

单位代码: 10359  
学 号: 2015111307

密 级: 公开  
分 类 号: B82



合肥工业大学  
Hefei University of Technology

# 硕士学位论文

## MASTER'S DISSERTATION

### (学术硕士)

论文题目: 大数据背景下数字身份的伦理问题及其克服研究

学科专业: 科学技术哲学

作者姓名: 程 昊

导师姓名: 董军 副教授

完成时间: 2018年3月5日



Y3473571

单位代码: 10359

密 级: 公开

学 号: 2015111307

分类号: B82

合肥工业大学

Hefei University of Technology

硕士学位论文

MASTER'S DISSERTATION

(学术硕士)

论文题目: 大数据背景下数字身份的伦理问题  
及其克服研究

学科专业: 科学技术哲学

作者姓名: 程昊

导师姓名: 董军 副教授

完成时间: 2018 年 3 月 5 日

合 肥 工 业 大 学

学历硕士学位论文

大数据背景下数字身份的  
伦理问题及其克服研究

作者姓名：程昊

指导教师：董军 副教授

学科专业：科学技术哲学

研究方向：生态哲学理论与方法

2018 年 3 月

A Dissertation Submitted for the Degree of Master

**Research on ethical problems of digital identity and  
their overcoming in big data era**

By

Cheng hao

Hefei University of Technology

Hefei, Anhui, P.R.China

March, 2018

# 合 肥 工 业 大 学

本论文经答辩委员会全体委员审查, 确认符合合肥工业大学硕士学位论文质量要求。

答辩委员会签名: (工作单位、职称)

主席: 徐宏 安徽大学 教授

委员:

茹洁玲 合肥工业大学 副教授

李才华 合肥工业大学 副教授

导师:

李军

## 学位论文独创性声明

本人郑重声明：所呈交的学位论文是本人在导师指导下进行独立研究工作所取得的成果。据我所知，除了文中特别加以标注和致谢的内容外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得合肥工业大学或其他教育机构的学位或证书而使用过的材料。对本文成果做出贡献的个人和集体，本人已在论文中作了明确的说明，并表示谢意。

学位论文中表达的观点纯属作者本人观点，与合肥工业大学无关。

学位论文作者签名：程昊 签名日期：2018年4月20日

## 学位论文版权使用授权书

本学位论文作者完全了解合肥工业大学有关保留、使用学位论文的规定，即：除保密期内的涉密学位论文外，学校有权保存并向国家有关部门或机构送交论文的复印件和电子光盘，允许论文被查阅或借阅。本人授权合肥工业大学可以将本学位论文的全部或部分内容编入有关数据库，允许采用影印、缩印或扫描等复制手段保存、汇编学位论文。

（保密的学位论文在解密后适用本授权书）

学位论文作者签名：程昊 指导教师签名：李军  
签名日期：2018年4月20日 签名日期：2018年4月20日

论文作者毕业去向

工作单位：

联系电话：

通讯地址：

E-mail：

邮政编码：

V

V

## 致 谢

首先感谢合肥工业大学，给了我学习的机会。非常感谢董军老师在本次论文写作的过程中，从选题指导、论文框架到细节修改，都给予了细致的指导，提出了很多宝贵的意见与建议，董军老师以其严谨求实的治学态度、高度的敬业精神、兢兢业业、孜孜以求的工作作风和大胆创新的进取精神对我产生重要影响。他渊博的知识、开阔的视野和敏锐的思维给了我深深的启迪。本篇论文是在老师的精心指导和大力支持下才完成的。

感谢所有授我以业的老师，没有这些年知识的积淀，我没有这么大的动力和信心完成这篇论文。感恩之余，诚恳地请各位老师对我的论文多加批评指正，使我及时完善论文的不足之处。

最后，我要向百忙之中抽时间对本文进行审阅的各位老师表示衷心的感谢。

作者：程昊

2018年3月XX日

## 摘 要

互联网技术、物联网智能技术和大数据技术呈几何指数增长，让全球发生了翻天覆地的变化，也让人类已经正式迈进了大数据时代。数字身份作为人在数字网络中交流和沟通的起点，也在大数据的影响下爆发了一系列的伦理问题，如数据权力被侵害，隐私安全、公平正义和身份认同问题等。这些伦理问题不仅给大数据技术的发展带了负面影响，也给人类社会造成了全球性的困扰。因此如何约束并规范不同数字身份的人们在大数据背景下的网络行为就是解决解决数字身份伦理问题的重中之重。

本文首先介绍了国内外相关问题的研究现状，而后对大数据和数字身份及其相关伦理问题做了概念性的阐述。大数据技术带来的海啸式影响力，一方面加剧了数字身份的伦理问题，另一方面也让数字身份产生了一些新的伦理问题，这些伦理问题不但复杂多样，而且盘根错节、相互交织，一旦疏于治理，必将引起严重的灾难，因此构建大数据时代中数字身份伦理问题的伦理原则和规范是十分必要的。本文在对身份伦理问题的对策构建中，指出数字身份伦理问题基本表现为数字权利被侵害、隐私安全、公平正义和身份认同四大类问题，并分别对其进行了分门别类的阐述，大数据使用中的各种“鸿沟”让不同主体之间的公平难以得到保障，大数据的“数据万能”倾向让主体对自身数字身份的认同产生了单一化、同质化的趋势等。然后进一步从主体根源、客观原因、社会背景三个方面来分析产生上述数字身份伦理问题的根源，主体根源是主体的复杂性、客观原因是技术的负面效应、社会背景是外部规约的不足，并将其进一步细化为 12 个方面进行探讨。最后本文提出了构建大数据背景下数字身份伦理问题的对策和实践路径。一是要坚持 5 大伦理原则，包括：坚持以人为本原则，以无害为前提，向善为目的；坚持权责统一原则，确保用户的知情同意和隐私安全；坚持公平正义原则，平衡大数据利益相关者的利益；坚持开放共享原则，实现过程的透明与公开，同舟共济；坚持道德与法律相结合的原则，完善法律制度，形成道德自律。二是 5 条实践路径，包括：制定和完善法律法规，加强政府监管力度；提倡行业自律机制，建立社会监督模式；调整隐私观念，强化自我保护意识；努力进行技术创新，降低隐私泄露风险；全民提高社会公民素养，消除数据信息不公平现象。

**关键词：**大数据；数字身份；伦理问题



## Abstract

Internet technology, Internet of things intelligent technology and big data technology showed exponential growth, so that the world has undergone tremendous changes, but also let mankind have officially entered the era of big data. As a digital identity in the digital network communication and the starting point, also broke out a series of ethical problems in the impact of big data, such as data privacy abuses, security, justice and identity issues etc.. These ethical problems not only have negative effects on the development of big data technology, but also cause global trouble to human society. Therefore, how to restrict and regulate the network behavior of people with different digital identities in the background of big data is the most important problem to solve the ethical problems of digital identity.

This paper first introduces the research status of related issues at home and abroad, and then gives a conceptual description of big data and digital identity and related ethical issues. A tsunami influence brought about big data technology, on the one hand, exacerbating the ethical problems of digital identity, on the other hand also let digital identity have some new ethical problems, these ethical problems not only complicated and intertwined, complicated and difficult to deal with, once poor governance, will cause serious disaster, so the construction of digital ethical principles and norms the era of big data in the identity of ethical issues is very necessary. Based on the construction of Countermeasures of identity ethical issues, ethical issues pointed out that digital identity is basically digital rights infringement, privacy and security, justice and the identity of four kinds of problems, and has separately carried on the elaboration be arranged a variety of big data, "in the gap" between different subjects so that it is difficult to get the fair security, big data "universal data" tendency makes the subject on their own digital identity produced a single, homogeneous trend. Then, from the three aspects of the main causes and objective causes and social background to analyze the origin of the digital identity problems, the main root is the complexity

of the subject, the objective reason is the negative effect of technology, the social background is lack of external specification, and the further refinement of 12 aspects discussed. Finally, this paper puts forward the countermeasures and practice path of constructing digital identity ethics problems under the background of big data. One is to adhere to the 5 ethical principles, including: adhere to the people-oriented principle, as harmless as the premise, good for the purpose; adhere to the principle of unified responsibilities, ensure that the user informed consent and privacy; adhere to the principle of fairness and justice, balance the interests of big data; adhere to the open sharing principle, the realization process of transparency adhere to the moral and legal, pull together in times of trouble; the principle of combining, perfecting the legal system, the formation of moral self-discipline. Two is the 5 path of practice, including: the establishment and perfection of laws and regulations, strengthen government supervision; promote self-discipline mechanism, establish social supervision mode; adjust the concept of privacy, strengthen the awareness of self-protection; efforts to carry out technological innovation, reduce the risk of loss of privacy; the improvement of social citizenship, eliminate data unfair phenomenon.

**KEYWORDS:**big data; digital identity; ethical issues

# 目 录

摘 要.....	1
Abstract.....	2
第一章. 绪论.....	7
1.1 研究背景.....	7
1.2 研究的目的和意义.....	8
1.2.1 研究目的.....	8
1.2.2 选题意义.....	8
1.3 国内外研究现状.....	8
1.3.1 国内研究现状.....	8
1.3.2 国外研究现状.....	10
1.5 可能的创新点.....	11
第二章 大数据及数字身份概述.....	12
2.1 概念.....	12
2.1.1 大数据的概念.....	12
2.1.2 数字身份的概念.....	13
2.2 关于数字身份特性的探讨.....	14
2.2.1 多元性.....	14
2.2.2 可伪性.....	15
2.2.3 流动性.....	15
2.2.4 超域性.....	15
2.3 大数据对数字身份的影响.....	16
第三章. 大数据背景下数字身份的数据权利被侵害的问题.....	17
3.1 数据人格权受到侵害.....	18
3.1.1 知情同意权受到侵害.....	19
3.1.2 修改权受到侵害.....	20
3.1.3 封锁权受到侵害.....	21
3.1.4 遗忘权受到侵害.....	21
3.2 数据财产权受到侵害.....	22
3.2.1 采集权受到侵害.....	22
3.2.2 查询权受到侵害.....	23

- 3.2.3 可携权受到侵害..... 23
- 第四章. 大数据背景下数字身份的隐私安全问题..... 25
  - 4.1 隐私暴露问题..... 25
    - 4.1.1 身份隐私暴露问题..... 26
    - 4.1.2 行为隐私暴露问题..... 26
    - 4.1.3 喜好隐私暴露问题..... 27
  - 4.2 信息安全隐患问题..... 28
    - 4.2.1 数据信息可信度降低..... 28
    - 4.2.2 身份信息数据被泄露..... 29
    - 4.2.3 数字身份被盗用..... 30
- 第五章 大数据背景下数字身份的公平问题..... 31
  - 5.1 技术不相等..... 31
  - 5.2 知识不对等..... 32
  - 5.3 贫富不均等..... 34
  - 5.4 社会不平等..... 35
- 第六章 大数据背景下数字身份的认同问题..... 35
  - 6.1 数字身份认同的概念..... 36
  - 6.2 数字身份与现实身份相脱离..... 37
  - 6.3 数字身份认同的价值观与社会认同的主流价值观相背离..... 38
  - 6.4 数字身份在网络中的地位与现实身份在社会中地位有差距..... 40
- 第七章. 关于大数据背景下数字身份伦理问题的原因的探讨..... 41
  - 7.1 主体根源：主体的复杂性..... 41
    - 7.1.1 网络主体被数字化遮盖了主体的品性..... 41
    - 7.1.2 企业利用大数据逐利无视他人利益..... 42
    - 7.1.3 主体隐私观的缺失..... 43
    - 7.1.4 主体自身的道德素的缺失..... 44
  - 7.2 客体原因：技术的负面效应..... 44
    - 7.2.1 大数据技术尚不成熟..... 45
    - 7.2.2 互联网的固有特性..... 45
    - 7.2.3 大数据技术自身的独特性..... 46
    - 7.2.4 数字身份的多元性..... 46
  - 7.3 社会背景：外部规约的不足..... 47
    - 7.3.1 法律机制不健全..... 47
    - 7.3.2 行业内缺乏统一的伦理规范..... 48

- 7.3.3 监督机制不健全..... 48
- 7.3.4 社会功利主义风气盛行..... 49
- 第八章. 大数据背景下数字身份伦理问题的对策性研究..... 50
  - 8.1 大数据背景下数字身份伦理问题的伦理治理..... 50
    - 8.1.1 人类应遵循的伦理原则..... 50
    - 8.1.2 服务商应遵循的伦理原则..... 51
    - 8.1.3 市场行业应遵循的伦理原则..... 51
    - 8.1.4 政府应遵循的伦理原则..... 52
    - 8.1.5 社会应遵循的伦理原则..... 53
  - 8.2 大数据背景下数字身份伦理问题的具体措施..... 53
    - 8.2.1 制定和完善法律法规，加强政府监管力度..... 54
    - 8.2.2 提倡行业自律机制，建立社会监督模式..... 54
    - 8.2.3 调整隐私观念，强化自我保护意识..... 55
    - 8.2.4 努力进行技术创新，降低隐私泄露风险..... 55
    - 8.2.5 全民提高社会公民素养，消除数据信息不公平现象..... 56

# 第一章 绪论

## 1.1 研究背景

随着运算技术，云技术等高新科技的不断发展，大数据时代已经到来。大数据已经影响着人们生活的方方面面，谷歌公司开发出 GFT（Google Flu Trends）基于对关键词的搜索技术来预测流感疫情，美国麻省理工学院通过结合交通数据和用户手机的定位数据来规划城市的建设方案等等，大数据应用早已经渗透进了我们的生活之中。放眼未来，由于大型数据源的数量增加和大数据分析的技术能力提高，对大数据的研究的兴趣，预计将不断增加。

