

北京城市实验室（BCL）2017年年会

从信息化走向智能化

李 栋

北京清华同衡规划设计研究院 技术创新中心

2017年06月03日

Q1 如何定义 “信息化”

- 规划大数据分析是否属于规划信息化范畴？
- 信息化工作服务的主体是谁？客体是谁？
- 从信息化到智能化距离还有多远？

两种场景

- 流程信息化 / 狭义信息化

- 既有机制下对业务的信息化改造，如委办局的办公OA.....

- 决策智能化

- 采集整理决策所需信息、设计多种决策方案、对方案开展模拟与预评估、对方案进行比选优化，对方案实施开展周期性监控、方案实施后评估等，如城乡规划编制与实施.....

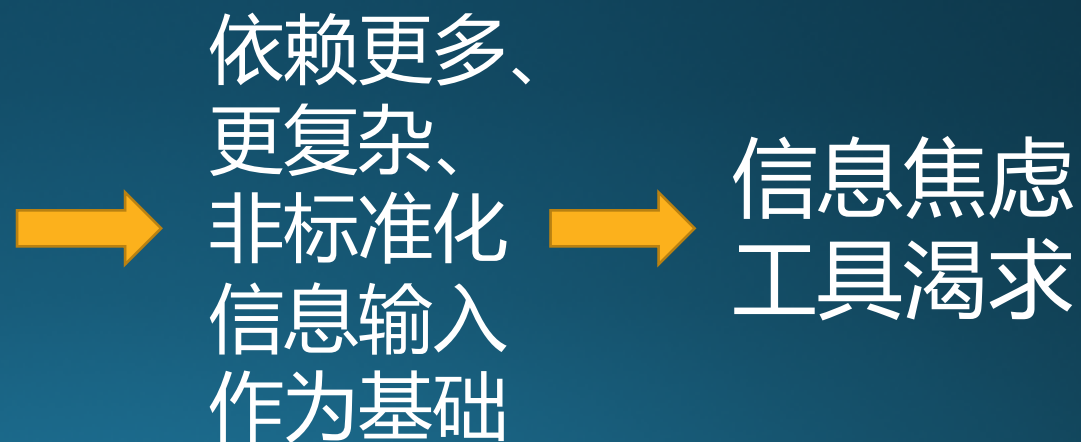
差异所在

- 流程信息化

- 规则相对清晰、条件相对封闭、目标相对明确

- 决策智能化

- 规则相对模糊、条件相对开放、目标不够明确



一个更大的、数据的视角

- 流程信息化

- 自有数据解决自己的问题

- 决策智能化

- 自有数据+他类数据 (alternative data)
解决自己的问题

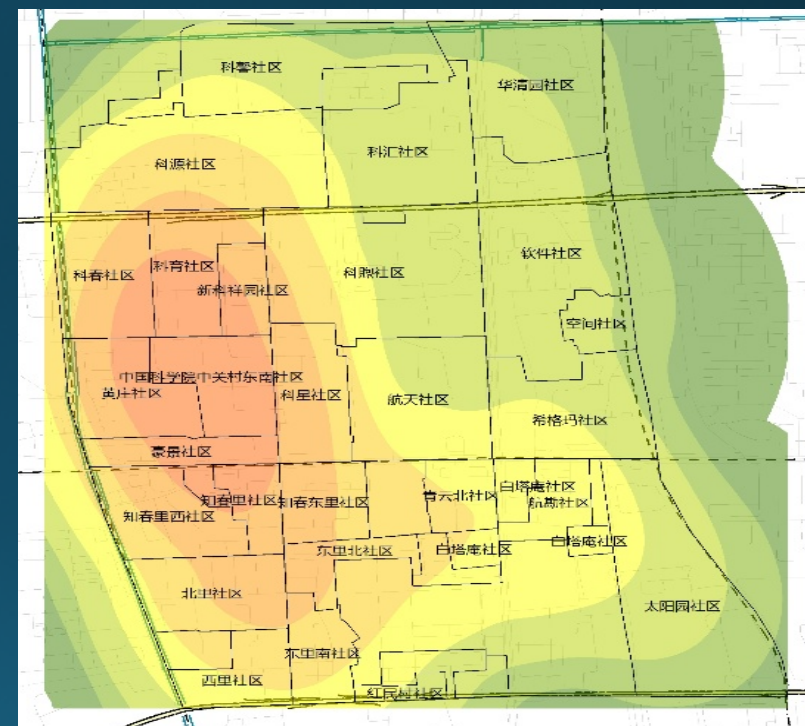
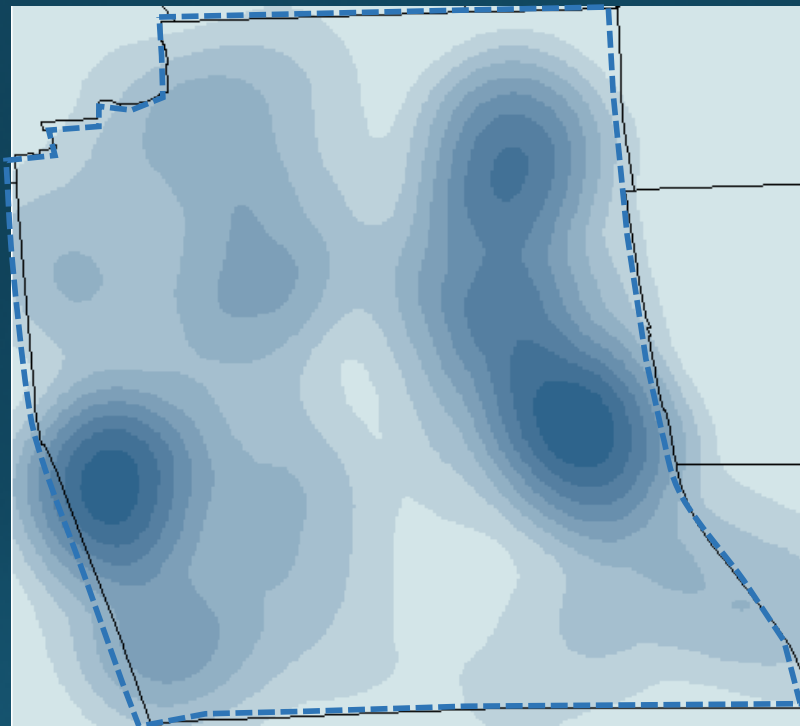
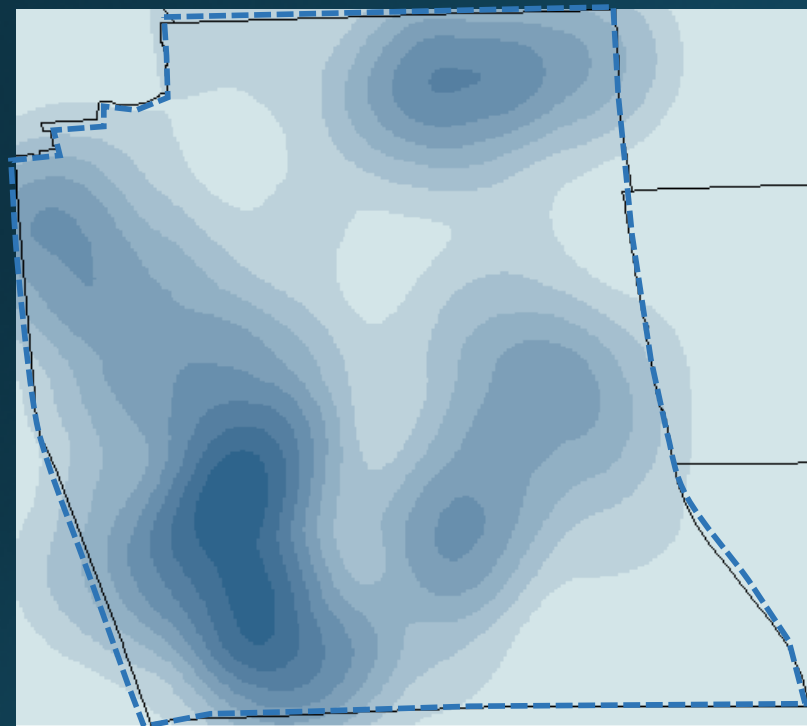
一个更大的、数据的视角

