

# | 数据隐私

北京理工大学  
邱万勇

# C CONTENTS

1

信息与大数据伦理问题

2

数字身份困境

3

大数据时代的个人隐私

4

数据权利

# 大数据时代新的伦理问题

## ○ 数据安全

- 从采集、存储、关联计算、发布到交易、存档，全流程中数据能不能保证不被盗、不被破解、不被篡改、不主动泄露？安全、可靠？

## ○ 身份困境

- 数字身份与社会身份，可以分离还是必须关联？

## ○ 隐私边界

- 当你听说“相比遭遇恐怖袭击、破产和财产被盗，美国人更担心网络在不经意间泄露了自己的隐私”，你怎么理解大数据时代个人隐私？法律该如何提供保护？

## ○ 数据权利

- 大数据是资产吗？在“我”、BAT等企业、政府、他人或公众之间，拥有权、采集权、使用权、处理权、交易权、分红权，这一整套的数据权利成立吗？可以定价吗？符合伦理吗？

## ○ 大数据公共治理

- 政府主导的公众数据是否应当无条件开放共享？基于大数据的公共治理创新如何才能避免歧视、不当得利、威胁个人自由？

# C CONTENTS

## 大数据伦理新问题：新型数据安全

# 新型数据安全

传统的数据安全：加密、传输、存储等

新型的数据安全：有意识知道的（自己提供的，如购物）

无意识的被记录被采取（移动基站、应用 App 记录信息、医疗机构、服务公司等等私自获取利用）



# 新型数据安全风险

- 个人数据 “被提取” “被记录” “被滥用” “被关联处理”，其获取过程无意识，使用边界不清晰，常超出客户最初授权范围，综合信息、敏感信息安全风险
- 网络条件下，各种应用系统被“撞库”成功后的数据泄露风险
- “云端”安全管理与审计体制安全风险

# 数字身份困境



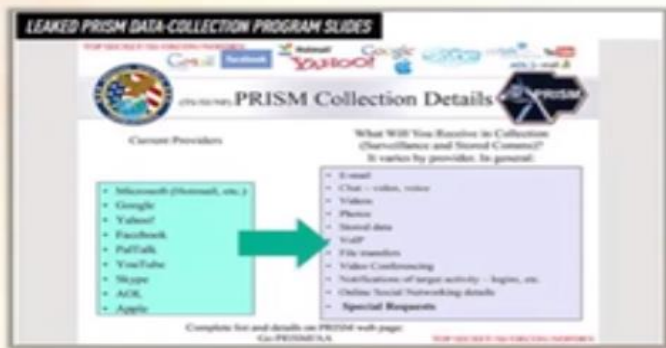
- 数字身份特点
  - 多样、可变、允许匿名/假名
  - 易被盗用，易被追溯
- 匿名还是实名？

## 辩论实名制

群体	利	弊
用户	网络发表言论时更加谨慎、更加合乎法律、道德规范；更利于青少年习得良好的社会行为	言论自由受到限制；个人数据泄露后隐私权、名誉权、财产权受到伤害的风险增加；接受不当个性化推送服务的频次增加
网络服务/运营商	更易于管理和运行，如向未成年人拦截不适合的网络游戏、暴力内容；更利于开展精准商业服务	服务吸引力受影响（如失去用户、失去粘度），进而减损价值；对信息和网络安全的投入要大大增加
政府	更利于提供精准公共服务；更利于减少网络不良信息，使得言论空间更加晴朗；利于青少年和知识水平不高的网民的生存、学习和成长；侦查和惩治网络犯罪更快	便于实施类似“棱镜门”计划，而失去部分公民的信任；“寒蝉效应”使言路闭塞
他人	发生被不当“人肉”时易于找到事主并追责；被有意无意网络侵权的风险降低	盗取、兜售或伪造公民信息的新型网络犯罪可能更加多发
法律/伦理学者	利于发扬他律与自律共治的道德作用；维护正当的合法性与必要性原则	以不信任作为获得信任的前提；以限制自由来保护自由；以正价值信息全面否定附加值信息；以用户个体的潜在风险换取网络空间的安全



