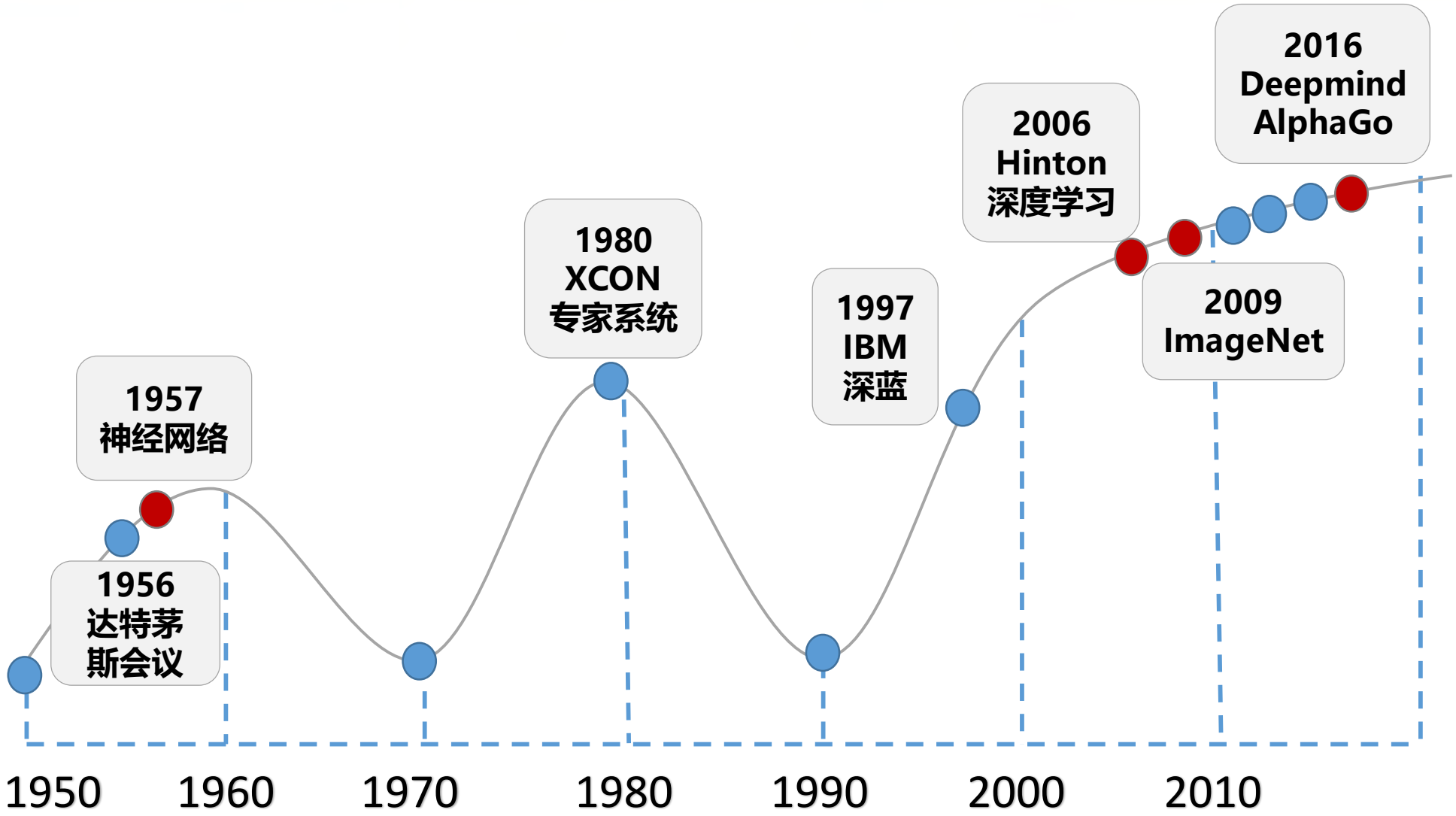
弱人工智能	强人工智能	超级人工智能
特定领域、既定规则中，表现出强大的智能	不受领域、规则限制，具有人类同样的创造力和想象力	超越人类的智能