- 他类数据 (alternative data) 如何产生的？
 - 全社会的、多部门的数字化、信息化趋势

数据应用环节

1. 感知：数据清洗收集
2. 测度：数据分析评价
3. 挖掘：数据综合解读

感知：数据清洗收集



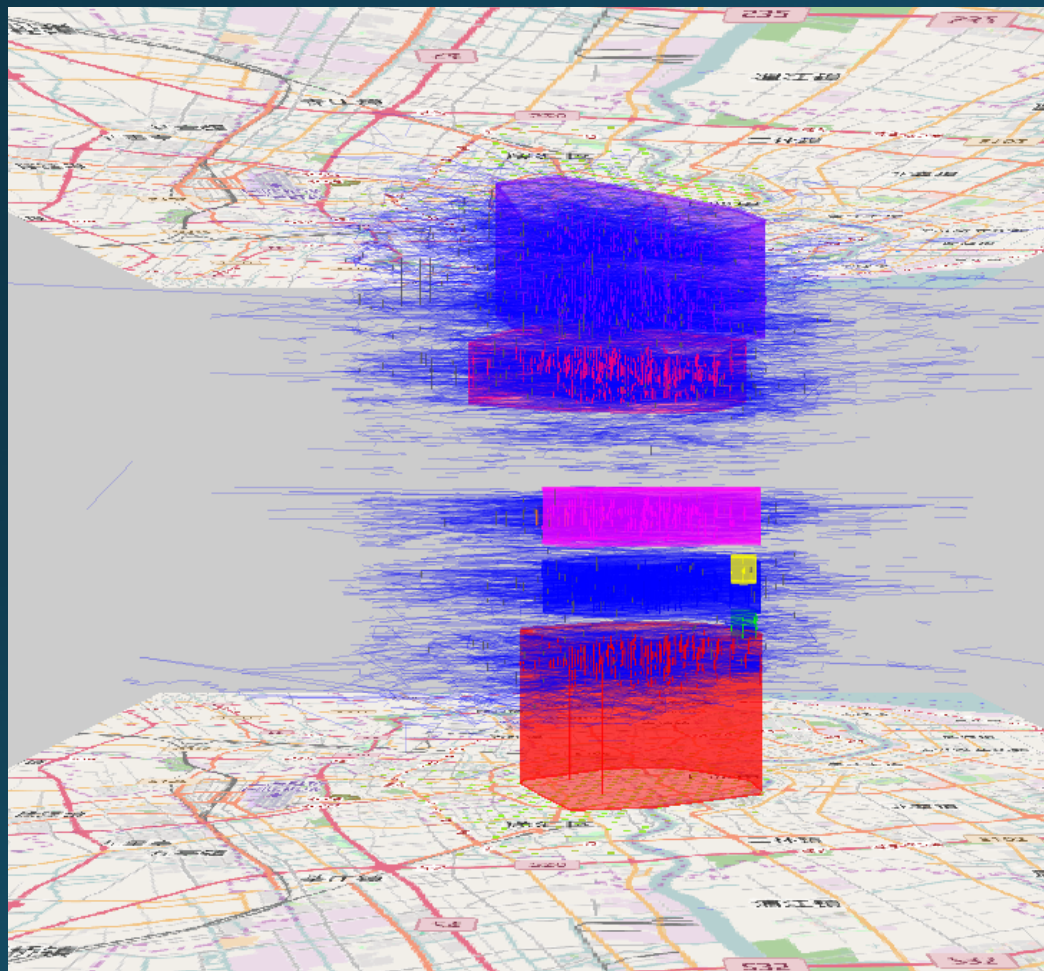
中关村街道住户、居住人口、老人卡用户三类空间分布

测度：数据分析评价

通州多种职住平衡指标的设计与对比（基于手机信令）

指标		计算方法	通州区	核心区	最小值	最大值
区内就业人口 (万人)		就业地位于通州区内的人口数量（含居住地在通州区外）	73.6	27.9	0.08 新华街道	11.39 马驹桥镇
区内居住人口 (万人)		居住地位于通州区内的人口数量	168.25	59.8	0.43 新华街道	24.03 马驹桥镇
昼夜人口比		白天与夜晚瞬时人口比 选取时间段：白天14:00，夜晚：4:00	0.87	0.91	0.81 西集镇	0.98 中仓街道
职住平衡状况	平衡系数 (数量上)	区内就业人口与区内居住人口之比	0.44	0.47	0.18 新华街道	0.64 永顺地区
	本地就业率 (质量上)	区内居住人口中，就业地位于通州区外的人口占就业人口的比重	61%	58.42%	47.94% 北苑街道	80.73% 西集镇

挖掘：数据综合解读



采用时空聚类算法识别共享单车时空轨迹热点

广义信息化需求越来越旺盛

- 数据可用性极大增加：超30亿人、约170亿设备联网在线
 - 人类从直立行走走到2003年所创造的数据量，总计5艾字节*
 - 到2007年，短短几年数据量已超过了300艾字节
 - 到2015年，全球网络数据流量接连翻番，数据量达到966艾字节
 - 到2025年，全球数据总量将增长至170,000艾字节
 -