大数据彻底改变了人们的思维方式，把人们的注意力从因果关系转移到了相关关系，从“WHY”转为了“WHAT”，人们只需要知道是什么，而无需知道为什么。以翻译工具为例，IBM 公司在业界拥有强大配置的硬件条件、优秀的优化软件和超多的人工时间，因此 IBM 公司也成为了世界上在线翻译的巨头公司。为了更好地实现中英文互译，IBM 曾经将所有的《人民日报》中的文字输入超级电脑，希望能够破解汉语结构的秘密，但很可惜效果并不理想。谷歌公司制定了与 IBM 公司完全不一样的战略计划，谷歌采用各种在线网站，各民族的官方语言文本以及海量的云端图书等信息，来不断优化改善自身软件——谷歌翻译的翻译性能，并且谷歌翻译还能不断通过云端进行智能优化，终于后来居上超越了 IBM，谷歌凭借并不够精确的大数据战胜了 IBM 足够精准的小数据。

大数据虽然具有近乎无限的潜力与价值，但也存在着相当的安全隐患和难以解决的伦理问题。2012 年，著名社交网站 Facebook 就与加利福尼亚大学和康奈尔大学强强联合，在用户不知情的情况下，针对七十万名用户进行了一项名为“情绪感染”（emotional contagion）的测试，旨在确定来源于他人的情绪是否会影响用户自身的情绪，并不断在 Facebook 上形成大规模的“传染”现象。据英国《卫报》评述，此项实验一经披露就引发了大众强烈的抵触情绪，很多用户言辞激烈的对 Facebook 进行了批评，表示这项实验让人感到十分后怕，英国通讯信息管理局也表示可能会介入调查，Facebook 此举可能违反了数据保护法案。

身份是界定一个人是谁或是什么的一个特征或属性集。身份有社会身份（同伴、家庭和朋友）、法律身份（出生证、驾驶执照）以及物理身份（DNA、外观等）。数字身份则是大数据时代中人类与数据的桥梁，是人类在网络中的化身，它起源于人们对在线交易的需求，其来源可以追溯到因特网开放之初，在当时，人们便拥有了相应的数字身份，并且根据身份的不同也具有不同的权限，比如网站的管理员和访问者等等。当然数字身份的定义也远不止涉及管理者和访问者这么简单。

在大数据时代的背景下，数字身份的伦理问题也不断凸显。大数据市场目前

的混乱和无序让人们的数据权力极易遭到侵害，大数据中的相关性技术让人们面临着信息被泄露、身份被盗用的风险，大数据对未来的预测性让人们的隐私安全变得形同虚设，大数据使用中的各种“鸿沟”让不同主体之间的公平难以得到保障，大数据的“数据万能”倾向让主体对自身数字身份的认同产生了单一化、同质化的趋势。因此，如何解决这些伦理问题，以期能够最大化发挥的大数据潜能，是当今社会难以回避的研究课题。

## 1.2 研究的目的是和意义

### 1.2.1 研究目的

通过对数字身份在大数据背景下的伦理问题的探索性研究，认识和把握当代社会数字身份所带来的一系列伦理问题，并提出应对策略，以期完善关于数字身份的公共政策建构，促进社会的和谐稳定。

### 1.2.2 选题意义

#### ①理论意义

一方面梳理数字身份在当前大数据背景下的概念和范畴，明确数字身份在大数据伦理问题中的重要性、特殊性。

另一方面通过对数字身份伦理问题的研究，积极回应大数据技术进步对当前社会发展提出的挑战，丰富科技伦理、科技哲学的研究内容。

#### ②实践意义

一方面从现实层面分析数字身份的客观存在，探索大数据背景下数字身份所带来的社会伦理困境。

另一方面从社会层面把握数字身份的伦理问题，提出化解数字身份伦理问题的对策和建议，促进社会和谐。

## 1.3 国内外研究现状

### 1.3.1 国内研究现状

国内数字身份伦理研究一般分为数字身份被盗用、数字身份信息的隐私安全和数字身份需要被认同三个方面的问题。

#### ①数字身份被盗用的问题

杨山从科技进步的角度分析了网络身份盗用的概念，认为网络身份盗用是身份盗用的网络形式，“是利用现代信息技术及网络资源进行的身份盗用行为”<sup>1</sup>。杨诚从盗用身份流程和信息载体的角度将其表述为身份犯罪，“指非法获取、持有、收

<sup>1</sup> 杨山. 网络身份盗用类型及防范机制研究[D]. 西南交通大学, 2008.

集、传播、买卖、使用和伪造身份信息、文件和标志的行为”<sup>1</sup>。林荫茂从条件、内涵和目的三个角度总结了身份盗用的注意事项，身份盗用的成立条件是“以身份欺诈为目的”，包含“在世人的身份信息还包括已故人的身份信息”，目的是为了“非法获得利益和非法逃避义务”<sup>2</sup>。

## ②数字身份信息的隐私安全问题

薛孚从数据挖掘的方面讨论了身份信息泄露，最终侵犯隐私安全的问题。数据信息的不断被二次甚至多次利用，大数据的相关性让这些看似无关隐私的信息被整合和关联起来，暴露出个人身份信息，从而威胁他人的隐私安全，“数据挖掘能够将原本看似不涉及个人隐私的数据关联起来，暴露个人的身份信息”<sup>3</sup>。安宝洋从影响后果方面分析了泄露隐私的问题，他认为不管数据挖掘技术是有意还是无意泄露了他人隐私，都会打击用户的热情、降低用户的信任感，最终会“会迫使大量使用者因惧怕安静生活被干扰而最终叛离数据服务”<sup>4</sup>。黄欣荣将当今大数据世代泄露隐私威胁安全的现象形容为“裸奔”，面对大数据，我们的隐私安全不复存在，隐私已经消失在了历史之中，我们每个人“都成了透明人”<sup>5</sup>。

## ③数字身份需要被认同的问题

张欧阳从差异性的角度出发，分析了网络身份与现实身份的不同性，认为人们在网络和现实的感觉并不同，因而网络身份也是自我认同的一部分，“用户在长期使用这个网络身份的过程中形成了一个独立的身份认同与人格特质……网络和现实的双重身份同时存在，变成自我认同的部分”<sup>6</sup>。文贤庆从个人真实存在的角度出发，指出无论如何网络终究不可能脱离现实社会而存在，所以网络身份的认同最终还是要回归到社会交往关系之中，“一个活着的有机生命体，它能够进行感觉、认知、思想，形成自我反思和社会交往，这才可以称之为一个自我身份认同的主体和实体”<sup>7</sup>。宋辰婷从理论变迁的角度分析了社会对数字身份的认同问题，并认为社会认同已经开始从身份和归属认同转变为评价性认同“认同感不仅仅停留在对周围社会的信任和归属、对有关权威和权力的遵从，而且增加了更多的评价性因素”<sup>8</sup>。

<sup>1</sup> 杨诚. 身份犯罪立法的国际动态及其启示[J]. 政治与法律, 2008(08):35-39.

<sup>2</sup> 林荫茂. 从信用卡犯罪看身份信息犯罪[J]. 政治与法律, 2008(09):83-89.

<sup>3</sup> 薛孚, 陈红兵. 大数据隐私伦理问题探究[J]. 自然辩证法研究, 2015, 31(02): 44-48.

<sup>4</sup> 安宝洋. 大数据时代的网络信息伦理治理研究[J]. 科学学研究, 2015, 33(05):641-646.

<sup>5</sup> 黄欣荣. 大数据技术的伦理反思[J]. 新疆师范大学学报(哲学社会科学版), 2015, 36(03):46-53+2.

<sup>6</sup> 张欧阳. 网络民主的要素分析[J]. 兰州学刊, 2012(12):176-180.

<sup>7</sup> 文贤庆. 信息理论视域下身份认同的再审视[J]. 云南社会科学, 2017(03):41-47.

<sup>8</sup> 宋辰婷. 网络时代的感性意识形态传播和社会认同建构[J]. 安徽大学学报(哲学社会科学版), 2015, 39(01):149-156.



### 1.3.2 国外研究现状

#### ①数字身份被盗用的问题

Clare Sullivan 在 2011 年提出了将数字身份 (“Digital Identity”) 作为一个新兴的法律概念。随后的美国身份证法案证实了 Sullivan 的论点, 并展现了涉及数据库身份和交易身份的新的法律概念。Clare Sullivan 认为数字身份滥用和欺诈的后果是非常严重的, 因为它直接影响的是该身份所对应的能够负法律责任的该人。世界隐私论坛 World Privacy Forum 认为身份盗窃可用于便利或资助其他犯罪, 包括非法移民、恐怖主义、网络钓鱼和间谍活动等。美国联邦贸易委员会 2006 身份盗窃调查报告说明, 确定数据泄露和窃取身份之间的联系是非常具有挑战性的, 主要是因为身份盗窃的受害者往往并不知道他们的个人信息, 与此同时定性 “身份盗窃” 这一问题是由美国联邦贸易委员会根据受害者个人提交的报告所检测认定的。

#### ②数字身份信息的隐私安全问题

美国印第安纳大学情报学的副教授 L. Jean 在 2004 年就提出了有关自动人脸识别、标注、跟踪和许多个人行为相关的广泛的数字认证系统, 会让其与个人身份相关联变得很容易, 因此, 有时候隐私将会失去, 而安全性将被颠覆。一个建立在确认假名基础上的身份系统可以为数字业务和交易提供隐私和安全。而网络空间创造身份盗窃的机会, 在网络上在数字通信信道上发送的一切的都是可以被记录和保存的。因此, 网络空间需要一个系统, 允许个人验证自己的身份, 而不勇向他人透露自己的数字身份<sup>1</sup>。约翰·帕克 (2014) 将大数据时代我们面临的安全与隐私困境总结为 “全民监控” 四个字, 将如今无所不在的监控措施形容为天罗地网, 让任何人的生活 and 行为都无所遁形, 此时人们的私生活似乎已经名存实亡<sup>2</sup>。

#### ③数字身份需要被认同的问题

McRae 指出网络的虚拟性使得不仅让心灵、身体和自我之间的分隔加剧, 也让其变得复杂化了。这种分隔投射到虚拟世界时, 心灵, 身体和自我意识都通过人与人之间的相互作用成为一种主动的构建<sup>3</sup>。López, Néstor<sup>4</sup>认为青年人在互联网上谈论他们的生活和烦恼, 并将之设计成最优化和电子化的被社会认可的形式提供他人, 并评估其他人对它的反应。在他们的每个帖子, 上传的图像或视频中,

---

<sup>1</sup>L. Jean. "Digital Identity". IEEE Technology and Society Magazine. IEEE.

<sup>2</sup>约翰·帕克. 《全民监控》[M]. 金城出版社

<sup>3</sup>McRae, Shannon (1997). "Flesh Made Word: Sex, Text, and the Virtual Body". In Porter, David. Internet Culture. New York: Routledge. p. 75.

<sup>4</sup>López, Néstor, Oportti, Renato; Vargas Tamez, Carlos (2017). Youth and changing realities: Rethinking secondary education in Latin America (PDF). UNESCO. pp. 44 - 45. ISBN 978-92-31 00204-5

他们都有可能问自己是谁，来尝试与真实世界不同的印象。因此他们从他人所反应的接受和责难之中认同自己的身份，创造归属感，这是身份建构过程中的一个重要标志。

综合国内外学者的研究来看，研究者对于数字身份的伦理问题从概念、特征、表现、原因到解决之道都做出了十分细致和系统的研究。但是由于研究时间较短，再加上大数据技术创新较快，关于大数据背景下数字身份伦理问题的研究依然存在一些不足。

理论性较强，实践性不足。大数据背景下数字身份伦理问题的研究基本困于学术圈内，理论性很强，但关于大数据背景下数字身份伦理问题的讨论和研究，学术圈外部的氛围明显要低于学术圈内部。且这些研究成果较难应用于商家和企业之上，大部分商家和企业对于这些伦理问题依然仅限于喊口号、打标语，对这些理论成果并不能够践行。

研究大都集中于某一方面，比如数字身份盗用问题，数字身份的认同性问题等，将这些问题整合起来的研究较少，缺乏针对数字身份伦理问题的综合性系统性研究，同时针对大数据对数字身份伦理问题的影响研究也相对较少。

因此从整体来看，我国的大数据背景下数字身份的伦理问题研究正处于发展阶段，随着大数据技术的不断进步，数字身份引发的伦理冲突必将进一步加剧，如何解决这些伦理问题是当下紧要的问题，这正是本文研究的必要性。

#### 1.4 研究方法

本文的研究涉及哲学、经济学、伦理学、社会学、计算机学等诸多领域，主要采用文献研究法、专家访谈法和跨学科研究法三种。

文献研究法：本文通过查阅大量国内外相关资料，梳理了目前国内外对数字身份伦理问题的研究现状，弄清大数据背景下，数字身份伦理问题的表现形式，文献主要来源于知网期刊、网络文献和各种书籍等。

专家访谈法：参与国内外的学术会议，并对相关专家进行访谈，从伦理学、哲学、法学等各方面了解专家对数字身份伦理问题的看法和观点。

跨学科研究法：本文认为大数据背景下的数字身份伦理问题是跨越哲学、伦理学、社会学、法学、经济学和心理学等诸多学科的综合性问题，因此本文力求通过各种角度多元化的分析数字身份的伦理问题，以期能够从个方面对数字身份的伦理问题提出相应的解决对策。

#### 1.5 可能的创新点

一是根据目前对数字身份的定义较为混乱的情况，梳理数字身份的概念，从多角度定义“多元化”的数字身份的外延、内涵和特点，并揭示数字身份在大数据背景下的新特征。

二是针对当前对数字身份伦理问题的研究缺乏系统性的情况，综合性的对数字身份伦理问题进行整合研究，在总结他人观点的基础上，提出自己的看法，以寻求大数据伦理问题的解决之道。

## 第二章 大数据及数字身份概述

### 2.1 概念

#### 2.1.1 大数据的概念

随着信息科技的进步，互联网公司的崛起，云计算的普及和物联网热潮的出现，大数据强势的登上了历史的舞台，并迅速占领了公众的视野，引发了各行各业的参与和革命。大数据对经济的发展、科技的进步和社会的繁荣具有巨大积极意义，但是在大数据本身却是一个宽泛的概念，缺乏统一的定义，根据各个研究方向和研究目的的不同，都有不同的定义和范畴，而且看起来似乎都有其合理性。

麦肯锡全球研究院（McKinsey Global Institute）将大数据定义为：一种规模大到在获取、存储、管理、分析方面大大超出了传统数据库软件工具能力范围的数据集合。

第一家信息技术研究和分析公司 Gartner Group 把大数据定义为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来适应海量、高增长率和多样化的信息资产。

在《大数据时代》中，舍恩伯格提出了大数据具有的三个特点：1. 需要全部数据样本而不是抽样；2. 关注效率而不是精确度；3. 关注相关性而不是因果关系。

国内学者方环非<sup>1</sup>认为，就大数据本身指的是一种数据集合或信息资产。在这个方面麦肯锡全球研究所和 Gartner 的这一定义是正确的，但是如果细究之下会发现，大数据这一词时多指大数据技术，或者至少是在技术层面讨论这个概念。而且由于方向与研究目标成果的不同，对大数据的界定也各不相同，这些不同的界定又同时都具有其合理性。

虽然因为大数据具有这样或者那样的不同个性，难以精确的对大数据下普遍的定义，但是业界还是基本认同由著名的国际数据公司（IBM）定义的海量的数据规模（Volume）、快速的数据流转和动态的数据体系（Velocity）、多样的数据类型（Variety）、巨大的数据价值并且价值密度较低（Value）四个特征，合称 4V 特征，是大部分学者与专家都能够认同的大数据所具有的共性特征。

然而，由 IBM 一开始提出的是“5V”特点，即是 Volume（海量）、Velocity（快速）、Variety（多样）、Value（价值）和 Veracity（真实）。随着近些年大数据技术

<sup>1</sup> 方环非. 大数据: 历史、范式与认识论伦理[J]. 浙江社会科学, 2015(09): 113-120+160+2.