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

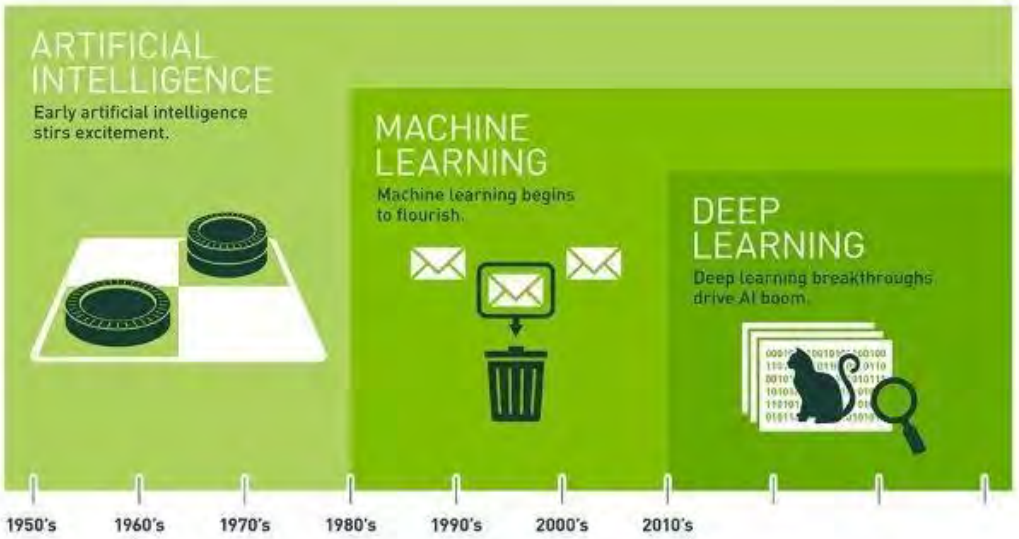
→

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

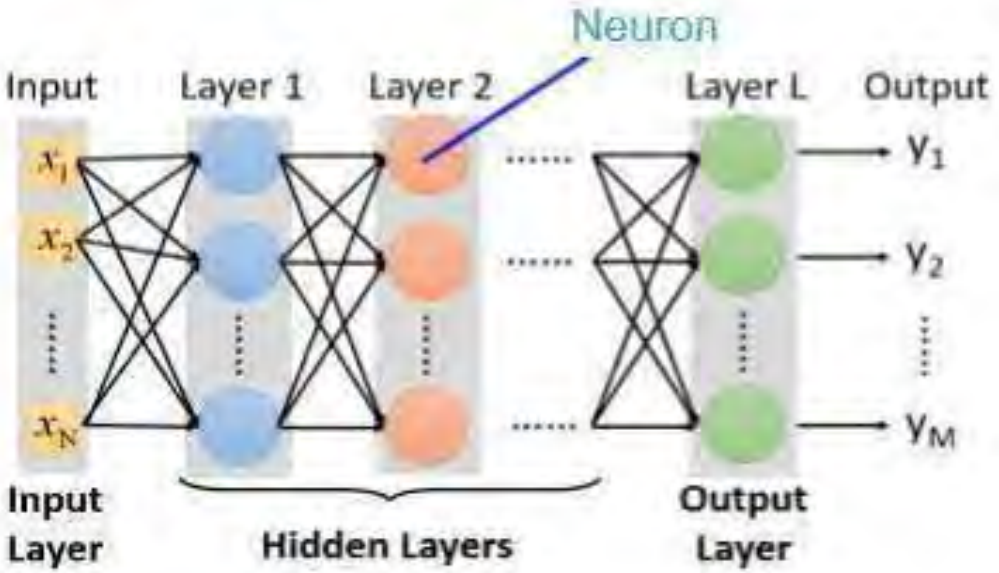
AI：70年浮沉史

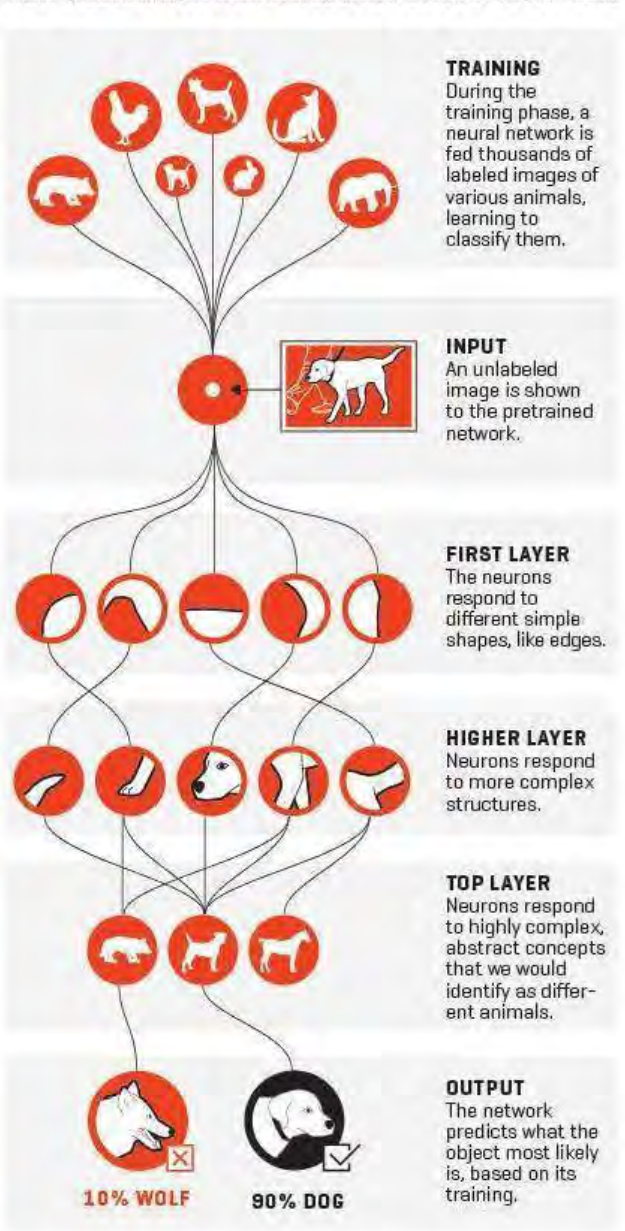

1957
神经网络
1956
达特茅
斯会议
1980
XCON
专家系统
1997
IBM
深蓝
2006
Hinton
深度学习
2009
ImageNet
2016
Deepmind
AlphaGo
1950
1960
1970
1980
1990
2000
2010

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.