* 1艾字节 (exabyte) 相当于10亿GB

广义信息化需求越来越旺盛

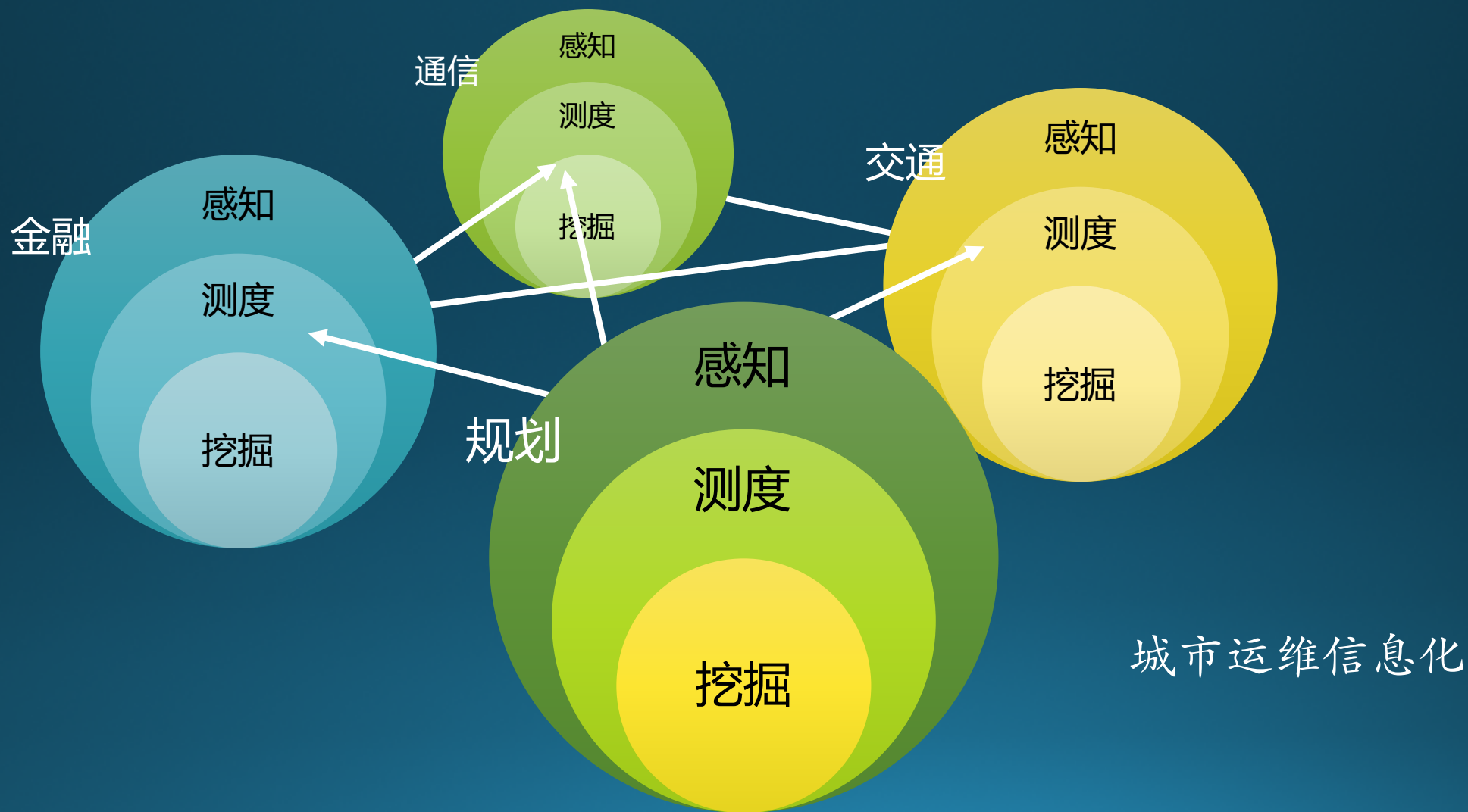
- 计算能力得到强化

- 强大的计算机通过网络连接，可以处理大量数据的机器学习成为可能

- 算法得到全面创新

- 十余年机器学习发展基础上，可以将这些庞大的数据用于学习算法的训练，开发新规则来执行日益复杂的任务。以深度学习为代表的新的技术得到持续优化，启发了新服务，刺激了对AI领域其他方面的投资和研究。