的发展，Veracity（真实）不再成为人们追求的目标，人们逐渐认识到数据质量才是能否发挥大数据价值的关键因素。

### 2.1.2 数字身份的概念

身份是由多种属性构成、定义人是谁或者是什么的立体的无限集合。具体可以分为生物身份，包括指纹、虹膜和面部特征等遗产信息；社会身份，包括父母、儿女和亲戚朋友等；地域身份，包括出生地、居住地和工作地等；法律身份，包括身份证、户口本和驾驶证……

而数字身份起源于人们对在线交易的需求，其来源可以追溯到因特网开放之初，在当时，人们便拥有了相应的数字身份，并且根据身份的不同也具有不同的权限，比如网站的管理员和访问者等等。当然数字身份的定义也远不止涉及管理者和访问者这么简单。对于数字身份的理解较为多样化，没有一个确定的定义，主要有以下几种观点：

#### ①从技术角度理解数字身份

从技术角度理解，数字身份指的是一种数字签名技术，是计算机领域的专用术语，指一种在互联网中，双方或者多方交换数据时，用来互相保证身份的一种数字签名技术，可谓保证数据信的第一道安全门户。赵翔认为，和人们现实中用个人签名来表示自己的身份一样，数字签名技术是人们在网络上用来表示自己身份的一种方式，“数字签名成为了手写签名有效合法的替代者”<sup>1</sup>。石贵珠认为，数字签名技术不但可以是为了在网络中确认彼此的身份，还是为了确认文件的完整性，“确保传输电子文件的完整性、真实性和不可否认性”<sup>2</sup>。李丽新认为源于密码学的数字签名技术同时也是为了防止他人伪造身份或者篡改数据，“数字签名就可用来防止电子信息因易被修改而有人作伪；或冒用别人名义发送信息”<sup>3</sup>。

#### ②从直观角度理解数字身份

从直观的角度理解数字身份，就是人们在网络上数字化的身份信息，比如各种证件号码、证书号码、卡号、手机号、IP 地址、数字密码等等以数字形式体现的某个个体或某个共同体在社会系统中的定位。谢刚认为数字身份是一数字化形式存储于网络上、用以标识个人身份特征的一种方式，“数字身份是……以数字形式存储和传播的信息所构成的个人身份”<sup>4</sup>。黄雯认为，数字身份是一组描述一个

<sup>1</sup> 赵翔.数字签名综述[J].计算机工程与设计,2006(02):195-197.

<sup>2</sup> 石贵珠.利用 RSA 密码系统进行数字签名[J].太原科技,2007(03):82-83.

<sup>3</sup> 李丽新,袁烨.网络安全通讯中的数字签名[J].现代情报,2007(08):53-56.

<sup>4</sup> 谢刚,冯纛,李治文,李文鹤.大数据时代电子公共服务领域的个人数字身份及保护措施[J].中国科技论坛,2015(10):34-38.

人或事物的数据，是“一个在线身份，或者说是网络身份”<sup>1</sup>。肖俊芳认为，数字身份不仅仅包括人的身份信息，还应该包括所有证明此人身份和地位的信息，比如“职业、学历、资产等一切表明社会身份及地位的相关属性的认证”<sup>2</sup>。

### ③从自我表述的角度理解数字身份

也有很多学者认为数字身份是由个人在线建构的网络中多元化的虚拟身份，是一种社会身份在网络上的体现，是通过参与网络对自我的表达。莫颖怡就将其表述为一种身份的建构，并认为这不仅仅是单方面的表述自我，还“包含了一种对反馈的渴望”。钟瑛将其表述为一种网络上的角色扮演，是“个人选择怎样的角色来表现自己”<sup>3</sup>。张欧阳认为其脱胎于传统身份却摆脱了传统身份上诸多社会因素的限制，“只关注于自己内心的情感和直观的利益诉求”<sup>4</sup>。

### ④从更宽泛的角度理解数字身份

从更加宽泛的角度来理解，则可以将一切的个人数据信息都包含在数字身份的体系之内，包括微博所发的博文、QQ空间的日志、微信的朋友圈和购物的浏览记录等个人在网络空间所留下的一切信息。Kinderlere将数字身份定义为“有关一个人的所有在数字上可得的信息的总和”。维基百科上把数字身份定义为作为一个人的网上活动产生的信息的整个集合，包括用户名和密码，在网上搜索活动，出生日期，社会保障，购买历史等……如果这些信息是公开的，那么它们可以被别人利用来发现人们的个人身份信息。

## 2.2 关于数字身份特性的探讨

关于数字身份的特性一般可以认为有以下四点：

### 2.2.1 多元性

数字身份具有多元性，一个人可以有很多个不同的数字身份，比如一个人喜欢在豆瓣上讨论电影，那么他就是作为电影爱好者的身份，如果一个人爱在淘宝京东上购物，那么他就是作为网购达人的身份，如果一个人喜爱在社区中与人聊天，那么他就是作为社交达人的身份。邱仁宗认为，个人可以根据环境条件的不同有很多个不同的且有效的数字身份，“一个人可不止一个有效的数据身份，其特征可根据情境、应用的目的或所获服务种类而有不同”<sup>5</sup>。钟光荣认为，人们在网络几乎可以拥有任何自己想要的身份，并且这些身份可以并行不悖，“可以随

<sup>5</sup> 黄雯,翟晓梅.数字身份与数字裂沟的伦理分析和治理研究[J].中国医学伦理学,2014,27(01):15-17.

<sup>6</sup> 肖俊芳.从郭美美伪造网络身份谈开去[J].上海信息化,2011(09):18-20.

<sup>1</sup> 钟瑛,刘海贵.网络身份的意义探析[J].复旦学报(社会科学版),2003(06):78-82.

<sup>2</sup> 张欧阳.网络民主的要素分析[J].兰州学刊,2012(12):176-180.

<sup>3</sup> 邱仁宗,黄雯,翟晓梅.大数据技术的伦理问题[J].科学与社会,2014,4(01):36-48.

心所欲地扮演想要扮演的任何角色，有时甚至可以同时选择两、三种角色参与同一种游戏活动”<sup>1</sup>。王贵斌认为，网络是与现实一脉相承的，是一种“多元文化社会”<sup>2</sup>。

### 2.2.2 可伪性

数字身份的构建受主观因素影响很大，由个人提供的身份属性可以真实的，可以是片面的、可以是历史的、可以是匿名的、可以是篡改的、更加可以是伪造的，因此网络空间中的数字身份自然而然具有可伪性。杨晓楠认为黑客可以窃取用户信息，利用软件伪造他人身份，“犯罪分子仍然可以利用身份伪造软件或者其他的技术手段躲避追查”<sup>3</sup>。方环非认为如果个人数据不受到保护，那么大数据公司也可以伪造用户身份，“我们的个人数据将会影响我们的身份认证……数据公司能够形成用户的身份，后果是不可想象的，甚至会出现身份的混乱、伪造等”<sup>4</sup>。

### 2.2.3 流动性

数字身份并不是一尘不变的，而是可以随着时间、地域、工作环境和家庭生活等情况变化而变化的。黄雯认为人们创建数字身份所提供的文字和图片信息可能会随着时间的流逝而变得片面化，时过境迁，即使当时较为完整的资料也会变得陈旧和不全面，“这些个人资料可能会被保留很长的时间，可能会被其他人在非原始情景之下使用，这时历史性描述和信息同步之间的区别就很容易变得模糊，造成当前身份和历史信息的混淆”<sup>5</sup>。刘卫红认为其具有流动性和可变性“身份不是与生俱来，而是在社会化过程建构而成的，具有多重性和流动性的特征”<sup>6</sup>。马云驰认为网络技术使得人的网络身份变得难以确定，“流动性一定程度上更加强化了现代社会的这一特征”<sup>7</sup>。

### 2.2.4 超域性

数字身份作为网络环境中人的社会身份，自然而然的可以使得人们越过千山万水在网上互相交流、学习和交易等等，因此数字身份借由互联网得到了超越地域的特性。祝菲认为数字身份让人们不用面对面的进行可以用来传递信息，印证身份，进行商品交易，只需要通过网络进行数字身份确认就可以了，“无论身处哪一个城市，人的身份都可以用数字身份来加以证实，用数字身份来进行沟通和

<sup>4</sup> 钟光荣.网络“自我”与现实“自我”的比较[J].边疆经济与文化,2005(03):68-70.

<sup>2</sup> 王贵斌.网络参与的身份建构[J].新闻知识,2011(04):49-50.

<sup>1</sup> 杨晓楠.互联网实名制管理与公民个人信息的保护[J].情报科学,2012,30(11):1613-1616.

<sup>2</sup> 方环非.大数据:历史、范式与认识论伦理[J].浙江社会科学,2015(09):113-120+160+2.

<sup>3</sup> 黄雯.大数据时代信息通信技术应用:伦理管理和政策研究[D].北京协和医学院,2014.

<sup>4</sup> 刘卫红.社会性别与身份认同的语言建构[J].江西社会科学,2014,34(07):238-242.

<sup>5</sup> 马云驰.匿名、流动性与道德需求[J].伦理学研究,2007(03):38-41+67.

交流”<sup>1</sup>。鲍跃华认为这种超地域性让人们具有了多重的身份，“每个个体身份都成为‘没有明确染色体的阶级’是一种阶级与身份的多重体”<sup>2</sup>。

### 2.3 大数据对数字身份的影响

由于大数据的巨大影响力，相关性收到了极大的重视，根据大量人们在网上冲浪所留下的痕迹可以精确到追踪到个人，由于数字身份是包含宽泛的个人所有数字信息的总和，大数据的强大能力可以让人们根据以往不受人们重视的部分数字信息精确追溯到重要的数字身份信息，个人的信息隐私存在着极大的暴露可能，甚至很可能会直接影响到人们的物理身份，数字身份也在大数据背景下具有了新的特征。

第一，数字身份内涵宽泛化。不单是数字身份的范畴扩大，其概念也不断的泛化。大数据背景下，数字身份的内涵已经由当初计算机网络领域的专业术语变为了个人在网络留下的整个数据集合，其概念也被定义为一个宽泛的网络集合，同时数字身份的涵盖范围由单一的专业领域扩展到几乎所有在线网络领域，内涵范围得到了极大的扩大。

第二，数字身份越来越容易被追溯。大数据时代的相关性使得通过大量的网络痕迹追溯到人们的隐私信息成为可能，数字身份同样也是如此，通过其他信息追溯到数字身份中的交易身份已经切实可行，利用大数据技术并不需要耗费多少功夫，尤其是现今大数据技术不断呈几何指数增长的环境下，追溯数字身份变得越来越简单。

第三，数字身份中其他信息的重要性得到极大提高。Snllivan C 把数字身份分为“交易身份”和“其他身份信息”，而交易身份是可以从单纯的个人身份信息中提取出来，同时“其他身份信息”不具备交易功能，提供个人基本资料和交易目的等一般信息。在大数据时代的背景下，通过一些看似无用的网络痕迹可以轻而易举的找到人们的其他身份信息，甚至是交易身份，固然交易身份是极其重要并直接关系到个人隐私的，此时的大数据背景下其他身份信息的重要性也得到了极大的重视，其在安全和隐私方面的重要性甚至已经不亚于单纯的交易身份问题，必须指出这里的其他身份信息指的是，除交易身份以外的一个人在网上活动产生数据的集合，而非 Snllivan C 定义的由上传信息系统的个人基本信息组成得“其他身份信息”。

第四，数字身份信息永久保留。不同于储存在我们自己电脑中的数据，我们可以随时删除，大数据的特性让我们的数据一旦接入互联网，就会永久留下痕迹。

---

<sup>1</sup> 祝菲.信息时代不能忽视的文化身份[J].理论月刊,2008(08):79-81.

<sup>2</sup> 鲍跃华.从身份认同到角色认同:党外代表人士代表性的现代转换[J].中共浙江省委党校学报,2012,28(05):76-81.

