

信息技术和大数据创新具有**双刃性**，其负面效应往往是模糊、滞后却又层出不穷的。

技术本身可以是中立的，但做技术的人需在法律、伦理框架下做正确的而不仅仅是最经济的方案选择——技术和工程伦理在信息技术和大数据创新领域存在且十分重要！

信息与大数据背后的问题：获取是否有逐一明确、授权？分析结果是否合理可信？存储和使用是否安全？应用是否平等、惠民？收益能否让参与数据提供个体共享？

在信息技术和大数据创新设计、开发、应用阶段，怎么能有效规避这些风险、进而保障创新成果可持续呢？（反面案例：共享单车）

1.1 信息技术的发展

什么是信息技术？

不同角度的描述：

信息技术是指有关信息的收集、识别、提取、变换、存储、处理、检索、分析和利用等的技术。

信息技术是指利用计算机和现代通讯手段获取、传递、存储、处理、显示信息和分配信息的技术。

信息技术是指研究信息如何产生、获取、传输、变换、识别和应用的科学技术。

1.1 信息技术的发展

近当代重要发展历程

电话：1900年纽约电话公司卡尔迪将普宾发明的装配载荷线圈成功用于扩大电话传输范围且减少交叉通话干扰；1914年地下电缆连接波士顿、纽约与华盛顿，进而发展成电子交换、**环球通话**

广播电视：1901年马可尼用莫尔斯码向2000英里外发送无线**电报**；1912年阿姆斯特朗研制出第一台调幅无线**收音机**；1925年拜尔德成功地传了一个可识别图像，**广播、电视**随后成重要大众传媒

电子化：1904年弗莱明发明**真空管和二极管**；1947年巴顿、布拉泰因和肖克利发明**晶体管**；1958年诺伊斯开发出能够可靠制造的小型**集成电路**，后创办英特尔公司；此后，IC芯片集成度依照摩尔定律每18个月翻番

1.2 当代信息技术的特点

颠覆性是信息技术的独特性质

连接能力：在无线、有线、局域网、广域的通信网络技术和手机、智能终端、计算机、嵌入式设备支持下，人、机、物形成**全时空、可追溯、可预测**的互联互通的网络

交互能力：符号、命令、文字、语音、图像乃至手势、表情，都可以被计算设备感知、识别，人机之间可以更加**自然“对话”**

渗透特性：家电可以上网，汽车可以联网，农作物生长态势及销售情况可以由农业物联网送达农技人员、采购人员和百姓、政府……各种**嵌入式设备**被戴在手上、穿进鞋里、藏在筷子里。跨界、颠覆，成为信息科技重要特性

融合能力：信息科技以数字化的0和1为基本形式记录、存储、传输、转换各类信息，不同信息可以方便地传输到同一个设备上，进而进行匹配、关联、融合等深度处理，产生**新的使用价值**

1.3 信息技术与社会变革

工业4.0来源及概念

“工业4.0”是德国联邦教研部与联邦经济技术部在2013年汉诺威工业博览会上提出的概念。德国学术界和产业界认为，“工业4.0”概念即是以智能制造为主导的第四次工业革命，或革命性的生产方法。该战略旨在通过充分利用信息通讯技术和网络空间虚拟系统—信息物理系统（Cyber-Physical System）相结合的手段，将制造业向智能化转型。

工业4.0描绘了制造业的未来愿景，提出蒸汽机的应用、规模化生产和信息技术等三次工业革命后，人类将迎来以信息物理系统(CPS)为基础，打通所有生产环节的数据壁垒，无线网掌控一切，旨在提升制造业的智能化水平，建立具有适应性、资源效率及人因工程学的智慧工厂，在商业流程及价值流程中整合客户及商业伙伴。其技术基础是网络实体系统及物联网。