广义信息化需求越来越旺盛

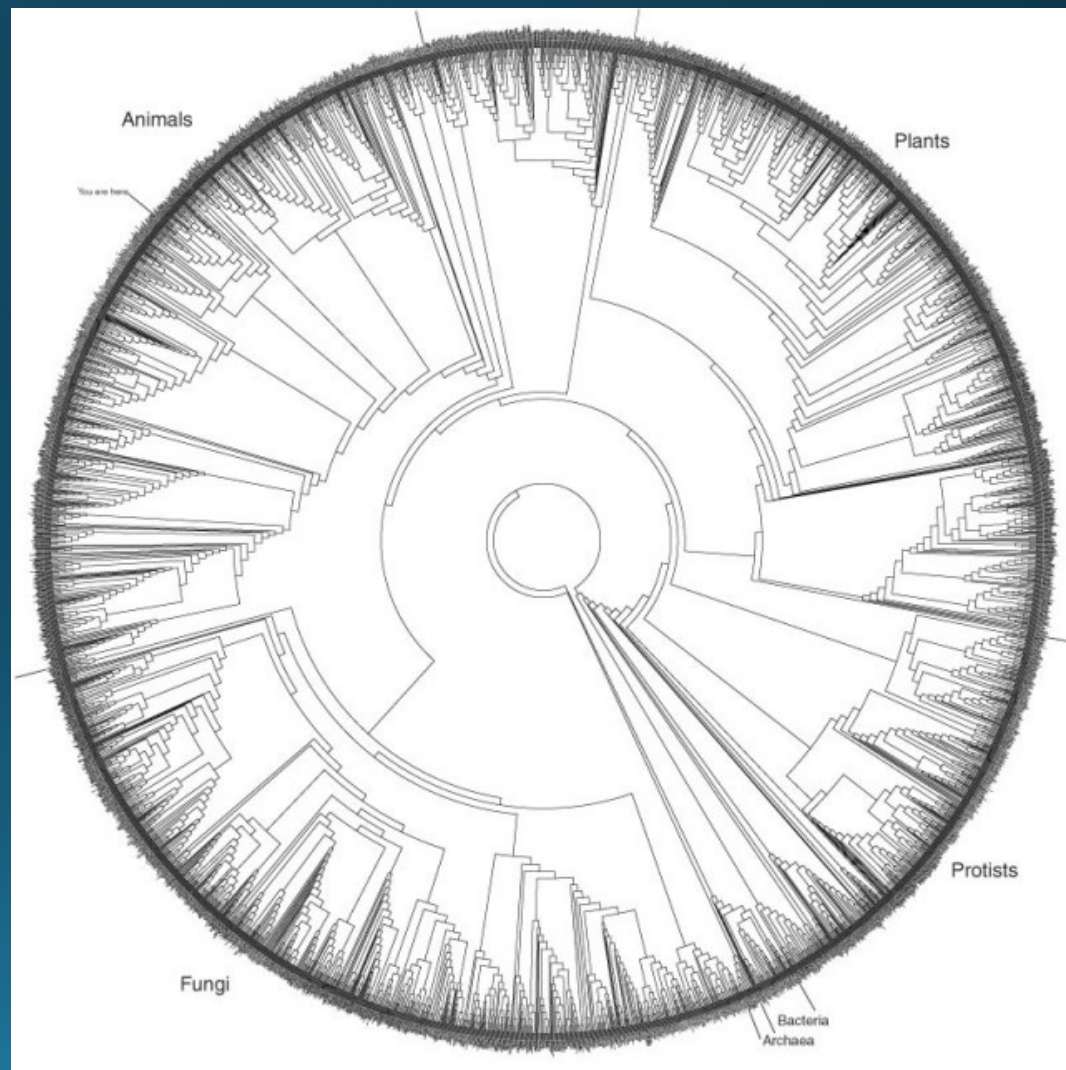


Q2 是否存在终极的 “智能化” 银弹？

- AI更聪明吗？
- 通用AI还是专用AI？
- 偏差、公正及其他陷阱？
- 如何协同？

AI更聪明吗？

- AI只是和人类不一样而已
 - 智能是多维度的，不存在“强弱”的单一维度
- AlphaGO依赖的是蛮力
 - 基于反馈进行优化后的穷举法



通用AI还是专用AI（强AI/弱AI）？

- 专用AI

- 在某一领域（如语言翻译、图像识别）
执行特定任务的AI

- 通用AI

- 尚未实现的理念假设，不受领域约束，
在所有领域都可以学习并执行任务



城市规划管理AI（如有）
是专用还是
通用？

偏差、公正及其他陷阱？

算法并没有让事情变得更加公平.....不过
就是让现实里存在的偏见变得自动化了。
——“Weapons of Math Destruction”

- 透明度、偏见和责任
 - AI做出的决策可能会对人们的生活产生重要的影响
 - AI所依赖的信息过于简化
 - AI可能歧视某些个人或由于带有偏见的训练数据而造成错误
 - AI的决策往往难以理解/解释，使得偏见的解决和追责问题变得更加困难