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

计算机视觉
语音识别
自然语言处理

自动驾驶
医疗健康
金融
零售
娱乐
AR
VR

核心算法

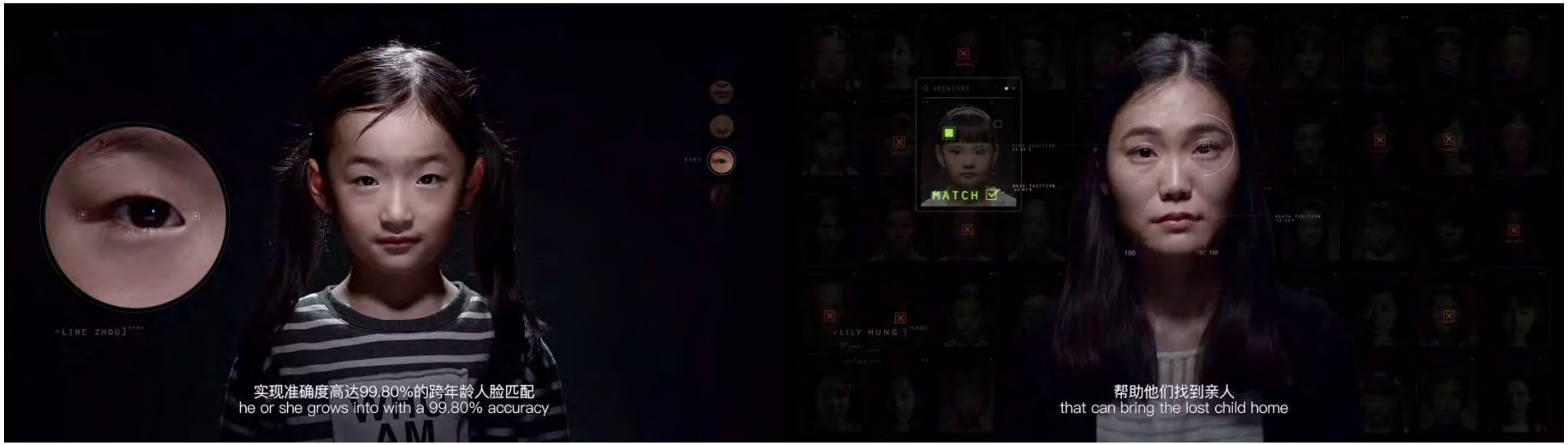
基于经典数学
基于统计模型
基于神经网络

计算能力

海量数据

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

AI：未来已来

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

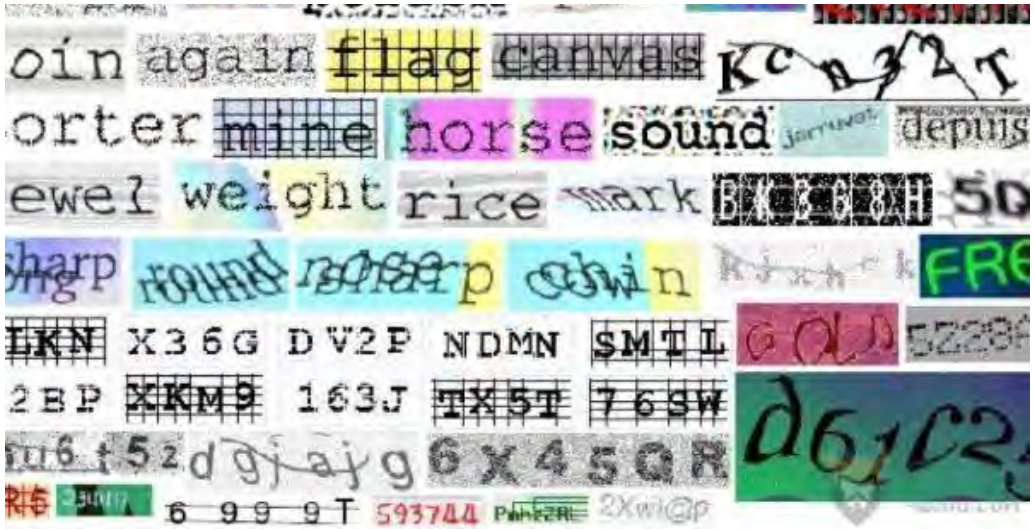
Artificial Intelligence

All in AI
AI in All
All can AI

AI：未来已来

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

中华人民共和国中央人民政府
www.gov.cn

国务院 | 总理 | 新闻 | 政策 | 互动 | 服务 | 数据 | 国情

首页 > 信息公开 > 国务院文件 > 科技、教育 > 科技

索 引 号：000014349/2017-00142　　主题分类：科技、教育\科技
发文机关：国务院　　成文日期：2017年07月08日
标　　题：国务院关于印发新一代人工智能发展规划的通知
发文字号：国发〔2017〕35号　　发布日期：2017年07月20日
主 题 词：

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

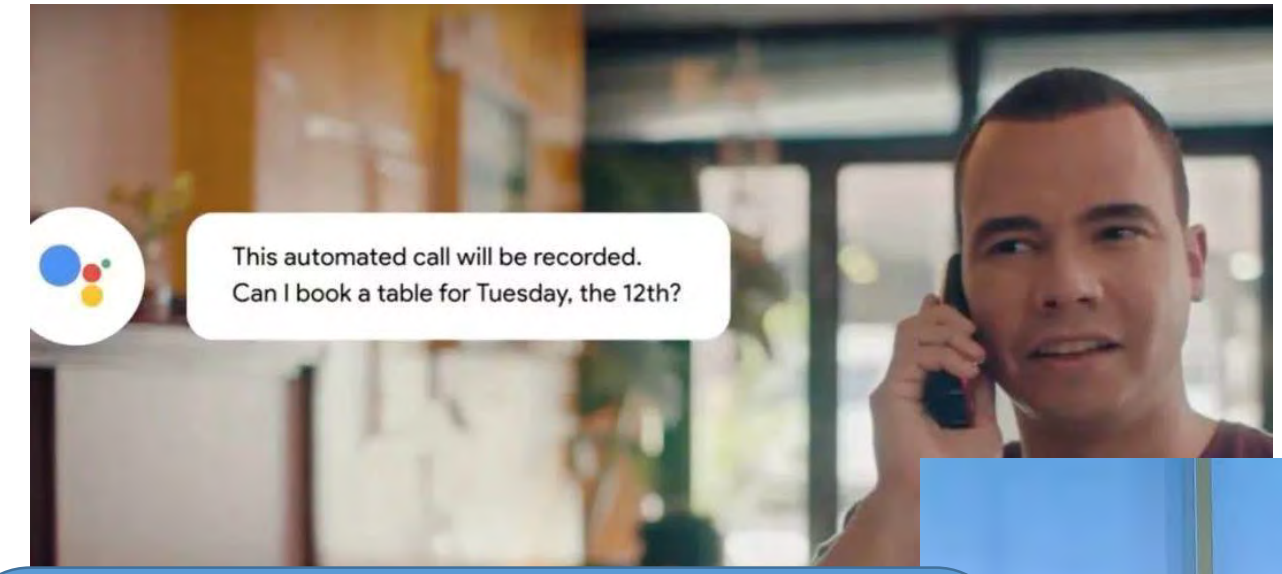
制造业创新中心建设工程　智能制造工程　工业强基工程　绿色制造工程　高端装备创新工程