工业4.0提倡以生产高度数字化、网络化、机器自组织为标志的第四次工业革命。



1.3 信息技术与社会变革

信息与大数据伦理问题

1. 信息技术的社会影响

1.1 信息技术的发展

1.2 信息技术的特点

1.3 信息技术与社会变革

1.1 信息技术的发展

重要发展历程——5次革命



1.1 信息技术的发展

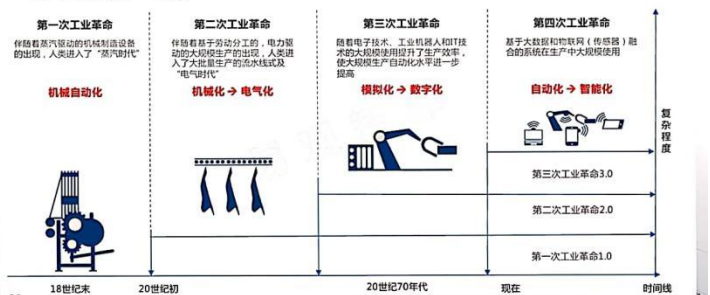
近当代重要发展历程

计算机：1939年阿塔纳索夫和拜利在衣阿华大学发明第一台**计算机**；1945年**图灵**发表论文阐述现代计算机原理，**冯·诺依曼**独立地写出描述存储程序的计算机文件，两人共同奠定了计算机工业基础；1946年第一台**电子数字计算机ENIAC**在宾夕法尼亚大学诞生；1975年盖茨和艾伦创建**微软公司**；此后30多年里，Wintel架构统领个人计算机时代，直到2007年乔布斯推出苹果手机iPhone，开创**智能手机和移动互联网**应用新时期

互联网：1969年美国军方研制的ARPANET公诸与众；1972年汤姆林森推出电子邮件；1991年荷兰的迪姆·伯奈斯-李发明万维网(World Wide Web)，从此，信息互通，地球成村，应用创新，势头正盛——雅虎（门户）、谷歌（搜索）、易贝（电商）、推特（社交）、脸书（社交）、优步（交通）……以及基于中国的阿里、百度、腾讯为代表的互联网企业

1.3 信息技术与社会变革

工业1.0-4.0



1.3 信息技术与社会变革

工业与互联网融合

Table with 4 columns: Industry/Region, Industry 4.0, U-Japan, China 2025. It lists various companies like IBM, Siemens, Denso, Haier, etc. and a table showing market value projections for different industries.

1.3 信息技术与社会变革

未来的制造将是基于大数据、互联网、人，通过各种信息技术进行柔性制造。

互联网+

以互联网为主的一整套信息技术在经济、社会生活各部门的扩散和应用过程,其本质是传统产业和生产过程的在线化、数据化。

互联网+作用下的产业,信息(数据)将作为独立的生产要素存在,并成为驱动产业发展的核心要素。

互联网+变革传统行业。

视频2: 什么是互联网+



影响两面性

信息技术是社会进步的加速器: 提供新的技术手段、经营业态、思想观念、社会网络, 支撑市场经济改革和向现代化的转型

信息技术创造社会生活新方式: 在线学习、电子商务、电子政务.....

信息技术引发社会新问题、给社会科学研究带来新机遇:

- 1、信息安全隐患经由CPS (信息物理系统)而扩大到物质社会系统的巨大风险 (风险社会)
2、数字鸿沟进一步拉大发展不平衡 (区域、代际、贫富) (公平和正义)
3、社交网络正在挑战社会结构、社会秩序、社会控制、社会道德伦理的理论构建和分析结果

历史回顾

1950年, 维纳出版《人有人的用处: 控制论与社会》, 率先提出信息与计算机伦理, 追问信息技术对诸如生命、健康、快乐、能力、知识、自由、安全、发展机会等人类核心价值的意义, 并采用“伟大的公正原则”作为信息伦理的基石

60年代, 计算机科学家唐·帕克收集计算机专业人员利用高科技犯罪和从事不道德行为的案例, 为美国计算机学会起草计算机工程师职业伦理规范并宣讲、培训

信息与大数据伦理问题

2. 信息与大数据伦理问题

- 2.1 信息与网络空间伦理
2.2 大数据时代伦理新问题
2.3 数字身份困境
2.4 大数据时代的个人隐私
2.5 数据权利
2.6 大数据公共治理伦理

历史回顾

70年代, 既是哲学家后又成为计算机教授的瓦尔特·曼纳提出 Computer Ethics 作为新的应用伦理方向

80年代后期和整个90年代, 计算机伦理学发展迅速。1985年, 詹姆斯·穆尔发表论文“何谓计算机伦理学?”, 德博拉·约翰森撰写《计算机伦理学》经典教材, 学术会议、大学课程、研究中心、专业期刊和讲席教授席位应运而生; 实际问题层出不穷; 还出现了负责日常甄别信息技术使用情况、监管滥用的专门组织, 定期发布报告, 提出降低风险的举措

相关伦理主题

人际关系虚拟化: 线上真假互存社交活动增加, 线下面对面真实社交时间减少, 人对生活满意度的内涵有什么变化? 符合现有社会伦理吗?

正当的网络行为: 实名与匿名看法尚未统一, 政府的技术鉴别或管制行为对否获得授权?