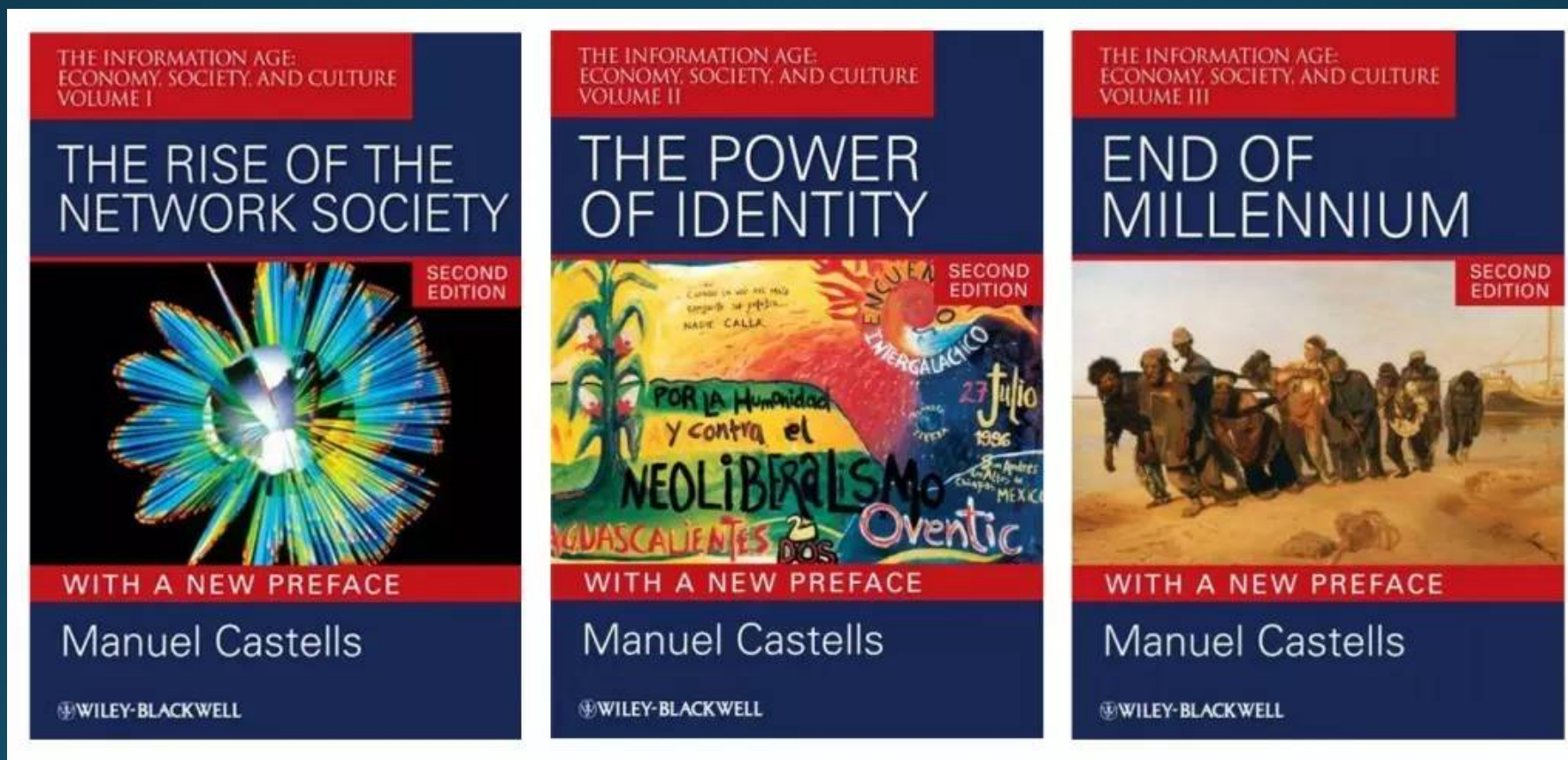


通过预测模型，左边的犯罪风险为高，右边为低。但事实并非如此

偏差、公正及其他陷阱？

- 信息成为生产资料产生数据资本主义 (Datenkapitalismus)
 - 以收集、整理、储存和加工数据为名，行垄断数据之实：名利、话语权
- 数字鸿沟 (digital divide)
 - 在信息技术和工具的占有和使用上所产生的一种普遍的社会不公，在AI中可能被持续放大
- 数据化便利 = 数据化监控
 - 私人领域被侵蚀，人类社会可能发生结构动摇