中国政府网制作　　设计：刘啸萱 设计：徐菁

AI：理想与现实的距离

Artificial Intelligence

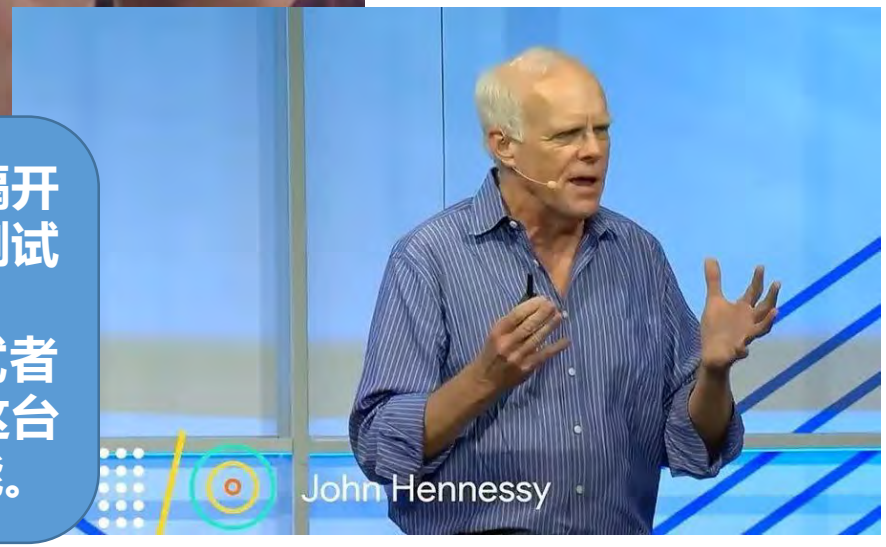
如何知道一个系统是否具有智能?
1950年，计算机科学家图灵提出了著名的“图灵测试”。