比如，我们上网购物和微信的聊天记录，都被阿里巴巴和腾讯存储在云端，我们随时可以收到的各种推荐商品和服务的广告，就是这些公司分析我们的数据之后定向给我们推送的。我们可以很容易删除我们自己的聊天记录和浏览信息，但却很难删除我们在云端所留下的踪迹。例如，FACEBOOK 中虽然可以删除自己的一条信息，但是你却无法删除被别人的转发或者提及，他们甚至可以讲你的记录截图，做成文档等存储，总之就是一旦上网，你不要指望能完全删除你的数据和信息。

根据国内外的研究情况，笔者把数字身份在大数据背景下的伦理问题分为 4 个部分，分别是数据权利被侵害的问题，隐私安全问题，公平问题和数字身份认同问题。

### 第三章. 大数据背景下数字身份的数据权利被侵害的问题

数字身份既然是一个人在网络活动上所留下的所有数据信息的集合，那么有关这些数据的权利自然而是数字身份伦理问题所要讨论的话题之一。在小数据时代，个人数据的种类和数量都比较少，同时也缺乏必要的整合与分析工具，所以实际上数字身份的可追溯性较弱，实际价值偏低，因而数据权利并不为大多数人所重。而大数据时代所带来的数据爆炸，整合分析工具的不断涌现，使得数字身份的可追溯性极大增强，使我们能够不断地从数据中挖掘价值，因此大数据就成为了一座宝库，人们在不断地认识到数据价值的同时，数据权利也成为了我们不容回避的话题。

数据权利指的是以人民为主体的私有权利，学术界一般认为个人的数据主权就是数据权利，即用户对自身数据应有自我决定的权利和自我控制的权利。数据权利（Right of Data）这一概念，由 2010 年卡梅伦就任英国首相时首次提出，他同时指出数字权利是公民的一项基本权利。当今中国法律并没有数据权力这一具体定义，笔者所提及的数据权力从根本上来说，是一种由多种权利互相组合而成的权利集合体。

由于大数据技术的特性，数据信息往往掌握在少数企业或者政府手中，人们对自身的数字身份信息缺乏实际的掌控权，很多被数据主体并不能够顺利的使用和支配这些数据信息，因为数据信息根本就不在他们手中，全部在数据记录者手中，这些资产与所有者二者是相互分离的，这就造成了数据主体难以支配或使用这些数据信息，甚至是知道这些数据信息被存储在哪里，被用来做什么，怎么被使用都是很困难的，以网店交易为例，很少有平台提供商和网店愿意告诉顾客他们的交易数据和浏览信息都用来做了什么，被交给了谁。



所以数据主体虽然在理论上拥有数据信息的所有权，但是实际上根本无法支配这些数据信息，而实际上并没有这些数据所有权的企业或者商家却可以实际上支配这些数据，却完全不需要经过数据主体的同意，甚至不需要告之他们。另外，单个的，孤立的数据主体的数据毫无价值，只有当这些数据主体们的数据信息大量集中在一起时，这些数据信息才能产生价值。这些数据信息被企业在虚拟空间上集中在一起，但是在物理空间上，这些数据主体确实互相分离的，他们几乎不可能可以集中在一起面对商家或者企业，来明确的保障自己的数据所有权。还有，数据主体缺乏分析大量数据信息的工具和能力。大量数据信息的分析必须借助大型计算机和复杂的高精度的专门的算法，这是一般人难以拥有的条件，所以即使我们拥有这些数据信息，我们也拿这些数据毫无用处，只能白白占据我们的硬盘空间。因此，我们的数据权利经常遭到侵害。

由于大数据技术的飞速发展，数据已经成为一种财产，数字身份作为个人在网络活动产生的所有数据的集合更是成为一座宝库；同时由于大数据能够暴露人的偏好的特性，数字身份也成为了个人现实人格的一种外化的形式。基于这两点理由，有学者认为个人数据权利可以分为数据财产权和数据人格权，刑法修正案（七）等关于个人信息保护条款，正是数据财产权和数据人格权的法律体现<sup>1</sup>。

### 3.1 数据人格权受到侵害

在大数据时代，我们无时无刻不在产生数据，这些海量的数据基于我们日常生活产生，通过复杂的算法分析与组合，可以追溯到我们个体，甚至可以洞察我们的行为习惯，消费模式，甚至能够预测我们的未来行动，理所当然的具有我们的人身特性，是一种新形式的外化的人格反映，所以我们对这些数据应该享有人格权。肖冬梅和文禹衡认为，“数据人格权不同于隐私权，因为传统的“隐私权”概念涵盖不了所有个人数据，个人数据既包含隐私数据也包含非隐私数据”<sup>2</sup>。研究人员一般将数据人格权分为知情同意权、修改权，封锁权和遗忘权四个方面。

数据人格权受到侵害的例子比比皆是，比如 iPhone 默认将“常去地点”设置为开启模式，此时手机会记录下使用者前往某地所需时间、造访频率和停留时长，以此推算出使用者经常去的地点，位置具体到街道、小区，并详细记录每处停留时间，精确到分钟。该功能选项默认打开状态，用户升级前未被充分告知。苹果收集的数据是否会用作他途，除了他们自己别人无法得知。很明显，苹果公司记录和收集用户的常去地点数据，但却并没有充分告知用户，侵犯了用户的隐私权。其实，不仅仅是苹果公司，很多公司都会“悄悄地”记录和收集用户的个人数据，比如腾讯公司通过 QQ 和微信来监控用户的聊天，根据聊天过程中发送的敏感词

<sup>1</sup> 张毅菁.数据开放环境下个人数据权保护的研究[J].情报杂志,2016,35(06):35-39.

<sup>2</sup> 肖冬梅,文禹衡.数据权谱系论纲[J].湘潭大学学报(哲学社会科学版),2015,39(06):69-75.

汇和文档等来决定是否给予用户封号或者警告等处理。

### 3.1.1 知情同意权受到侵害

知情同意权是数据人格权的基石，可再进一步细分为知情权和同意权。知情权是同意权的前提条件，同意权是知情权发展后续，二者缺一不可。任何企业和政府在收集和处理个人数据前必须告知数据主体并征得其同意。数据知情同意权不止是规范采集、处理行为本身，还应该包括采集数据的目的、用途及其后续的变化等，以及处理数据的方式、程度等及其后续的变化都需要告知数据主体并获得其同意<sup>1</sup>。

数字身份作为一种个人数据的集合，个人做为数字身份的主体自然而然具有数字身份的最终决定权，他人对个人数据的采集、整理、存储和使用等等均需经过主体同意，否则就是一种侵权行为。有学者<sup>2</sup>指出知情同意权并非绝对禁止他人使用自己的数据，而是数据主体对自身数据的控制权利。主体的数据一旦被公开后，主体必须拥有知情的权利——究竟是什么个人或机构因为什么原因而掌握了我的什么数据。

企业不经过用户同意就私自使用用户的数据信息是不合理的，上文所说的苹果公司进行的不充分告知用户的数据收集明显是不合理的，但是这里可能会有很多人辩解说，不论是苹果还是腾讯还是 FACEBOOK 都已经提前通知了用户他们的数据收集，记录和使用的详细情况，理由是他们在提供给用户相应的服务之前已经列出了详细的服务条款来让用户选择是否了解和接受，但是各种服务协议和同意书过于冗长和繁杂，更不用说其中还包含着许多令人费解的专有名词和专业术语，以及大量的法律条款，除非刻意去研究，否则很难有人能够耐着性子读完这些协议，更别提弄懂协议中的每一条款所包含的内容，以及分析这些条款是否合理了。

腾讯公司注册 QQ 的相关条款和隐私政策，三项规则和政策，共计有 1 万 2 千多字，而注册淘宝网所需阅读并同意的政策和规则有三篇，共计 2 万 7 千多字，堪比一部中篇小说。大部分人注册 QQ 只是为了和朋友聊天，注册淘宝只是想从网上买个东西而已，犯得着为此看一篇甚至几篇让人难以理解的无聊的文章吗？所以，绝大多数人都会选择忽略这些条款，直接点击同意，他们根本不关心条款中写了什么，也不关心协议是如何运作的，他们只是想要获得自己获得聊天，网购等服务而已，即使很多人都明白这些条款中所带有的法律词语的重要性，也很难有人能静下心来逐字逐句的看完，他们认为不值得为此浪费时间。而且，即使你能耐着性子看完也没有多大的作用，如果你不同意这些条款，你就无法享受到诸

<sup>1</sup> 肖冬梅,文禹衡.数据权谱系论纲[J].湘潭大学学报(哲学社会科学版),2015,39(06):69-75.

<sup>2</sup> 肖冬梅,文禹衡.数据权谱系论纲[J].湘潭大学学报(哲学社会科学版),2015,39(06):69-75.

如 QQ、淘宝等所提供的服务，这多少带有点霸王条款的味道。所以，在 1. 条款太长且术语过多，难以理解，浪费时间，2. 如果不同意便无法享受相应的服务的情况下，人们只有被迫接受这些条款，而不加以详细考察，实际上是我们是别无选择，除非你选择不接受这些服务所带来的便利。但是，当你看到身边的同事，朋友，亲戚，都在使用 QQ 和微信聊天，都在使用淘宝京东购物时，你会怎样呢？假设你还能经得起这种诱惑，那么，当他们不断告诉你 QQ 微信有多么方便，京东淘宝有多么便宜时，你还能忍受的住吗？我们再假设你还可以坚持底线，绝不用这些服务，但是，当公司里，单位里全都改用微信群和 QQ 群内通知重要事情时，你还能置身事外吗？

### 3.1.2 修改权受到侵害

数据修改权也称为数据更正权，是指数字身份的主体有权修改或者授权于他人修改自身数据的权利。主体不但有权利修改或者授予他人修改自身的数据，同时也有权利禁止他人未经本人授权而擅自修改本人的数据。如果主体认为有关自身的数据错误或者不完全，可以申请对自己的数据进行补正。

互联网上的身份信息经常出现滞后的现象，跳槽以致工作地点变动，升学让学籍和户口所在地改变，婚姻关系让家庭组成情况改变等等，这时候网络上的信息往往不会随着个人身份信息的变化而及时更新，这会给个人的工作和生活带来不便，比如甲在某网站注册了用户，并留下自己的手机号，让该网站提供信息推送的服务，在这期间甲更换了手机号，此注册手机号的户变成了乙，此时网站工作人员并不知道这个信息，依然继续向此手机号推送信息，此时受到推送信息的是乙，乙可能不需要此类推送，认为这种推送骚扰了自己的生活，而甲需要这些推送信息，但是因为手机号更换的原因已经无法受到推送信息，此时甲就有权利要求该网站工作人员将推送信息的服务转到自己的新手机号上，同时乙也有权利要求该网站停止对自己手机号的推送服务，如果该网站拒绝更改推送手机号，并继续向原手机号推送服务，就侵犯了甲和乙的数据修改权。

更为常见的是网络上随处可见的虚假信息，有些商家为了迎合用户的口味和偏好，往往并不是完全按照事实来报道新闻，他们会将事实情况稍加更改或者与其他无关事件相互联系，或者直接根据道听途说的信息凭空臆造，不追求信息来源的真实性，比如大家常见的各种花边新闻、明星绯闻、娱乐八卦等等，很多追星族和偶像明星们的粉丝非常喜爱这类新闻，但实际上这些报道可信性度并不高，商家如果不加以考察就随意在网上发表出来，无疑是侵犯了当事人的名誉权，当事人有权行使自己的数据修改权，要求商家更正自己的报道信息，并向自己赔礼道歉。如果商家不及时更正信息或者不愿意更正信息，就是侵犯了当事人的数据修改权，当事人可以运用法律武器来维护自己的权利。

### 3.1.3 封锁权受到侵害

封锁权是指数据主体可以合法的请求数据管理者暂停对特定数据的使用。值得注意的是不仅仅只是针对数据主体自身的数据，也可以针对非自身的数据请求封锁，但必须要有合法的正当依据，比如举报他人数据不真实或者有争议，要求对其数据进行封锁。

数据封锁权就是数据主体在法律规定或者与数据管理者约定好的事情发生时，可以要求数据管理者通过约定好的方式对特定的数据进行暂时性的冻结和封锁处理。一般来说，事先约定好的事情是数据主体相关数据的完整性或者正确性处于受威胁的状态时。

大数据时代，数据一旦出现在了网上就很难被自己完全掌控了，数据主体如果认为自己的数据信息已经被泄露或者可能有泄露的风险，在适当条件下可以向服务提供商请求封锁自己的数据，例如青少年玩网络游戏时可以向服务商要求绑定手机密保卡，在登录游戏的过程中必须输入密保信息，一旦连续多次输入错误就会将此账号锁定，以此来防止自己的虚拟资产遭受损失。我们常用的 QQ 和微信如果不是在自己经常登录的地点登录并使用，腾讯一般会提示用户非常用地点登录，提醒用户注意自己的账号安全，用户如果收到此种提示，可以向腾讯要求更改自己的账号信息，防止自己的账号安全收到损害。如今的智能手机一般都会要求用户注册手机账户，比如苹果手机的用户如果发生手机丢失的现象，可以第一时间从别处登录自己的苹果账户，要求苹果公司锁定自己的手机设备，封锁丢失手机上的信息，防止自己的数据信息泄露。

### 3.1.4 遗忘权受到侵害

数据遗忘权又称数据删除权，是指在合法的情况下，数据主体有权要求互联网企业删除自己的个人数据，一般是为了维护自身的名誉，保证身份安全等原因，最终目的是要使得互联网企业不能再通过数据识别和追溯到数据主体。

数据遗忘权分为狭义和广义两种含义，狭义的数据遗忘权是指数据主体不再需要数据服务提供商的服务时，有权要求数据服务提供商删除数据主体的所有数据；广义的数据遗忘权是在狭义的数据遗忘权的基础上，再增加用户有权要求数据服务提供商删除与自身有关的不真实的、有害于自身名誉或者有损于自身利益的数据的权利。首先应该由法律或者双方约定好删除的条件，一旦满足条件之后，必须由数据主体自身主动申请，否则数据管理方无权删除数据信息。