平等与公正: 在公正原则被置于工程伦理重要地位的现代社会, 谁有责任向“信息贫困”人员提供技术、服务和平等生存的伦理责任?

知识产权争议: 知识产权规则应当是什么?

全球化信息交互与治理困境: 乐观主义者 VS 悲观主义者

“我”和大数据

我享受了网络带来的生活服务、社会交往和信息获取的便利, 也在网络上留下了很多“足迹”

我也在生活中经受了不推销电话、不感兴趣的网络推荐等烦扰, 还不时担心个人账号失密

我有条件地同意研究人员使用我个人数据展开科技创造, 比如, 我不希望做对我个人行为 and 日常生活的挖掘、跟踪!

☆ 与您浏览过的商品相关的推荐



什么是大数据?

大数据 (big data), 指无法在一定时间范围内用常规软件工具进行捕捉、管理和处理的数据集合, 是需要新处理模式才能具有更强的决策力、洞察发现力和流程优化能力的海量、高增长率和多样化的信息资产

大数据的5V特点: Volume (大量)、Variety (多样)、Velocity (高速)、Value (价值)、Veracity (真实性)

视频3: 什么是大数据?



大数据与以往数据应用的不同

第一, 可以获得全体数据而非采样数据, 这既决定了大数据算法原理与样本分析方法明显不同, 也体现其复杂度迥异

第二, 允许获得的数据呈现混乱、复杂状态而不再强求干净、精确, 即大方向的正确比微观精确更重要

第三, 聚焦发现和分析事物的相关性而非因果性, 避免在因果性上劳而无获、止步不前。

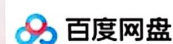
数据安全

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”安全管理和审计体制安全风险

百度网盘, 让美好永远陪伴



百度网盘为您提供文件的度快、安全稳固, 支持教习受15G的免费存储空间 https://pan.baidu.com/



给你一个好地方, 来保存照片、文件等资料。

iCloud 内置在每一部 Apple 设备中, 因此, 你的照片、文件、备忘录等内容都能得到安全的保护, 实时的更新, 并随处供你取用, 它能自动备份, 让你专心去做自己的事, 一开如使用 iCloud 时, 每个人都会获得 5GB 的免费存储空间, 并且随时都能轻松升级。

• 大数据伦理

20世纪90年代，人们认为“在互联网上，没有人知道你是一条狗”

不久的将来，你很可能是透明的！

大数据时代，一些被广泛珍视的伦理价值，如权利平等、交易公平、安全感、诚信、自由、公正等正在经受挑战

视频4：大数据、透明人



33



• 大数据伦理

大数据时代，作为技术应用提供方的数据工程师、大数据创新企业、政府部门，与作为使用方的普通用户、社会团体，面对新的伦理挑战：

- 1、**身份困境**：数字身份与社会身份，可以分离还是必须关联？
- 2、**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？
- 3、**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？
- 4、**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？

34



• 数字身份

一组独一无二地描述一个主体或实体的数据，是关于一个人的所有在数字上可得的信息的总和。

数字身份特点：多样性、可变性、允许匿名和假名。

数字身份管理技术：

- 1、用户名+密码+验证码
- 2、第三方认证，如U盾
- 3、预存的个性化问题
- 4、生物特性，如指纹、虹膜等



36



• 大数据引发数字身份新问题

易被盗用、易被追溯



37



• 大数据引发数字身份新问题

人肉搜索——反腐：

2012年8月26日，陕西省安监局局长杨达才在延安交通事故现场，因面含微笑被人拍照上网，引发争议并被网友指出杨达才有多个名表。

2012年09月21日，陕西省研究决定：撤销杨达才陕西省第十二届纪委委员、省安监局党组书记、局长职务；

2013年8月30日，西安中级人民法院在3号法庭公开开庭审理被告人杨达才受贿、巨额财产来源不明一案。同年9月5日获刑14年。



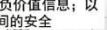
38



• 数字身份管理的伦理分析：匿名vs实名

		辩论实名制	
		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	盗窃、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；以正值信息全面否定负价值信息；以用户个体的潜在风险换取网络空间的安全	

39



白岩松：警觉沉迷于“投你所好式”网络 这是民族危险

2019年3月6日上午，白岩松与丁磊讨论，谈到了对于“投你所好式”互联网推送的担忧，“让你原地踏步，让你在自己的兴趣里头沉迷而不会提升，这是一个民族的危险。”