Duplex

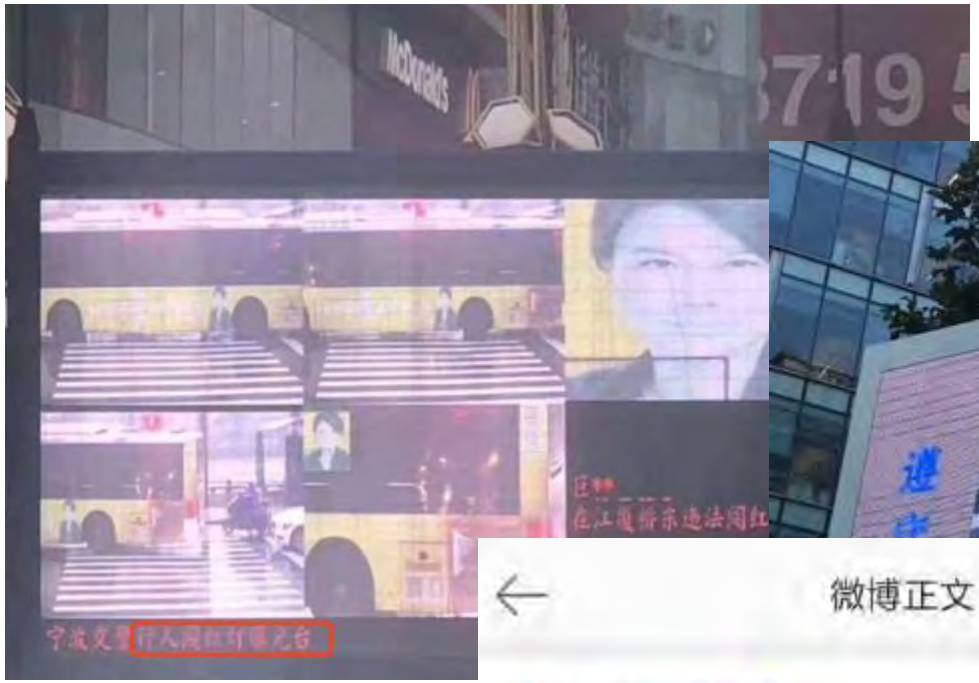
谷歌造假！通过图灵测试的AI实测：4次成功完成任务，3次靠的是人工

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？原来是一场乌龙一

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

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

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

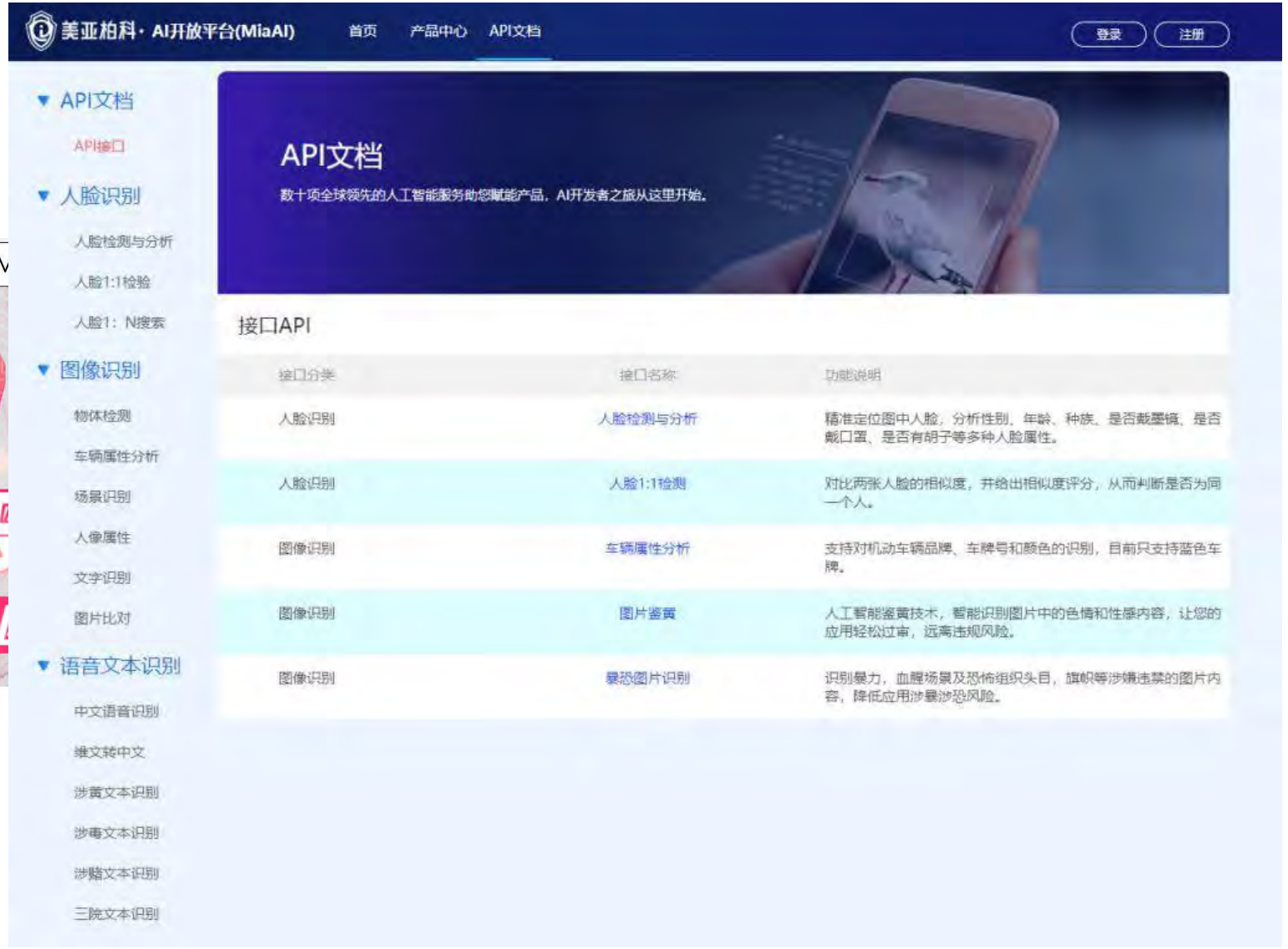
语言级　　工具级　　智慧级

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
国家市场监管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