# 大数据时代不再有个人隐私？



- 2013年6月6日，前美国中情局职员斯诺登披露PRISM计划，引起公众极度关注
- 互联网技术创新 vs. 个人信息/隐私泄露——>社会安全风险
- 2016年2月17日，《库克怒发公开信：苹果不会给美国政府开后门》

# 隐私之变

## 大数据时代再无隐私——大数据怎么获取



学堂在线  
xuetangx.com

采集方法	案例	示例技术	用户能感知到吗？	用户可自由选择退出吗？
收集公开数据	用爬虫软件“扒”近期微博	新浪微博开放API	不能	不能
公开收集数据	微博关键词云图应用，网站问卷	Web应用，Cookie...	能，确定	不用
日志文件	电商，搜索引擎，地图.....	Cookie...	不能	不能或很难
隐藏式收集	手机手电筒App索取获取精确定位信息权限	Android /ios等API	能，常被忽视	不能或难
攻击、破解	12306用户信息暴露	黑客攻击等	不能	不能
买卖	骚扰信息（出生、银行开户、手机开户...）	交易（公开或私密）	不能	不能
关联、推断	洛杉矶警方统计推断出某些小区较不安全	关联分析、聚类分析、机器学习	可能不能	不能

# 隐私之辩

- 大数据预测，是采用聚类、关联分析、统计学习等方法，对多来源、多形式、多维度的海量数据进行计算，从而不仅能“复制过去”，很多成功的案例见于感冒疫情预测、经济走势反转点预测、CPI指数预测等
- 然而，其所处理、分析的数据可能主要是人的活动，如：当看到求职者社交网上留着其“青春期”酗酒后的狂乱神态的照片，就此预测当事人未来还将缺乏自我把控能力而拒绝聘用，合理吗——网络遗忘权受法律保护吗？
- 有学者认为大数据预测技术存在伦理困境
  - 结果预判挑战自由
  - 隐私披露挑战尊严
  - 信息垄断挑战公平
  - 固化标签挑战正义

# 数据权利：大数据是资产吗？

- 从财务上看，自然人或法人的资产必须具备三要素：
  - 被他拥有和控制
  - 能够用货币来衡量
  - 能为他带来经济利益
- 大数据实践中，
  - 大数据在数据权属上存在模糊地带
  - 其货币价值与可用真实性、可信性、完整性、可用性等指标度量的数据质量密切相关
  - 大数据的价值多体现在关联价值上，即通过将数据不断聚合、加工后产生增值
  - 不同于一般物质性资产，数据的价值不随便使用次数的增多而减少，具有非消耗性



# 哪些权利？如何扶“弱”？

- UGC（用户产生的数据）是大数据应用重要依靠
  - 但，这些数据的权属在认识上和实践上存在多种分歧：
  - 未清晰定界数据所有权和使用权，且缺乏明确的数据授权、让渡机制
  - 缺少对数据是否按照预设目的和要求来使用、共享和删除的审计权
  - 未定义涉及财产性和声誉性汇报的数据分红权

哪些权利？如何扶“弱”？		
弱者	风险与威胁	保护
“数据鸿沟” 彼岸	丧失公平机会 人财物安全受侵害	政府普惠信息设施、教育、基本社会保障 培育和发展社会组织 提供人性化公益服务
个人一方	个人隐私被侵害 数据收益无保障	加快个人隐私、数据资产相关法治建设，尤其是司法实践 研究、试行数据权属交易规则 培育和发展坚守社会价值的社会组织
小微创业者	数据资产被垄断	加大数据共享、开放力度 规范数据市场

# 推动科学研究数据开放共享

- 传染病病毒结构、疫苗研制案例，加快阻止传播、治愈病人：AIDS、SARS、EBOLA ...
- 抗震救灾的地震科学数据共享案例：《地震科学数据共享系统运行规范》（中国地震局2009年颁布）
- 缩小研究数据造假、学术失范的空间：小宝方晴子、黄禹锡、汉芯造假等
- “众筹”力量开展科学研究

# 数据：从共享到开放

“谁占有数据，谁就占有先机。”从增进社会福祉角度考虑，大数据时代更应积极推动数据共享与开放，让人们能够更加平等、公正地访问和使用数据，让科学发现和知识生产过程更加快速，让更多产品和服务得以萌生，让更多的就业机会和更适宜的生活方式不断涌现。



感谢观看

THANK YOU FOR WATCHING