在大数据时代，数据信息不同于储存在我们自己电脑中的数据，我们可以随时删除，一旦进入互联网，就会永久留下痕迹。比如，我们上网购物和微信的聊天记录，都被阿里巴巴和腾讯存储在云端，我们随时可以收到的各种推荐商品和服务的广告，就是这些公司分析我们的数据之后定向给我们推送的。我们可以很

容易删除我们自己的聊天记录和浏览信息，但却很难删除我们在云端所留下的踪迹。例如，FACEBOOK 中虽然可以删除自己的一条信息，但是你却无法删除被别人的转发或者提及，他们甚至可以讲你的记录截图，做成文档等存储，总之就是一旦上网，你不要指望能完全删除你的数据和信息。为了规避垃圾邮件、消除不良事件的影响等，合法遗忘或者删除自身不需要的数字身份信息就变得愈发重要。

### 3.2 数据财产权受到侵害

在大数据时代，数据已经具有了财产性，被视为一种信息资产，越来越多的体现出了商业价值，并且能够进行商品交易，所以理所当然的具有了财产权的属性。研究者一般较为认同将数据财产权分为采集权，可携权，查询权，使用权，获益权五个部分。

有学者认为，数据是大数据时代的一种新型资源，它就像资本一样能够不断增值，数据已经成为一种“资产”，而包含个人所有上网数据的数字身份已经成为了巨大的资产，这为这些为数据法权化——数据成为法律意义上的财产并确立数据财产权提供了基础。这种数据资产收到侵害的形式一般有以下几种。

#### 3.2.1 采集权受到侵害

采集权是指数据主体有权利同意或者不同意对其个人数据进行采集的权利。

数据是大数据的核心，而数据采集又是数据的核心。无论是进行数据的存储、分析、挖掘还是预测，都要以采集数据为前提，因此数据采集权是数据权利中的第一道屏障，有学者认为数据采集权是数据权利束中最基础、最核心的权利。

荷兰隐私数据保护机构 DPA 与 2014 年发表声明，称 Google 不经过用户允许就随意采集和存储用户的网络数据（其中包含搜索引擎数据、YouTube 网络视频数据和其他互联网产品）并用于广告方面，这些举动谷歌从 2012 年已经开始。DPA 方面要求谷歌在 2015 年 2 月份之前规范其在荷兰市场采集网民数据的做法，同时谷歌方面将可能面临 1500 万欧元（约合 1900 万美元）的罚款。很明显，谷歌在未经用户许可的条件下对用户数据信息进行了非法采集，侵犯了用户的数据采集权。

用户的数据采集权不仅仅体现在数据采集行为的开始，也可以体现在数据采集的过程中，用户一旦在授权企业采集自己数据的过程中改变自己的想法，可以随时向企业要求停止对自己数据的采集，比如微软的 Windows 10 刚发布时，其系统强制收集用户数据，虽然微软不断表示只有匿名数据才会被提交到远程服务器，并且所有信息都是被加密存储在绝对安全的数据中心，因此不可能有人能访问这些收集到的信息，但是依然有很多非常重视隐私的用户对此感到极为不满，终于在 1 个月之后，微软迫于压力为企业版的用户提供了一个额外选项——可以关闭数据收集的功能——用户终于可以在关闭 Windows 10 对自己的数据采集功能了。

### 3.2.2 查询权受到侵害

数据查询权指的是数据主体有权利阅览查询数据管理者所持有的自身数据的权利。数据获取权也是数据查询权的另一种表述方式。具有查询权利的数据包含数据主体被收集、处理与利用的各个环节的数据，涵盖数据生命的整个周期。

数据查询权是用户监管自身信息状况的主要途径之一，用户可以通过对自身信息数据的查询来确认自己的身份信息是否正确，如果不正确或者过于陈旧，则有权利要求服务提供商予以更正。对于用户上传的数据信息，服务提供商有义务保持用户数据的完整性和准确性，以便于用户随时进行查询，如果未经用户同意就更改或者关闭用户权限，使用户不能查询自身数据信息，就是侵犯了用户的数据查询权。

阿里巴巴不但随时为用户提供浏览自己购物记录的服务，还未用户提供查阅自身浏览足迹的服务，微信用户可以随时查询自己的朋友圈并观看有多少人给自己点赞，有多少人进行评论，智能手环用户通过与手机 APP 的同步，可以观看自己的行走路线和睡眠信息等等，都体现了商家对用户数据查询权的重视。

### 3.2.3 可携权受到侵害

可携顾名思义就是可以携带，意指数据主体可以把自身的数据从一处迁移到另一处进行保存。可携权是指数据主体有权要求掌握其数据的相对方(如服务商、运营商政府等)协助其将个人数据在不同系统或设备等载体之间进行迁移、保存。

也有学者认为数据可携权就是数据主体能够在毫无阻碍的情况下将自身数据和其他资料从一个数据管理处移动到另一个数据管理处的权利。数据可携权的基础是数据查询权，如果连查询个人信息的数据都不可以，那更加谈不上对数据的迁移和保存了，因此数据的可携权包含两方面的内容，一是对个人数据信息的获取，二是对个人数据信息的迁移，无论是在数据获取还是在数据迁移的过程中，数据控制者都不仅不能干涉、剥夺用户的数据可携权，还必须配合、协助用户进行数据信息的获取和迁移。有些服务提供商拒绝为用户提供相关数据，或者对用户信息的转移请求不予协助，就是对用户数据可携权的侵犯。

比如微博中就可以随意导出自己的博文，腾讯提供将 QQ 的信息转移到微信中的服务，谷歌用户可以使用 Google+ 里的“数据自由”(Data Liberation)功能下载并导出自己数据文本，媒体信息，包括自己上传的照片、视频、自己的联系人好友等等。相应的 Twitter 也推出了下载个人信息文档的服务，Twitter 以邮件形式将发送一个压缩包给予用户。

用户解压压缩包之后，可以通过点击 html 文件清楚地看到自己之前所有的 Twitter 信息还可以按照日期顺序对这些信息进行排序等等，这些都是网络服务商尊重用户数据可携权的体现。

### 3.2.4 使用权受到侵害

数据使用权意指数据主体有权使用自身数据的给自己增加效益带来收益的能力。但是值得注意的是，我们一般人作为数据主体，很难完全通过自己利用这些数据给自己带来收益，因为或许大数据企业能够通过数据模型和复杂算法挖掘出数据中潜藏的我们的行为模式和消费习惯进而向我们进行精准营销，但是这其中必须的对数据进行二次分析和挖掘的能力并不为我们普通人所掌握，而是少数大数据企业和政府机关的“特权”，所以我们很难直接占有和掌控这些数据，但并不能因此而忽略用户的数据使用权<sup>1</sup>。

数据使用权目前的主要是被授予或者转让他企业和国家来使用，由企业或者国家来代替数据主体行使我们的数据使用权，因为企业和国家掌握着大数据的关键技术，可以将大数据转化为实际利益。企业使用用户数据的权利是用户所给予的，企业没有权利将用户的使用权私自赋予第三方使用，除非经过用户许可，同样也没有权利将用户数据用于非约定的其他目的，更加没有权利不经用户同意就随意行使用户的数据使用权力，否则就是对用户权利的侵犯。同时用户对于自身数据信息的使用权可以授予多家企业使用，并不必为一家企业所独占，企业如果要求独占用户的数据信息，也是对用户权利的一种侵害。在授予企业数据使用权时，并不影响用户本人对自身数据信息的使用权，个人依然可以将数据信息用作其他用途。

### 3.2.5 获益权受到侵害

获益权是指数据主体通过自身数据获得利益的权利。这种获益权即使不是因为数据主体直接使用数据而获利也依然存在，如果数据管理者通过数据主体的个人数据获利，那么数据主体就有权利要求数据管理者给予其适当的报酬。获益权也可称为数据请酬权，当信息管理者旨在以盈利为目的，采集、处理与使用个人数据时，数据主体有权向数据管理者请求支付对等报酬的权利。值得注意的是数据获益权仅仅是指以盈利为目的的使用个人数据，如果是非盈利目的则不可以直接请求报酬。

数据现如今已被视为一种商品，电商通过分析用户数据，得到用户的消费偏好和消费习惯，改善自己的服务方式，比如提高物流效率、推送顾客偏好商品、推出打折促销活动等增加了自己的收益，那么电商是否应该为此向用户提供一部分收益来对用户进行回报呢？虽然目前对此学术界争议较大，但是笔者还是认为用户是有权获取回报的，因为数据能够创造价值，数据是个人人格的一部分，电商在利用数据获益的过程就是讲数据商品化的过程，理所应当对用户提供服务。

<sup>1</sup> 肖冬梅,文禹衡.数据权谱系论纲[J].湘潭大学学报(哲学社会科学版),2015,39(06):69-75.

目前用户的获益权还难以得到保障，因为电商将部分收益提供给每一位用户缺乏操作性，应为第一，每个用户的数据信息相比于海量繁杂的样本来说，可以说是九牛一毛，冰山一角，电商难以对此给予高额报酬，第二，电商的用户数量虽然庞大，但是此种维权行动必须要大量人员联合起来向电商予以索求，其中的沟通交流、协调商谈好肥的时间和精力难以想象，大部分人都会觉得太麻烦而且报酬较低而选择放弃，第三，电商的所给予用户的注册时的声明中一般都会有授权本公司使用用户数据的权利而无需支付费用的条款，诚然这属于一种霸王条款，但实际上没有多少用户会因为这样一个条款就放弃接受电商所提供的服务，比如注册淘宝用户所必须同意的《法律声明及隐私权政策》中就明确有一条款，“向您提供更契合您需求的页面展示和搜索结果、了解产品适配性、识别账号异常状态，我们会收集关于您使用的服务以及使用方式的信息并将这些信息进行关联”，这实际上就相当于淘宝使用自己数据信息的权利，几乎没有人会因为这一小小的条款就拒绝淘宝账户的注册。

## 第四章. 大数据背景下数字身份的隐私安全问题

在大数据时代，我们的数字身份信息都被上传到因特网，无处不在的摄像头和“第三只眼”无时无刻不在窥视我们的隐私，电脑病毒、黑客攻击等都对我们的身份信息虎视眈眈，我们数字身份信息的隐私随时面临着被泄露被曝光的危险，信息安全岌岌可危。人们通过个人电脑，手机登各种终端在互联网上交流，学习，工作，休闲和娱乐，几乎每一次对互联网的访问都被记录了下来。

而数据的存储格式使得数据集合不太可能将随着时间的推移而优雅地消失，他们更可能随着时间而变成可以互相影响，这些被收集的数据之间可以为了完全不同的目的来比较和合并。一个用户删除的信息可能仍然保留在一个以前已经收集的数据集合中——并可能继续存在好几代。基于新的数据存储方法和新的分析方法相结合的基础上，今天是匿名的数据，明天可能会成为识别个人身份的数据。数据的寿命及其意料之外的用途，会质疑目前研究人员保证用户隐私和匿名的能力——这样的事情甚至已经发生了。对收集和记录了什么样的数据，以及这些数据是否有可能在一段时间后，与其他数据结合以威胁到用户的匿名和隐私，这些都必须予以小心谨慎的对待。在可能的范围内，任何必要的，但可能具有识别意义的东西都应该被伪装，以防止未来对个人数字身份的追溯。

### 4.1 隐私暴露问题

隐私的丧失几乎成了家常便饭。我们平常浏览网站时，QQ 微信聊天时，发布微博博文时，海量的数据中就混杂着很多个人隐私信息，这些信息发布在网上的



同时就宣称着我们的隐私信息已经暴露，由些学者据此认为“隐私权已死”。同时，由于网络的自由性、开放性和隐匿性，我们的隐私信息变得更加容易暴露，保护隐私信息变得愈加困难。我们在办理银行卡，各种注册和登记之时，就要交出自己的姓名，年龄，身份证号码，银行账号等等。

2011 年，黑客攻击了日本索尼公司，并窃取到了索尼公司的近三千名用户的姓名和电子邮件将之公诸于众，此举可能已经造成西班牙、德国、荷兰和奥地利等国家一万零七百笔转账记录以及一万两千七百个美国信用卡卡号遭到泄露。索尼的这大规模隐私泄露问题也提醒我们，即使是跨国大公司的隐私保护也并非那么安全。稍有不慎，大数据就可能变成大泄露。隐私一般可以分为三类，身份隐私，行为隐私，喜好隐私，日新月异的大数据技术，让这三类隐私随时面临被泄露的威胁。

#### 4.1.1 身份隐私暴露问题

我们的年龄、身高、体重、血型、健康状况等生物学特征和个人财产、家庭关系等社会学特征都属于身份隐私信息，大数据技术的信息搜集能力可以使得这些隐私信息唾手可得，毫无隐私可言。

联动挖掘模式是大数据的技术特点之一，很多不法分子在经济利益的驱使下，利用大数据技术的这一特点侵犯人们的隐私权。商企经常运用大数据关联分析技术分析用户的偏好，而分析的过程常常会触及人们的身份信息隐私，使用关联性的数据挖掘技术对海量的复杂数据进行相对简单的数据分析并辅以复合运算技术，不管是故意还是非故意都可以披露用户的隐私身份信息<sup>1</sup>。

同时，大数据技术的应用使得医疗信息被泄露的几率大大增加。医疗健康信息往往涉及到用户的大量隐私，这些海量数据不但是开放共享的，也是实时动态的，同时也兼具交叉检索的特性，这无疑增加了身份信息暴露的风险<sup>2</sup>。

#### 4.1.2 行为隐私暴露问题

行为隐私指我们的日常生活和社会交往中各种行为的动态集合，一般表现为具体的某一种或者几种活动。

大数据时代，不但有着 360 度无死角的各种摄像头不分昼夜的监视着我们的生活，就连我们日常生活中使用的电脑、手机和各种智能设备也都会实时记录下我们的一举一动。

我们的每一次使用搜索引擎进行搜索、发布朋友圈动态、查询百度地图，都会被后台的软件记录下来，上传到云端储存，我们的行为隐私随时可能被暴露。

<sup>1</sup> 宋振超，黄洁. 大数据背景下网络信息的伦理失范、原因及对策 [J]. 理论与改革, 2015, (02).