丁磊：每个人的注意力有限，人们肯定喜欢看好看的电影电视剧，看公众号，看资讯，现在都是碎片式阅读。

白岩松：首先要建立阅读习惯，这一点家长和社会都有责任，其次是阅读要方便、便捷，再者阅读的培养在校园里头应该是持续的。对大学生没有培养好阅读习惯，因时间碎片而不读书担忧。

40



• 保护数据隐私面临的技术和非技术挑战

可信性与可靠性：在大规模、分布、开放的信息基础设施内，存在为数众多的数据收集、处理和发布实体，难以确保各自具有可靠的、可信的数据管控能力

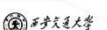
快速扩散性和放大器效应：数据除了存储在专有数据库系统外，还存在被出售、被快速扩散、快速覆盖的可能性。隐私痕迹很难消除。

挖掘技术与关联发现：可以把零散的、碎片化的数据重新关联、拼接起来，从而复原一个人的整体轮廓。

身份盗用与冒用：恶意使用偷盗来的数字身份，例如进行信用卡诈骗，甚至用于邪恶的网站败坏被盗人身份。

恶意攻击：现行数据管理系统无力防备黑客的犯罪行为或信息战侵略。

43



• 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“复现过去”，也能“预测未来”，很多成功的案例见于感冒疫情预测、经济走势转折点预测、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？

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

45



• 隐私权、个人信息权的法律实践

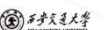
隐私权：自然人享有的私人生活安宁与私人信息依法受保护，不被非法侵扰、知悉、搜集、利用和公开的一种人格权

个人信息：与特定个人相关联的、反映个体特征的具有可识别性的符号系统，包括个人身份、工作、家庭、财产、健康等各方面信息

个人信息权：对自己个人信息享有的支配并排除他人非法利用的权利

2021年8月20日，十三届全国人大常委会第三十次会议表决通过《中华人民共和国个人信息保护法》，自2021年11月1日起施行。

46



• 数据的资产属性与数据权属——大数据是资产吗？

从财务上看，自然人或法人的资产须具备三要素：
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：

- 1、大数据在数据权属上存在模糊地带
- 2、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
- 3、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
- 4、不同于一般物质性资产，数据的价值不随使用次数的增多而减少，具有非消耗性

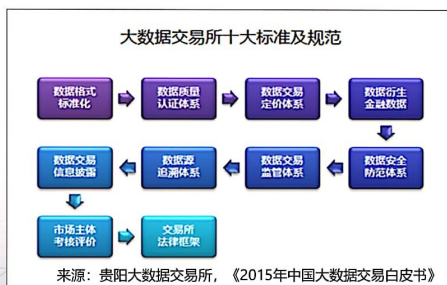
• 防范数据滥用与垄断

视频8：数据滥用



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 加快个人隐私、数据资产相关法治建设，尤其是司法实践
个人一方	个人隐私被侵害 数据收益无保障	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 加大数据共享、开放力度
小微创业者	数据资产被垄断	规范数据市场

• 探索建立大数据交易规范和公平市场



• 推动数据公开、共享与开放

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...

抗震减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



• 推动数据公开、共享与开放

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、公平地访问和使用数据，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。

类型	访问模式	数据要求	权利约定方式
数据共享	一对一、一对多和多对多 一般不公开访问	单独约定	逐一授权
数据公开	一对多，公开访问	包含元数据	公开声明文件约定
数据开放	一对多，公开访问	开放授权；数据结构化；开放格式；提供URI定位；能与其他数据链接	公开声明文件约定

• 中西公共治理价值观的差异

国家制度是当今人类社会的一种基本形态，现代国家保持政治权利、领土和人民三要素统一，强调国家主权及其合法性

现代国家基本职能：保护本国免受其他国家的侵犯，保护国内每个人免受他人的压迫和侵犯，承担个人或少数人不能完成的事情

不同国家自然资源、历史发展和形成的主流文化思想存在显著差异，使得在社会价值观和受此影响的公共治理伦理上各有侧重

中华价值观表现出四大特色：

责任先于自由、义务先于权利、群体高于个人、和谐高于冲突

• 大数据公共治理的伦理选择

中国政府数据资源共享开放工程

中国政府治理大数据工程促进计划——《促进大数据发展行动纲要》2015.8

坚持尊重、公正、有限记忆、无害和可持续原则，其中，信息惠民可能是大数据公共治理中的最高要求

信息惠民，惠什么？

方便，提供服务方便，老百姓感觉被服务的方便

安全，集中的、强大的风险防范强于分散的、薄弱的甚至缺失的风险防范

和谐美好，让社会运行更加和谐，让人民生活更加美好