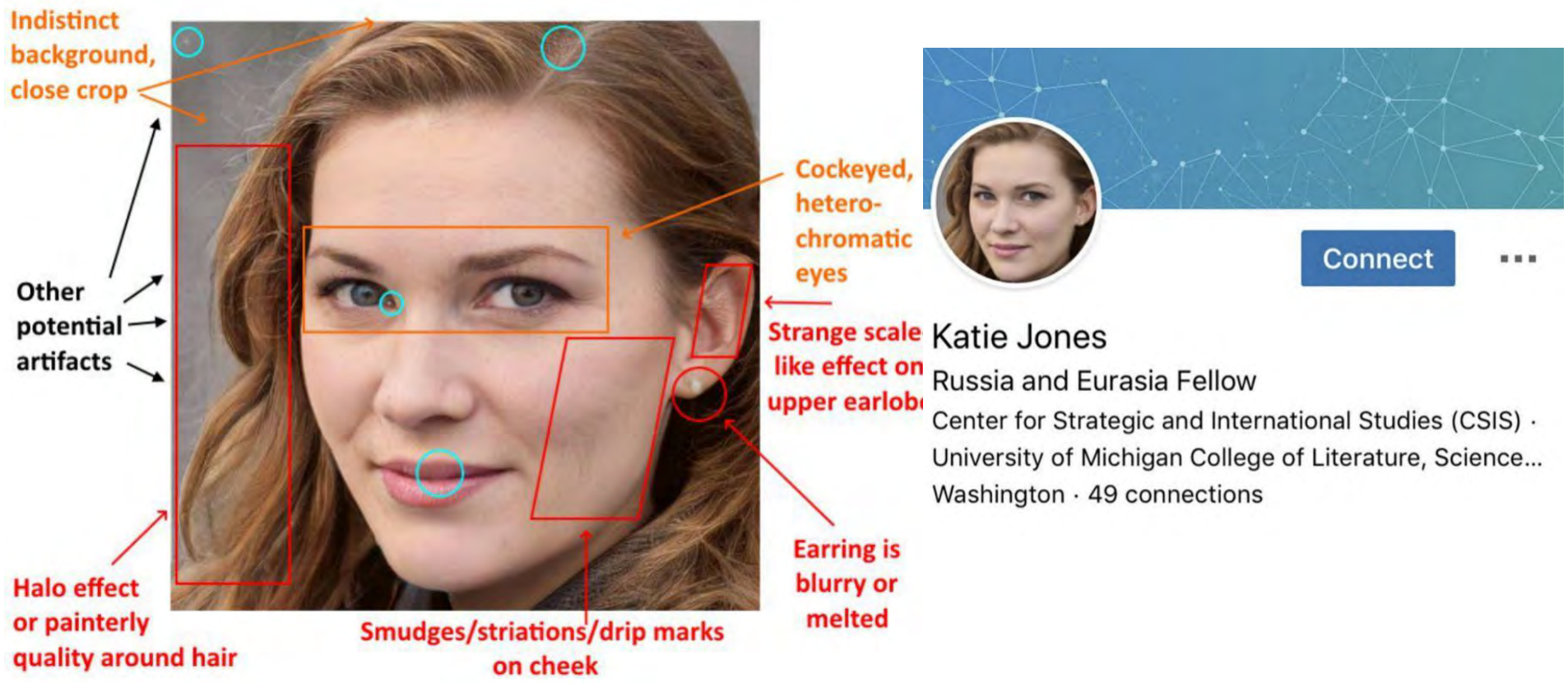


当电子数据取证遇见AI

- 语音识别
- 情景预测
- 人脸检测
- 图像识别（鉴黄）

机遇与挑战

可能改变鉴定规则
鉴定将更加依赖技术
训练样本缺乏



间谍用GAN生成假头像

机遇与挑战

图像、视频篡改、伪造

DeepFakes

http://deepfakes.com.cn/

https://github.com/deepfakes/faceswap

机遇与挑战

笔迹鉴定

GAN 可以被用来学习生成各种各样的字体

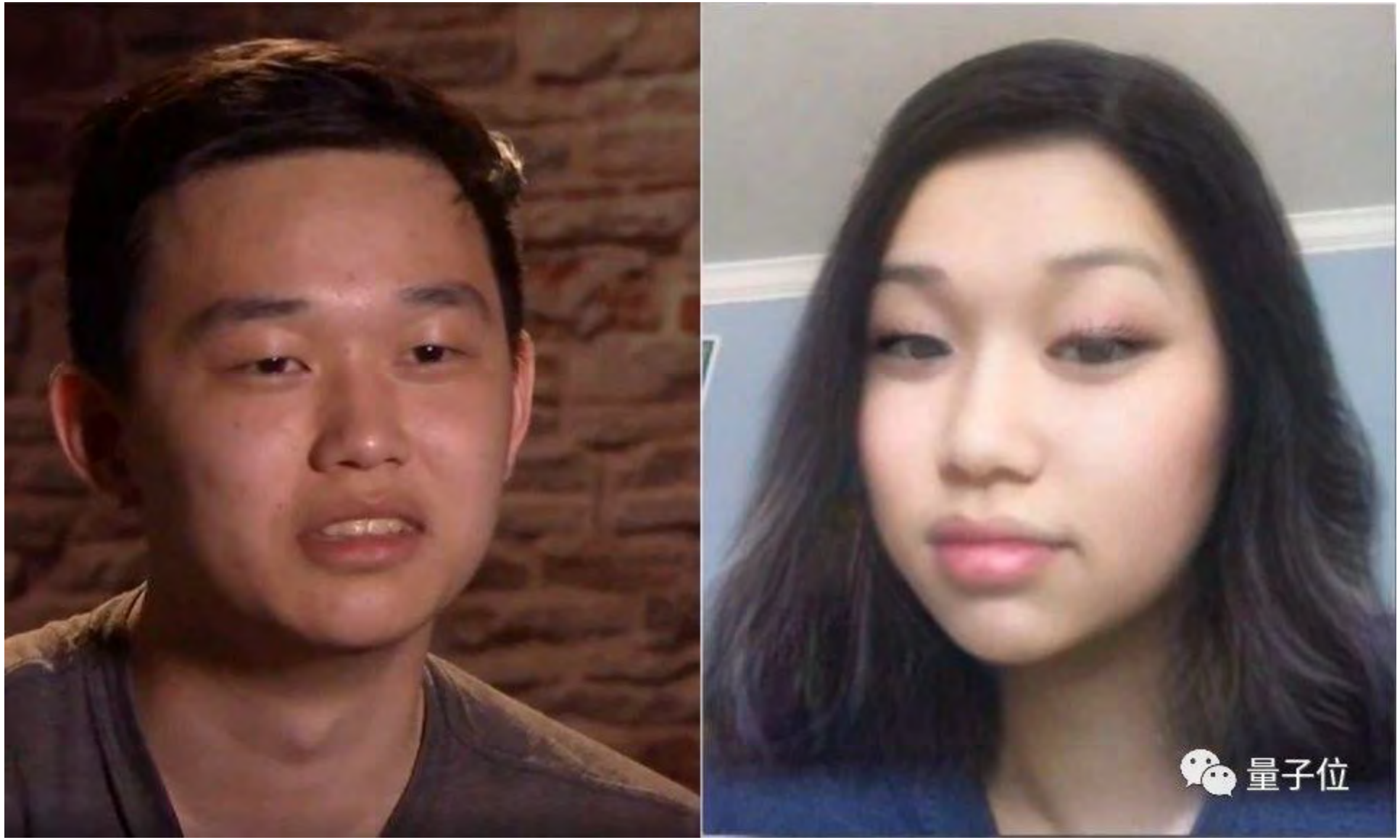
机遇与挑战

CycleGAN

一键变性

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

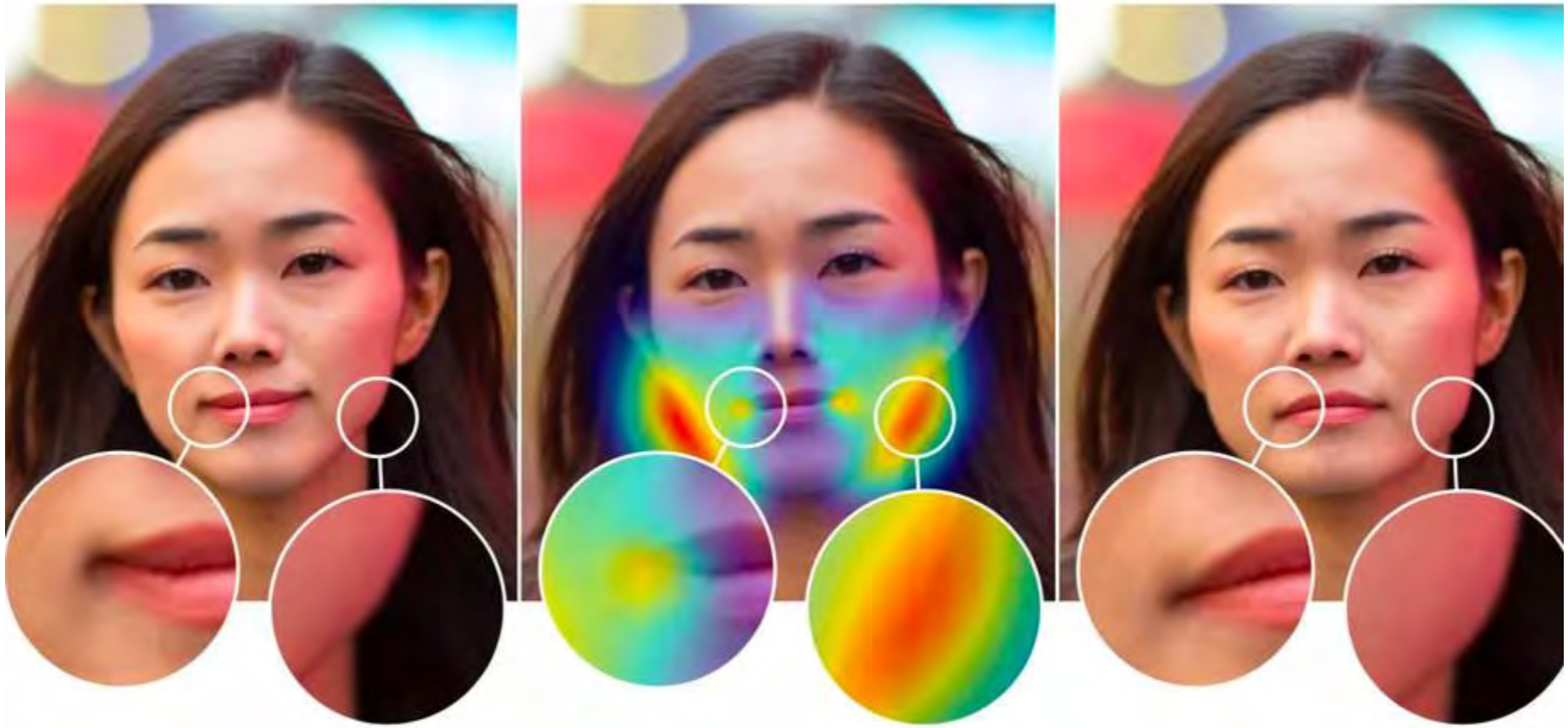
机遇与挑战



用AI变身16岁女孩，抓住40岁违法警察！

机遇与挑战

解铃还需系铃人

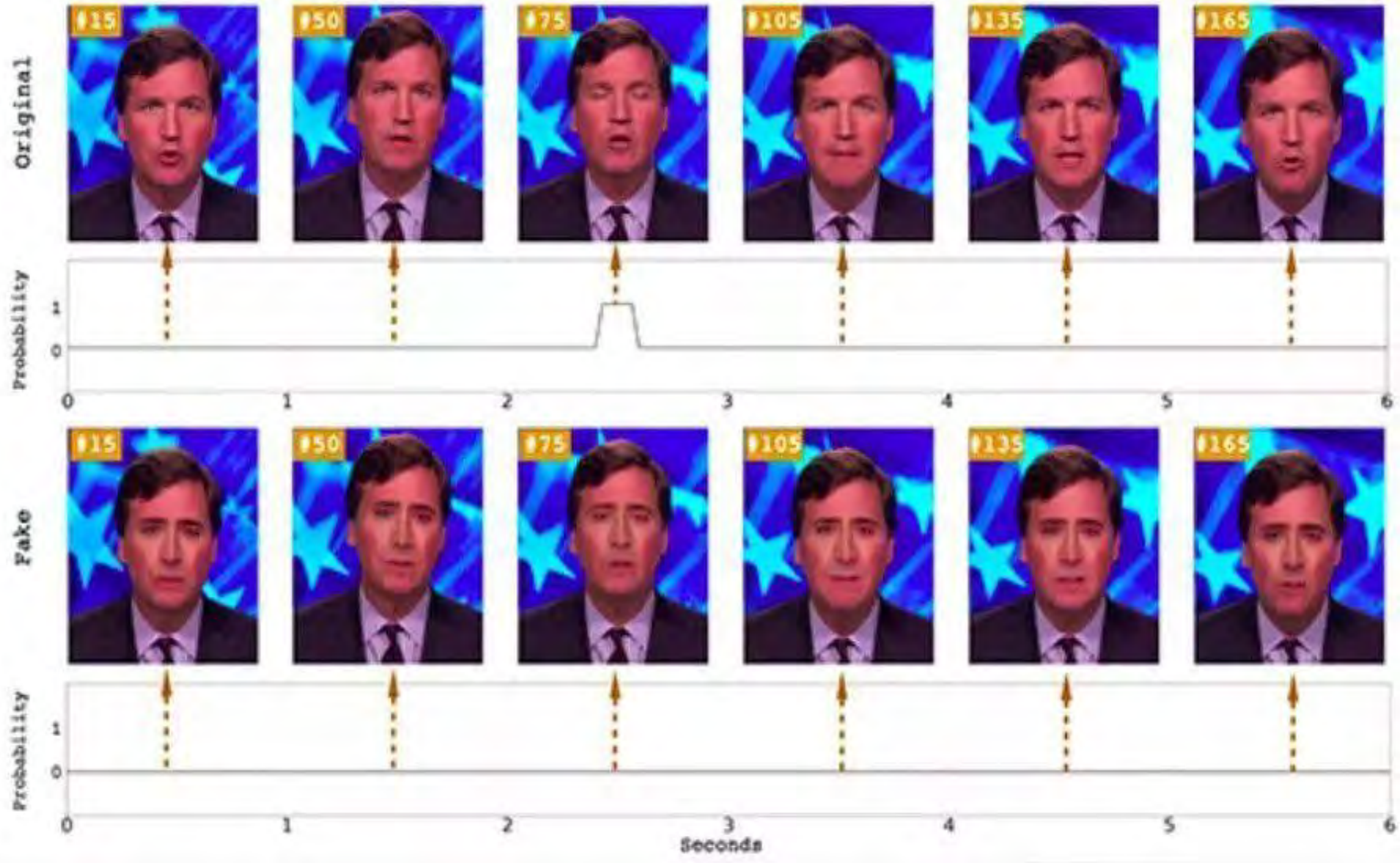


(a) Manipulated photo (b) Detected manipulations (d) Original photo

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）。

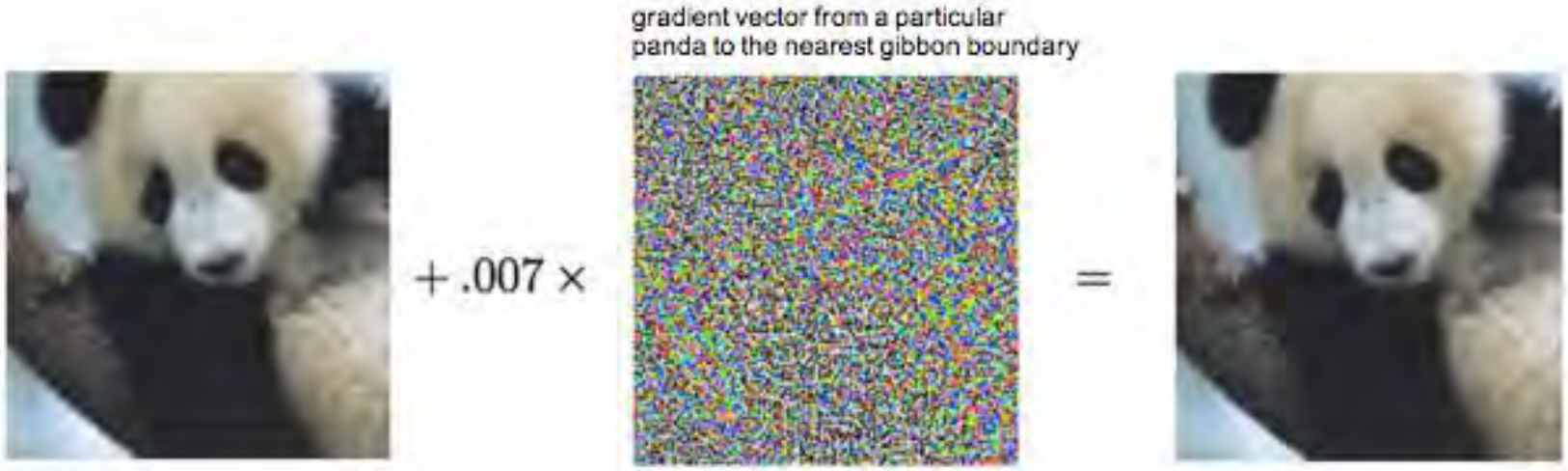
机遇与挑战