<sup>2</sup> 王灵芝，郝明. 医疗大数据的特征及应用中的伦理思考 [J]. 医学与哲学(A), 2017, (04).

薛孚从当今的全景监控现象分析指出，无论我们认识到还是没有认识到，面对无处不在的监控，我们的隐私在现今都已经无从躲藏了<sup>1</sup>。在通过大量终端传感器进行的全景监控之下，通过公开或者非公开的对网络浏览历史数据进行采集的情况下，所采集的数据不管我们意识到还是没有意识到，都肯定包含这个人的身份信息数据，“主体的隐私都被侵犯”<sup>2</sup>。

互联网自身特性让我们的行为所产生的数据足迹永久保存了下来，大数据时代这种第三只眼遍布的情况的确侵犯了我们的隐私权利。数据足迹是智能设备自动将人类的一举一动进行信息采集，这些数据信息一旦进入网络就难以彻底清除，因此也就容易永久保存，不易消逝，在不知不觉中大数据早已偷窥了我们的秘密。

大数据的预测功能也同样对个人隐私信息造成了侵害。大数据技术可以智能的识破数据信息中包含的各种“意图”，泄露这些“意图”无疑会对我们的隐私和尊严造成损害。大数据预测系统使巨量结构化与非结构化数据中分离识别真实资料与意图，这就让大用户陷入以隐私披露为代价获取便利的个性化服务之尴尬境地，尊严受损现象层出不穷<sup>3</sup>。

#### 4.1.3 喜好隐私暴露问题

电商时代，网络商企使用大数据技术采集我们的浏览记录和购买信息，使用数据分析工具确定我们的偏好，对症下药的进行定向广告的投放，对我们的手机和邮箱进行狂轰滥炸，对我们的生活造成了严重影响。唐凯麟<sup>4</sup>认为虽然大数据集市的使用可以提高商企的效益增加他们的利润，但是也极有可能会打扰用户的正常生活习惯，以逐利为目的的商企侵害用户隐私的行为也会不断增加。消费者的性格、偏好及其消费习惯等十分有用的信息都可以通过分析消费者的购物记录被商家得出，这些有用的信息对企业 and 商场来说的作用就是潜在的利润，但对于我们普通的消费者来说，则意味着我们的隐私被暴露了。

从选择权角度来说，我们可能始终被大数据限制着，而我们选择的的可能只是大数据让我们选择的，而并非我们真正喜欢的，我们不仅处于被限制的状态，而且一直被大数据牵着鼻子走，我们可能会丢失更好的其他的选择，而不是让大数据来帮我们选择，这严重损害了我们的选择权利<sup>5</sup>。

从获取路径的合法性方面来看，有些企业绕过用户许可，直接非法收集、利

<sup>1</sup> 薛孚,陈红兵.大数据隐私伦理问题探究[J].自然辩证法研究,2015,31(02):44-48.

<sup>2</sup> 薛孚,陈红兵.大数据隐私伦理问题探究[J].自然辩证法研究,2015,31(02):44-48.

<sup>3</sup> 蒋洁, 陈芳, 何亮亮. 大数据预测的伦理困境与出路 [J]. 图书与情报, 2014, (05).

<sup>4</sup> 唐凯麟, 李诗悦. 大数据隐私伦理问题研究 [J]. 伦理学研究, 2016, (06).

<sup>5</sup> 陈仕伟, 黄欣荣. 大数据时代隐私保护的伦理治理 [J]. 学术界, 2016, (01).

用数据，出售个人喜好信息，严重威胁用户的隐私安全。很多企业直接使用用户的个人信息资料，并将其随意公布和贩卖给他人，却根本不需要用户同意，这对用户隐私构成侵权。

## 4.2 信息安全隐患问题

数字身份信息安全一般包括两方面的内容，一指的是数字身份信息本身的安全，数据的完整性，可用性，可靠性和保密性等，二指的是数字身份信息保护的安全，通过对数据进行物理备份，架设磁盘阵列和构建多套数据库的现代化数据管理方法对数据信息主动进行防护。数字身份信息安全涉及数据生命的全部周期，从产生，记录，存储，到分析，使用，废弃，同时也包括防止数据遭到泄露，破坏，修改，伪造等等诸多方面。

第一，几乎所有个人身份信息大数据都会有所涉及，对大数据稍加处理就可能造成隐私泄露，因此我们有责任，也有义务保证这些敏感数据不被滥用。第二，不仅仅是个人，身份信息一旦遭到泄露往往就会涉及到一大片群体，大数据技术的进步性很可能威胁的是所有公民的信息安全问题。在大数据时代，政府不再垄断着数据信息，信息在国与国之间的流动变得十分频繁。大数据时代所具有的开放和共享特性使得动员全世界的民间力量成为可能，这些民间力量结合在一起甚至拥有超越国家机器的力量。因此数字身份的伦理问题在大数据时代涉及的不仅仅是个人，而是包括所有公民在内的整个集体。第三，我们也同样需要保障数据的真实性和有效性，维护数据的完整性，不随意篡改伪造数据，不能随意割裂数据，同时也要防止他人恶意对数据进行捏造和篡改。第四，大数据虽然具有预测未来的神奇功能，但我们如此对此过于信任，就会导致对大数据的过于依赖，进一步变为唯数据主义。殊不知不但大数据的预测未来功能尚且有待商榷，数据本身的真实性，有效性也值得怀疑，因为现今数据的来源过于庞大，不少不法分子，网络骇客和不良商家出于各种目的，对数据进行各种伪造的篡改，要确保每一条数据的真实性和有效性根本就是不可能完成的任务，如果数据真的经过篡改和伪造，那么数据的真实性和有效性也就大打折扣了。因此，我们不能对数据太过信任。数字身份信息的安全隐患问题主要包括三点。

### 4.2.1 数据信息可信度降低

互联网中有大量的虚假、错误数据。例如，很多电商就被发现对交易数据进行造假，他们增加虚假的商品流量和刷取商品成交量。从网络电子踪迹的角度来说，这些虚假的数据的的确确是真的，但它们根本不可能如实的显示出人们的交易行为。这反映出大数据信息安全中存在着非常重要的数据可信度问题。

数据传播的过程会造成角信息失真问题，人工误差与版本变更都会严重影响

数据的可信度。造成数据信息失真的原因一方面是由于对数据采集过程进行了人工干预，以致数据产生误差与失真，另一方面也有可能是由于在老版本更新到新版本的过程中数据出现错误，这些都会影响到数据信息的安全。

数据可信度降低很大程度上也来源于对数据的造假。某些意图不轨者可能会遵循自己的目的修改和伪造数据信息，让数据看来更好看、更有说服力，但这些数据已经不具有反映真实情况的能力了。在背景条件一定的情况下，也存在认为伪造数据的可能，比如前文提到的交易量造假行为，引诱数据分析者得出错误的结论。因为数据量过于庞大，虚假信息藏身于其中无异于泥牛入海，人们很难判断数据真伪，从而做出错误判断<sup>1</sup>。

数据超载也是网络中的数据信息安全隐患的问题之一。数据超载是指不对数据进行有效的分析和挖掘，支队数据进行简单的堆积处理，这种直接扔给受众海量未处理数据的行为很容易造成受众的负载过重，难以消化的负面情况。当今数据垃圾在互联网中的比比皆是，人们处于超载模式中，受众对于海量信息无法完全理解<sup>2</sup>。

#### 4.2.2 身份信息数据被泄露

由于对大数据简单加以处理就会暴露我们的隐私，所以数据泄露对据我们有着巨大的威胁，比如 2013 年查开房网站导致 2000 多万人个人信息泄露。不仅仅是私人企业，政府机关同样也是数据泄露的重灾区，2015 年 4 月，深圳、上海、河北、安徽等省涉及数千万人员的个人身份证、社保参保信息、财务、薪酬、房屋等敏感信息泄露，这无疑给我们敲响了保障数据安全的警钟。

现在申请一个注册网站的用户名，都需要提供自己的身份信息，一些网站由于各种原因比如黑客攻击、服务器故障等原因造成数据泄露也让人们屡见不鲜，比如 2014 年索尼电影业就在黑客攻击中泄露了大量用户数据，包含 30287 份索尼影视美国公司的文件和 173,132 封电子邮件。除此之外微信 QQ 和微博等社交平台也是数据泄露的重灾区，用户自身发布朋友圈或者博文的时候，可能会无意带上各种地点、声音、图片和视屏的等信息，这也可能会造成自己身份信息据泄露。

从非主观故意的角度来说，人员失误和管理不善是造成身份信息泄露的主要原因。

一方面是数据管理人员的不当操作问题可能出现在任何应用软件之上，甚至是简单的取消和确认两个选项之上，人工的失误是难以避免的；另一方面是以内越来越复杂的数据存储系统也对应要求越来越高素质的管理运维人员，这也是造成身份信息泄露的原因之一。

<sup>1</sup> 冯登国，张敏，李昊. 大数据安全与隐私保护 [J]. 计算机学报，2014，(01).

<sup>2</sup> 罗弦. 网络新闻生产中大数据运用的伦理问题及编辑对策 [J]. 科技与出版，2015，(01).

从主观故意的角度来看，内部人员窃取用户信息倒卖给第三方以此牟利也是身份信息泄露的主要方式之一。商家通过挖脚内部开发人员、测试人员和管理维护人员，甚至是内部用户人员来获取敏感数据，进行不正当的商业竞争行为从而以此获利，这种行为比创造新的软件要容易得多<sup>1</sup>。

黑客攻击也是造成身份信息泄露的重要原因之一，计算机技术便捷有效、大数据体量庞大难以有效防护、大数据容易与其他敏感隐私相关联也从客观上加剧了黑客的攻击频率和效率。大数据因其自身数据量过于庞大，价值含量较高，天然具有易被针对和进攻的特点，“这也导致了大数据信息安全风险的加大”<sup>2</sup>。

接收数据信息的一方有责任对个人数据信息进行有效的保管和防护，但即便是在不触犯法律的情况之下，个人数据信息依然有可能被泄露，比如中间的某一个环节出了问题，可能是工作人员操作不当造成数据丢失等等原因。

#### 4.2.3 数字身份被盗用

网络技术如此发达的今天，利用网络技术盗取他人身份信息可以说变得前所未有的方便。利用各种病毒、爬虫软件攻取他人电脑；利用基站伪装客服与受害者联系；利用网上的查开房、手机定位等非法查询软件获取他人身份信息，这些报道屡屡在我们身边出现，稍有不慎，我们的身份信息就会被盗取利用，可以说是防不胜防。

网络身份盗用是身份盗用的网络形式，是利用现代信息技术及网络资源进行的身份盗用行为。从盗用身份流程和信息载体的角度来说，数字身份盗用指非法获取、持有、收集、传播、买卖、使用和伪造身份信息、文件和标志的行为。身份盗用的成立条件是“以身份欺诈为目的”，包含“在世人的身份信息还包括已故人的身份信息”，目的是为了“非法获得利益和非法逃避义务”。

公司对用户的身份信息不够重视，保护措施不严密，相关制度不完善，会给一些不法分子可趁之机，利用公司保密机制不严密、不健全来盗取公司核心机密资料，或者利用工作之便，拷贝公司重要信息，进而变卖。

不法分子还可能直接伪装成他人的身份。身份盗用者可能会利用网络电话直接伪装成受害者的父母或者亲戚朋友等，骗取受害者的身份信息，利用工具将网络电话改为亲戚朋友的电话号码，对方接到电话就显示出亲戚朋友的号码，从而放松了戒备。

身份盗取者还可能间接伪装成网站或者企业，惟肖的伪装成某种官方网站，通过电子邮件发送给受害者，诱使他们访问假冒的网站，以获取个人身份信息，实施网络钓鱼活动，钓鱼者向网络用户发送声称来自中国建设银行的电子邮件，

<sup>1</sup> 谭峻楠，廉小伟，刘晓韬. 2012 年度数据泄露调查报告摘述与分析 [J]. 保密科学技术, 2012(9).

<sup>2</sup> 黄国彬，郑琳. 大数据信息安全风险框架及应对策略研究 [J]. 图书馆学研究, 2015, (13).

给出假的超链接网址，并且该网页的外观仿制的与真正的中国建设银行网址及其相似。

还有黑客通过病毒等技术手段攻击，可以完全操控他人的终端设备，届时所有的信息都会被黑客所掌握，电脑一旦遭受到了病毒攻击，如果没有相应的防护措施，就是瞬间变为完全受黑客摆布的“肉鸡”电脑，届时电脑内的所有信息资料都会暴露在黑客的眼中。

## 第五章 大数据背景下数字身份的公平问题

由于现实生活中人们的生活、经历、知识、经济水平、社会地位和国家环境等各不相同，个人形成的数字身份也迥然相异，个人对网络资源的利用方式、对信息的理解程度、对数据的占有有多寡都等会有所不同。

尤其在大数据时代，海量数据的获取和大数据技术的使用也存在着很大的差异。对数据信息的掌控和利用能力的不同，会直接影响到不同国家、地区、企业、社会群体、个人的经济收入、知识吸收和地位等级，社会也因此产生了一种新的不公平问题。

这种不公平问题很明具有马太效应。国家越是富有，联网设备和网络服务的价格就会越便宜；越容易上网，从网上获得信息也就越多越容易，在这个信息时代，拥有越多的信息就意味着拥有更多的财富；国家越是贫穷，联网设备和网络服务的价格就会越昂贵；越难以上网，从网上获得信息也就越少月困难，相对于富裕国家的差距也就会越来越大，形成滚雪球一样的状态。这种不公平问题不仅仅是因为技术方面的问题，还涉及到知识、经济、社会等方面，其表现形式有以下四种。

### 5.1 技术不相等

这种不公平问题首先体现在信息技术上，电脑技术和网络技术的研发和应用程度在不同的人群、地区、国家和社会之间差距明显，这就造成了“穷者越穷，富者越富”的马太效应。以国际互联网为代表的新兴信息通讯技术在普及和应用方面的不平衡现象造成了信息的不对等<sup>1</sup>。由于国家、地区 and 个人的经济条件、技术条件和环境条件等存在差异，所以新兴技术的普及程度也存在着一定的差距，先进的技术成果并不能同时被所有人、所有地区和所有国家的共同享受，以至于造成国家、地区和个人间的不公平问题。

当今世界的发达国家以绝对优势垄断着信息技术产业，美国在其中遥遥领先

<sup>1</sup> 胡鞍钢，周绍杰. 新的全球贫富差距:日益扩大的“数字鸿沟”[J]. 中国社会科学, 2002, (03).