偏差、公正及其他陷阱？



卡斯特尔（Manuel Castells）“信息时代三部曲”
《网络社会的崛起》、《认同的力量》、《千年终结》

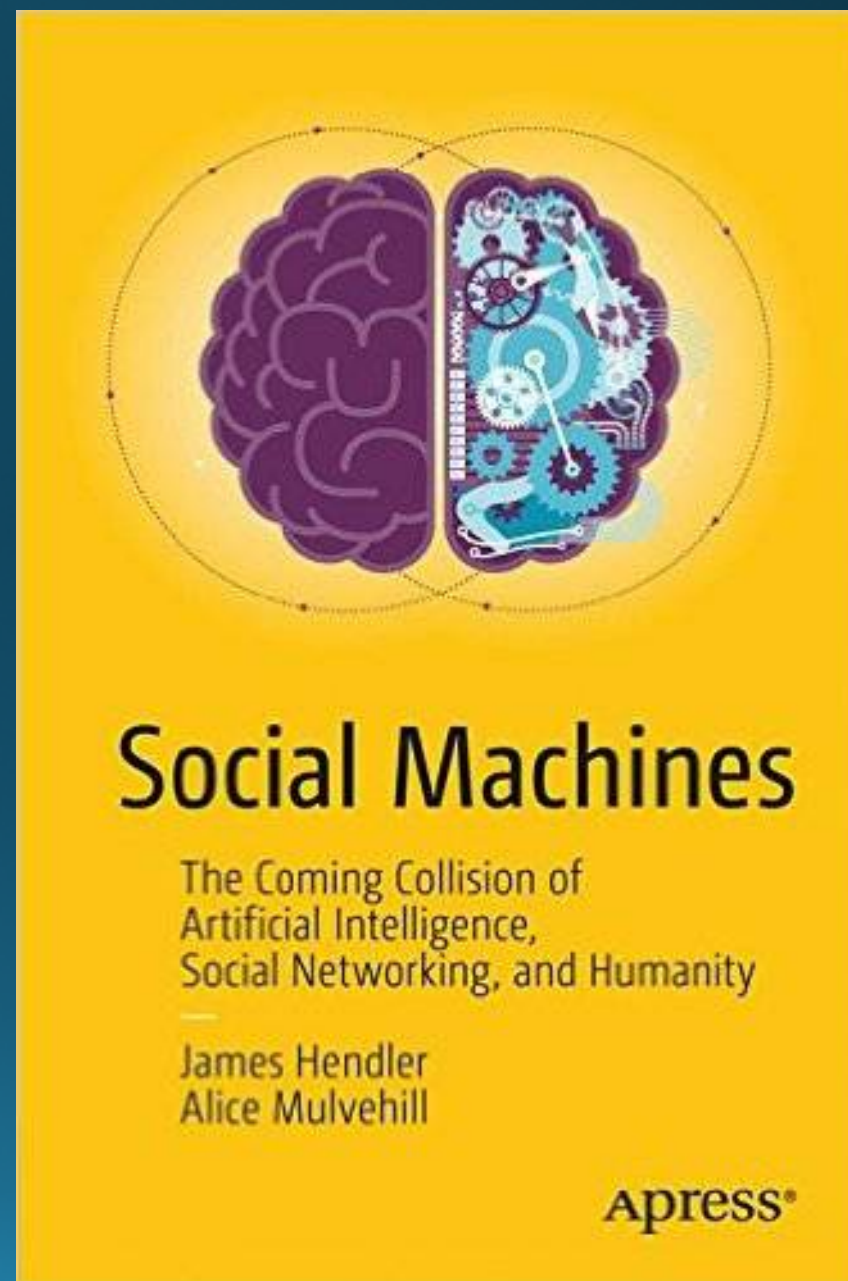
如何协同？

- 人类的优势：
 - 好奇心：基于大量无用联想的创造力
 - 但并非“思考主义”（thinkism）
- AI的优势：
 - 划定边界、给定规则、确定目标条件下，高效能地（甚至是带有启发式的）计算并搜寻答案



如何协同？

- Social Machine：人和机器构成的一个完整系统
 - 人擅长做有创造性的工作
 - 机器擅长做机械性、反复的重复劳动



数据应用的SMART原则

- **S**trategy
- **M**easures / metrics
- **A**nalyze
- **R**eport
- **T**ransform



Value