DARPA Media Forensics



机遇与挑战

没有完美的算法——深度学习逃逸



熊猫与长臂猿

机遇与挑战

- 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

机遇与挑战

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。

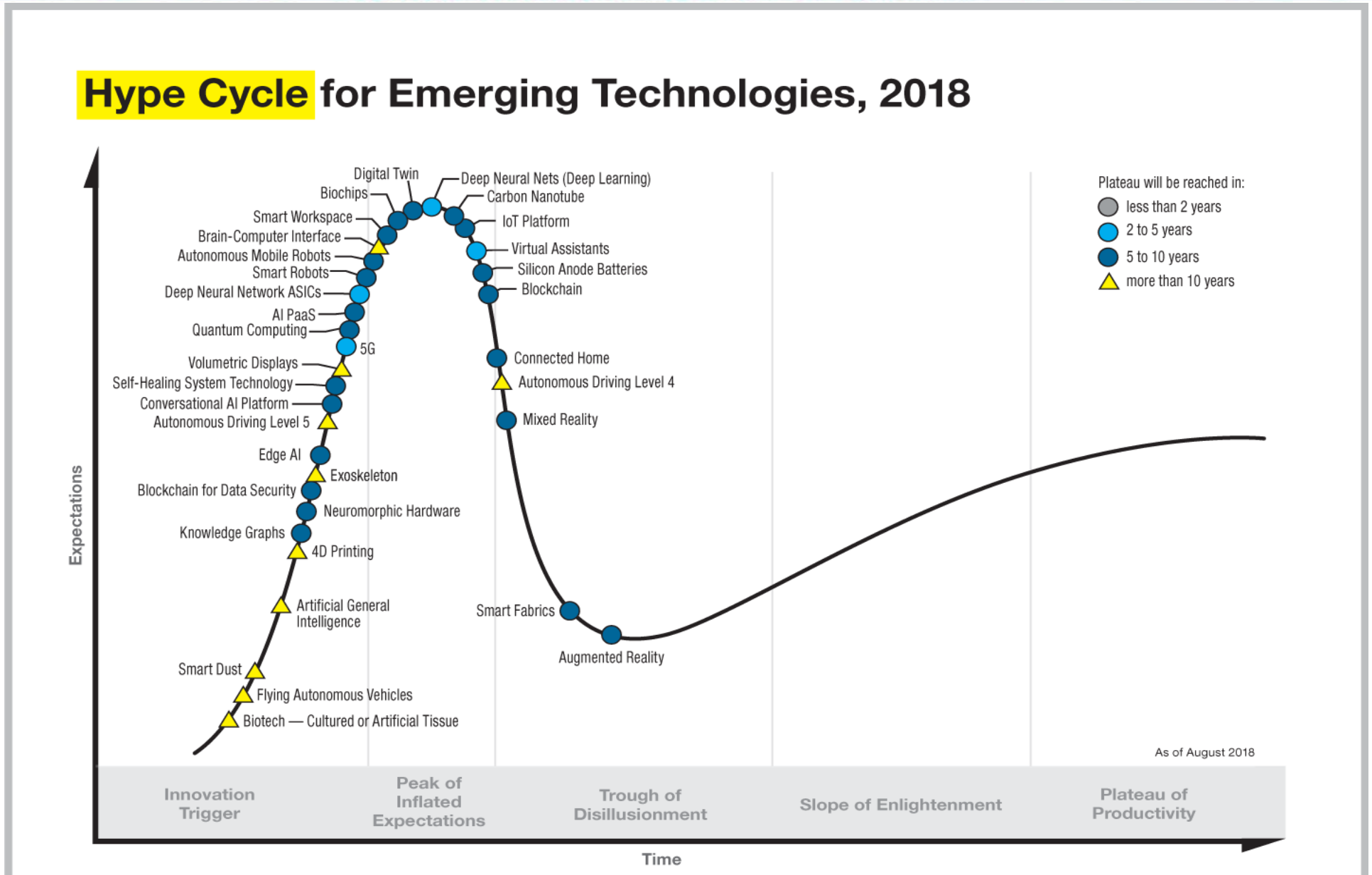
1）发展 AI 的过程中必须考虑 AI 对人类社会的影响；
2）AI 的应用是为了赋能人类，而非取代人类；
3）人工智能应该更多融入人类智慧的多样性、差异性和深度。

人类（Human）、赋能（Augment）、智能（Intellect）

机遇与挑战

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
Digital Twin
Biochips
Smart Workspace
Brain-Computer Interface
Autonomous Mobile Robots
Smart Robots
Deep Neural Network ASICs
AI PaaS
Quantum Computing
5G
Volumetric Displays
Self-Healing System Technology
Conversational AI Platform
Autonomous Driving Level 5
Edge AI
Exoskeleton
Blockchain for Data Security
Neuromorphic Hardware
Knowledge Graphs
4D Printing
Artificial General Intelligence
Smart Dust
Flying Autonomous Vehicles
Biotech — Cultured or Artificial Tissue
Deep Neural Nets (Deep Learning)
Carbon Nanotube
IoT Platform
Virtual Assistants
Silicon Anode Batteries
Blockchain
Connected Home
Autonomous Driving Level 4
Mixed Reality
Smart Fabrics
Augmented Reality
Plateau will be reached in:
less than 2 years
2 to 5 years
5 to 10 years
more than 10 years
As of August 2018
Expectations
Innovation Trigger
Peak of Inflated Expectations
Trough of Disillusionment
Slope of Enlightenment
Plateau of Productivity
Time

THANKS

2019北京网络安全大会

2019 BEIJING CYBER SECURITY CONFERENCE