其他国家。Intel 的处理器、英伟达的显卡、微软的 Windows 都集中在美国，这些计算机领域的巨头占据着各自领域内绝大部分的市场份额，主导者全球相关领域的发展。与此相对应的是，发展中国家不得不从发达国家引进昂贵的技术和设备，占据着优势地位的发达国家很容易坐地起价，以更昂贵的价格将技术和设备卖给发展中国家，有时候为了防止技术外泄和控制市场，发达国家只会出售一些陈旧的甚至是被淘汰设备给发展中国家，花更多的钱买更差的商品，这给发展中国家造成了沉重的经济压力。所以信息基础设施总是存在着数量和质量上的欠缺。

在大数据时代的背景下，数据信息是一座金矿，信息技术是从金矿中挖掘出金子的挖掘机，挖掘机性能差距越大，所挖到的金子差距也就越大。信息技术越发达，占有的数据信息也就越多，从中能够创造的价值也就越多，信息技术越落后，占有的数据信息也就越少，从中能够创造的价值也就越少。反过来创造的价值越多，则可以投入更多来发展技术，让技术飞速发展，创造的价值越少，则对技术发展的投入就会较少，技术水平发展的会更加缓慢，形成马太效应。我国的电脑普及水平，网络普及率，信息化投入和社会信息化指数方面与发达国家之间还有较大差距，如果不改善这一现状，中国与世界的差距还有进一步拉大的可能。

技术的不平等也推动了国与国之间的不平衡，诞生了数据霸权的现象。霸权在传统意义上主要包括经济霸权、文化霸权和政治霸权，由于大数据时代的到来，各国纷纷认识到数据的重要性，所以诞生了数据霸权这一新的意识形态概念。

数据霸权是指在国家与国家之间，数据更富有的国家干涉甚至是操控数据贫穷国家的信息。数据霸权实际上大数据时代中国际霸权现象的一种新的表现形式，也包含了数据强国的主观意图。大数据和云计算的结合推进了数据霸权的加剧，只有少数几个集中于美国的大数据公司提供重要的大数据技术服务，其他国家根本没有这样的能力，这些技术和服务让美国可以十分方便的采集数据，这在客观上加剧了美国的数据霸权，数据强国在大数据时代竭尽全力想要垄断数据信息，以求实现数据霸权的意图昭然若揭。

## 5.2 知识不对等

从知识方面来看，文化程度高的人，数据信息的获取速度比文化程度低的人快，这反过来会促进文化程度增长的快慢。获取知识速度快的人文化程度增长越快，获取知识速度慢的人文化程度增长较慢，双方的知识不平等现象会越来越严重。互联网是网络时代的生产力的新代表，它的存在使得知识知识不平等加剧，互联网飞速传播的特性使对于信息掌握的差不断加大，这一现象依托传播的载体是数字技术，其传播的对象是数据信息，表现形式是存与信息资源和数据知识方面的知识不平等现象。

不同身份的人之间获取数据信息的能力快慢是有差距的，主要包括：第一，获取信息能力的差距。通常一个人如果受教育程度越高，那么他的理解能力也就越强，浏览信息数据的数量也会越多，这更加方便此人从数据信息中获取知识，了解信息；第二，知识储备量存在差距。一个人的知识储备越多，那么他对新生事物的接受、理解、学习和掌握也就越快，就会变得更加见多识广。第三，交际圈范围存在差距。学习是一项活动，知识越丰富的人，通常交际圈的范围也就越大，参与更多的组织，认识更多的人，与更多的人进行交流，那么他们的知识就增长的越快，所获得信息就越多。第四，传播媒介存在差距。科学技术知识和政府公共业务的传播主要依靠纸质媒介，高学历人员是主力受众。新闻、科学技术知识和公共事务相较于电视娱乐而言的重复率较低，从中获取知识和信息的效率也就较高，重复较多，有利于知识水平低的人对新知识的了解和掌握，反之，重复较较，就不利于知识水平低的人对新知识的了解和掌握。

即使高度发达的美国，这种不公平也依然存在，知识水平方面，从人种来说，白种人高于黄种人和黑种人；从地理位置来说，美国沿海地区要高于内陆地区；从学历程度来说，高学历肯定要强于低学历。不同人之间知识的不平等现象，让不同的人对互联网技术的学习和应用程度也存在明显差距，造成了不同群体在社会中得到的机会不平等，人与人之间的差距可能会越拉越大。

大数据要求的数据样本并非仅仅是“大”，而是要求“全面”，相应的，大数据对每项数据的准确性要求就会下降了，数据样本容量增加的同时，也会增加大数据的不准确性，因此大数据分析的结果并非绝对精确地，大数据的受众只有具有较高的知识水平才能够比知识水平较低的人更准确更快速的分辨出大数据中的真实信息和非真实信息，从而更快速更高质量的提高自己的知识水平。同时，如果受众的知识水平较低，就难以分辨出其中的真实信息和非真实信息，如果误把错误的信息当做正确的来接受，反而会让你自己出现错误的认识，这些错误的认识还可能继续传播给其他人，造成更多的错误。

大由于受众的兴趣偏爱不尽相同，如果并非自己感兴趣或者自己需要的知识，那么用户针对大数据碎片化、个性化的信息几乎没有耐心去进行浏览阅读，更别提进一步来增加自己的知识了，因为每个人身份、兴趣和偏好的各不相同，决定了对信息种类的需求也不相同，对同一信息感兴趣和不感兴趣的、有需求和没有需求的用户之间，就出现了知识的不平等现象。另一方面，大数据虽然只需要知道“是什么”而不需要知道“为什么”，但是好奇心是人类的天性使然，有些人一旦知道了“是什么”，自然而然的就想知道“为什么”，甚至不达目的誓不罢休，而有些人在知道“是什么”之后，可能因为各种各样的原因，比如怕麻烦，懒惰，不感兴趣等等原因放弃追寻“为什么”，这样在继续深究“为什么”的人与不愿意



继续发掘“为什么”的人之间，就会出现知识不平等现象。

### 5.3 贫富不均等

归根到底，发展中国家和发达国家间之间巨大的经济差距，造就了发展中国家和发达国家间数据信息掌控和使用程度与网络技术发展的差距，是贫富差距现象和经济发展不平等现象在大数据时代的发展和延续。

无论是数据信息占有量的不平等、信息技术的不平等还是知识的不平等，都是从根本上反映了信息时代国际国内经济不平等现象和贫富不均等现象，是信息时代经济发展进程中出现的经济差距。不同的个体、企业、地区和国家之间存在所处的经济层面和互联网环境不同，对于新兴网络技术的获取机会、对数据信息的占有量和受教育的机会等也不相同<sup>1</sup>。

不仅仅是国与国之间，同一国家内的发达地区与欠发达地区之间也存在着这种贫富不均等现象，主要是指城镇与乡村之间，存在着较为明显的贫富不均等，同时大数据造成的地区之间技术不平等和知识不平等也在不断加剧地区之间的贫富不均等现象。在城镇与乡村之间，不管是信息技术意识，还是信息技术接入，亦或是信息技术利用和信息技术环境存在着较大的不平等现象，并且有进一步持续扩大的趋势，这主要是因为城乡之间教育文化、信息基础设施、信息技术利用、经济四个方面都存在着差距，技术间的不平等也会反过来造成贫富不均等现象。

人与人之间同样存在着贫富不均等现象，由于人与人的经济水平不同，在市场经济的大环境下，经济水平越高的人一般社会地位也会越高，社会上较为重要的消息，比如有关政治决策、经济环境等知识总是先被社会地位较高、积极实力较强的人所得到，然后再传向地位较低，经济实力较弱的人，因为经济实力较强的人一般获取重要信息的手段和途径要比经济实力弱的人要多得多，同时，经济水平较高者的交际圈范围相对更广，更有可能包含了解政治经济信息的人，因此经济实力较强的人能够借此获得更多的考虑时间来决定自己的对策，从而更好的展开行动，让自己的决策和行动更加合理和优秀，从而不断增强自己的经济水平，加剧贫富不均等现象。

大数据带来了新的经济价值生产方式，也影响到了经济价值的分配方式。这种新的分配方式会给部分群体带来客观的利益，从而不断加大现有的经济差距。那些有能力使用大数据技术的组织将会加剧这种不平等性。大数据仅在中国每年市值就超过了百亿元，大数据是个庞大的产业。在前期必须对大数据技术注入巨量的资源，不仅仅是金钱，还包含海量的数据资源和现今的数据复合挖掘技术，没有大量的资本根本无法完成这样巨大的工程。只有那些拥有强大经济实力的企

---

<sup>1</sup> 薛伟贤,刘骏.数字鸿沟的本质解析[J].情报理论与实践,2010,33(12):41-46.

业在这种情况下才能发展大数据技术，积极进行技术创新、了解客户需求，形成一种越来越大的“滚雪球”效应加剧贫富不均等现象。

#### 5.4 社会不平等

是信息社会中，数据占有量多的阶层与数据占有量少的阶层之间的存在着差距，这是一种信息分化现象，也是传统社会分化现象在信息时代的发展和延续。此时的社会不平等现象是信息社会的阶层分化和社会分化现象的的投射，归根结底是由信息网络社会不平等所导致的。

电脑和网络技术的普及程度的不同会让生活在在社会中人们的机遇也不相同，这极有可能导致人与人之间社会地位的差距。有条件利用计算机和互联网的人可以在网上寻找招聘信息，更容易找到好工作，也可以利用网络张贴交易信息，让自己的商品更容易卖出去，也可以在网上进行公司的管理和参与公共事务等，相较于没有条件或者不会利用计算机和互联网的人，更容易让自己处于优势状态，从而提高自己的社会地位。

大数据时代数据信息已经成为一种商品，能够上网并懂得利用数据的人能用数据信息产生价值，在网上获取更好的交流和获取信息的机会，而相对贫困的人没有条件利用互联网，也就失去了在网上沟通交流的机会，失去了在网上了解信息获取数据的机会，这就使得社会阶层开始出现分化，造成了社会的不平等现象。

在大数据时代，种族歧视问题也很可能被大数据技术所放大，有些犯过错误的人可能根本没有改过自新的机会。Google 提供给一些企业进行定制数据库业务，主要是企业用于招聘员工之用。在这些数据库中，黑种人相较于白种人被列出了更多更为详细的信息数据，不但包含了这些黑种人的年龄和身高等信息，甚至连婚姻历史、宗教信仰、性欲取向和药物使用等信息都被包含在内，这就让这些企业更为谨慎的对待这些黑人受聘者。更加过分的是，数据库中竟然还“贴心”的收集到了法院和警察局等公布的有关这些黑种人的信息，一旦检测到某个黑人曾经有过违法行为，就会“智能”对企业进行提示：这个人过去曾经有过违法行为，雇佣时需要谨慎。大数据技术让这些曾经有过违法记录想改过自新的人几乎很难找到正当的就业机会。

## 第六章 大数据背景下数字身份的认同问题

数字身份也是人们在网络空间建立起来的一个在线身份，由于网络的匿名性和多元性，人们可以拥有很多个不同的数字身份，这些数字甚至可以和现实中的身份毫无关联，但是这些建立起来的数字身份，同样也和现实中的身份一样，存在认同的问题。互联网就像一面镜子，每个人把自我折射在其中，就形成了数字

身份,从这个角度来说,这种数字身份就是网络中的“另一个自我”。身份认同就是一种对自我身份的确认,通过区分自我与社会间的差异性,从而对自我身份的与众不同获得识别与认可,换句话说,也就是个人知道自身在社会中处于怎样的地位或者扮演怎样的角色,并且个人的这种角色可以得到某些人或者某些群体的认可和接受。

## 6.1 数字身份认同的概念

身份认同(Identity)问题是西方文化中一个及其重要概念,新左派、女权主义以及后殖民主义都对这一问题青睐有加。身份认同的含义是对个人以及特定社会文化的认同。英文的身份认同是 Identity 与身份是同一个单词,牛津词典将 Identity 解释为两种含义,第一种含义是身份,本身,本体,其中蕴含着对“我是谁”的追问,第二种含义是同一性,相同,一致性,包含着本体对自我与他人或者种群“一致性”的追求。

从主客观来看,身份认同分为主体对自身的认同和客体对自身的认同,这里的客体指的当然是由他人所构成的社会,也即是自我认同和社会认同。自我认同是根据自身经历来不断反思对自我的一种理解,是一种内在性的认同,是对“我是谁”的追问,社会认同是人处在一定的社会环境之下,发自内心的贴近该环境内的价值观和思想文化的一种状态,是对自己所属集体的共同性和与其他集体之间的差异性的认知,追求的是“我们是谁”的问题。

著名的英国社会学家安东尼·吉登斯(Anthony Giddens)认为身份的认同总是被建构的。“它是个人依据其个人经历所形成的,作为反思性理解的自我。虽然认同也可以由社会规范产生,但只有社会行动者将规范内化,且围绕着内化过程建构意义时,它才会成为认同。”

因此身份认同如何被构建也有“自我建构”和“社会建构”两种观点。

弗洛伊德是身份由“自我构建”的代表,他提出把“认同”作为一个“心理过程”,是个人通过模仿另外一个人或者团体所具有的价值、规范和面貌,并内化为自己的行为模式的一种过程。

泰弗尔(Tajfel)则提出了与之相对的“社会身份认同”的概念,即“个体认识到自己所在群体的成员所具备的资格,以及这种资格在价值上和情感上的重要性”<sup>1</sup>。一个属于社会群体一员的人的身份,与其所属的群体是什么类别,二者在自我概念中具有极其重要的地位。泰弗尔还呼吁人们通过对社会进行分类、比较和积极的区分来获得具有正面积意义的社会认同,这样可以获得并提升自信和自尊。

<sup>1</sup> Tafel H, Turner J C. The social identity theory of intergroup behavior. In: Worchel S, Austin W (eds). Psychology of Inter group Relations. Chicago: Nelson Hall, 1986. 7- 24.

信息网络时代的到来让身份认同发生了巨变。网络媒体相较于传统媒体，身份信息的流传速度、深度和广度都大大提高，值得注意的是，不同文化领域，包含主流文化、亚文化与多文化之间的接触和互动更为密切，因此网络时代的数字身份认同变得更为差异化和复杂化，不同身份背景和文化背景的冲突与融合特性尤变得为突出。

雪莉·特克（Sherry Turkle）<sup>1</sup>认为，人们业已经把网络变成不断自我构建和在构建的一种社会化的实验室，人们在虚拟的网络世界中不但可以对原有自我进行塑造，还可以创造新的自我。在虚拟的网络上，用不同以往的方式去看待身份认同问题已经成为了人们的新习惯，身份认同的意义绝不像是传统含义一样稳定了，所有人都能够不断的构筑自己的身份和创造新的身份，身份的含义不再是一成不变的了。

在大数据时代，数字身份变得容易被追溯，这就让人们在网络上交流和行动上产生了顾忌，担心会暴露自己的数字身份，一定程度抑制了网络的自由性，但是同样的，人们的随意性也被降低了，这就减少了很多不当言论和图文的出现，也对网络环境的健康起到了积极的作用。大数据同时也让人们变得极为重视数据，这可能会造成人的数据化，坚持“唯数据论”，进一步割裂数字身份与现实身份，但同时大数据也依托于个人的喜好不同，带来了更多更便捷的个性化服务，使得人们生活的更加舒适。

大数据对数据信息的二次挖掘，能够找出用户潜在的兴趣和偏好，提供完善的个性化服务，通过推送或者打广告的形式，将这种服务完全推广到用户手中，使人能够更加舒适方便的完善自己的数字身份。

从另一方面来说，认为大数据将我们的思想与见解数据化，利用数据分析将其还原为我们的需求与偏好，这同样使得我们可能会乐观的认为我们的心理、需求、爱好、生活习惯等方式都可以被大数据所推测出来，让人产生一种数据万能的错觉，最终沦为数据的奴隶，变成“唯数据主义”观点的拥护者。

## 6.2 数字身份与现实身份相脱离

因为人们在网络和现实的感觉并不同，因此数字身份身份虽然也是自我认同的一部分，但却与现实中的身份相区别，形成了一个已具有独立个性和特征的独立人格，这种独立人格与现实生活中的人格并行不悖，同样属于自我认同的一部分。

网络所具有的技术特征，比如匿名性和去中心化等，人们在网络中与现实相脱离的程度比在现实社会中与现实相脱离的更为容易，也更为完全。人们原有的

---

<sup>1</sup> Turkle, S.(1995).Life on the screen: Identity in the age of the internet. New York: Touchstone, p184.

社会角色、身份和外貌、身高甚至性别等物理特征都可以轻易地被隐藏，人们完全可以脱离原有的阶层和位置，脱离原有的群落，包括邻居、职业，甚至是民族和国家等过去的认同方式。由于互联网的零碎化的交往特性，人们在不同的网络社区、环境和位置完全可以展现出不同的自我特征和认同偏好，多元化的自我的多样性的认同早已是现今网络所具有的普遍特性，网络中数字身份与现实相脱离和身份认同偏好的多样化比其在现实中体现的更为显著。

人们现实生活中的种种不顺会让造成人们的失落感觉和挫败感，很多人自然而然的会去到网络中去寻找成功与慰藉，通过电子游戏和网络社区等方式来使得自己暂时逃离现实的失败。很多“网络沉迷者”花费在网络中的时间和精力是超出正常时间的，他们将大量的时间和精力花费在从事与正常生活和工作无关的虚拟世界活动，比如电子游戏，网络社交和网上购物等行为上。这些人不断地逃离现实，并在网络中建立虚拟的自我身份，来获得他人的认同，比如通过在网络游戏的高成绩与高排名引发他人的称赞与羡慕，通过不断与他在网络社区中聊天与发帖来获得他人认同从而提高自己的等级与威望等。这些行为适量进行是有积极作用的，而一旦超过了应有的度量，就会让很多人越发的沉迷其中不能自拔，甚至认为现实中自己失败的人生毫无意义，造成网络数字身份与现实身份互相对立冲突。

更可怕的是一个人对于网络上虚拟的数字身份越是认同，他就越会花费更多的时间和精力去维持或者提高这种认同，那么他分配给现实身份的精力和时间就会越来越少，结果会导致对自己的现实身份愈来愈不认同，造成虚拟的网络数字身份与现实身份脱离的越来越严重。

有学者认为自我对数字身份的认同是通过建构数字身份本身而实现的，自我构建就是通过网络以各种方式对自我连贯性的叙述，包括建立自己的主页和在社区与他人互动等方式。虽然网络中数字身份的建立具有很大的随意性，充分体现了主体的主观能动性，但是我们应该清醒的认识到，寄托于网络中的数字身份并不能独立于现实中的真实身份而存在，过于夸大自我对网络身份的主体性会让人们只注重个人的存在，忽略社会对数字身份认同的影响力，网络主体构建数字身份的确具有很大的自由性和多重性，但这种数字身份的构建依然受背后存在着的现实世界中的真实身份的极大影响，真实身份可以说是数字身份的现实基础。

### 6.3 数字身份认同的价值观与社会认同的主流价值观相背离

价值观是个人或者群体从思想上对周边事物进行的一种认知、认定、或者判断等的总体看法和总体评价。价值观不仅引导着人们的价值取向与追求，也是一种价值尺度和衡量标准，人们据此判断事物价值的有无和大小。不同的个人、不同的阶层、不同的社会具有不同的价值观，虽然现今价值观不断变得多元化，但

不论何时，总有一些价值观点成为主流选项，成为整个社会和国家所崇尚的精神，当今中国社会的主流价值观是以社会主义核心价值观为核心内容的主流价值观。

主流的价值观包括了一股向善的意志力，它要求公民群体变得更加爱国敬业和诚信友善，要求国家变得更加富强民主和文明和谐，要求社会变得更加自由平等和公正法治。其内容包括政治、经济、法律、道德和技术等各个方面，要求它们必须全面积极向上发展，这些美好的希望通过互联网技术形成了一股巨大的力量，能够改变民族，改变国家，改变社会。

在网络中，社会对数字身份的认同实际上是个体对社区的一种归属感，是认同自己的群体，并随之产生对群体内部的偏好和对全体外部的偏见，获得个体自我身份的确认，通过自身群体与他人群体的差异排斥他人群体，获得全体也就是社会自我身份的认同。

一方面互联网和多媒体技术为多元化的文化传播提供了宽广的平台，人们通过网络获得各种影像与音乐，通过网络足不出户就可以领略世界各地丰富多彩的民族文化，在这个时代真正做到了“秀才不出门，尽知天下事”，同时人们借助互联网的交互能力，能够随心所欲的在网络中进行创作和交流，比如撰写 QQ 日志、发表微信朋友圈和分享微博好文章等；但是另一方面，互联网的开放性和自由性也为色情、暴力等丑恶的反面价值观大开了方便之门，人性中的贪婪和暴力被互联网的虚拟性和匿名性披露的淋漓尽致，人们对于正确价值观的分辨能力变得极为困难。淫秽、情色、暴力等低俗行为在当今已经成为了社会公敌，再加上外国敌对势力的经济、技术和文化的霸权在通过网络不断影响人们价值观的正常判断，比如近期闹得沸沸扬扬的“扮鬼子事件”，两名精日分子竟然在南京抗日碉堡前身穿二战日本军服进行拍照，亵渎了民族精神、忘记了民族伤痕，严重违背了主流价值观。

网络中的每个个体通过积极的社会认同来得到自我确证和提高自尊，自尊也分为积极和消极的两种，在群体内部与全体外部进行正面的比较能够提升积极的自尊，但通过贬低其他群体来抬高自身群体则只能够提升消极的自尊，并不符合主流的价值观。个体对自身所处的群体痴迷，认为自身群体强于他人群体，在寻找与其他群体的差异化的过程中，很容易会引起群体之间的成见与冲突。比如与同一个 QQ 群内的人拥有相同的兴趣爱好，互相分享交流，使得自己更加融入这群体，或者通过参与自己贴吧与他人贴吧之间的骂帖和攻击，来支持自己的群体，打压他人的群体等。

无论如何，网络中数字身份的主体是活生生的人，而人毕竟是一个活着的，能够感觉、认识、思考和反思的社会性动物，所以网络终究不可能脱离现实社会而存在，因此网络中数字身份的认同最终还是要回归到社会交往关系之中。而价

价值观一般具有一定的稳定性和持久性，在一定时期的环境当中，我们的价值观一般也是稳定和持久的，很难突然就发生改变，因此在我们从网络中的数字身份回归现实世界的社会身份时，网络中所形成的不良价值观是一时难以摒弃的，这些消极的价值观无疑是与主流价值观相背离的。

#### 6.4 数字身份在网络中的地位与现实身份在社会中地位有差距

地位通常由该社会的道德、法律和习俗等规定的，某个或者某些社会成员在该社会系统中所处的位置，地位的高与低一般表示在社会中取得的威望的高与低，财产、权利等的多寡等。

社会地位一般受家庭出身、受教育程度、职业能力、权利大小、收入多寡等影响，但是高收入、高学历、高能力等这些现实中能够影响个人地位的因素都被网络技术给最大化的稀释了，人们塑造的数字身份的地位与现实社会地位并不一定划上等号，甚至差距很大。现实中拥有较高社会地位的个体在网络中可能无法获得相应的尊敬与待遇，不仅人人羡慕的铁饭碗公务员可能在社区里被很多人鄙视为靠爹的关系户，坐拥过亿身家的大老板也可能在群聊中被所有人骂成为富不仁的大奸商，现实中人见人爱花见花开的大美女也可能被喷为样貌全靠整容照片全靠美颜的骗子。反之，在网络中拥有较高身份地位的人在现实中也并不一定很受人待见，比如电子竞技中知名战队的著名选手就被很多人当成是玩游戏不务正业的，网络直播的当红播主也被很多人认为是卖肉博眼球的道德败坏者等。

很多人在网络上成为了知名人士是因为他们犯了一些并不太大的错误，但是这些错误被爆聊出来之后，由于互联网极速的传播速度与广度，这些人的糗事很快都会被很多人知道，其中不乏很多对其进行辱骂、人身攻击甚至是进行人肉搜索的人。比如李开复的学历造假事件，江西“史上最毒后妈”事件等。

李玫瑾教授因为从犯罪心理学角度分析了药家鑫的行为，尝试性的为被告辩护了几句，就遭到了铺天盖地的谩骂和诽谤，成为了被公众口水所淹没的专家。随后她的资料就全部被曝光出来，甚至其学历也遭受了莫须有的质疑，她随后发表的每天一条微博，都会在网群之中引发激烈的讨论。

网络地位的差距甚至还与性别有关，性别也是互联网中虚拟要素性别最受人关注的对象之一。能够在网上一对一聊天时忍住不询问对方性别的人机会是不存在的，对方在询问之后的回答极有可能是文字形式的，有时候就算是发送照片也很可能是从网上盗来的图，无法加以检验。尤其是在网络游戏中，很多假扮成女生进行游戏的男生对得到一大堆男性玩家的喜爱和关照，但是一旦他开诚布公的公布自己是个男生，那么那些以前照顾他的男性玩家大多会直接离开，甚至是对其进行谩骂和侮辱，认为对方欺骗了自己。

数字身份认同的地位差距实际上就是一种歧视，据研究说明，在受到歧视的

时候，人类大脑对痛苦的感觉与人体被打时感受到的痛苦是一致的，歧视在现实生活中就是一种暴力的直接体现，遭受歧视与遭受物理暴力相比固然没有留下明显的外伤痕迹，但对心理的创伤比外伤更加难以愈合。

## 第七章. 关于大数据背景下数字身份伦理问题的原因的探讨

虽然大数据强调相关性，但这并不意味着大数据就不关心因果性。恰恰相反，大数据最重视的就是因果性，因为大数据的全部理论都建立在因果性之上，如果忽视因果性，则意味着大数据的理论失去了科学依据，将变为无根之萍，无源之水，因此，因果性是大数据理论的基石。有果必有因，大数据背景下数字身份带来的各种各样的伦理问题必然有其原因，其根源或许是复杂的，难以捉摸的，难以说清道明的，但是无论如何，数字身份问题的根源一定存在。弄清数字身份伦理问题产生的根源，我们才好看准目标，对症下药。大数据背景下的数字身份伦理问题的成因可以分为主体根源、客观原因和社会背景三个方面。

### 7.1 主体根源：主体的复杂性

人类作为网络世界的主体，在大数据的时代背景下不断地被数据化、标签化和碎片化，被虚拟化的人格正在不断异变。像所有的技术一样，大数据技术也是中立的。大数据并不包含有内置的视角预先判断什么是对，什么是错。同样的，数字身份本身也并没有对错之分。虽然大数据技术和数字身份本身没有价值框架，但是，使用大数据技术和利用数字身份的人却拥有自己的价值系统，所以才会引发一系列的伦理问题。

人类作为技术的主体是复杂的，一方面希望利用技术改善自己的生活条件，让自己生活的更加舒适，另一方面却难以把控技术的尺度，技术遭到滥用威胁人类生存的事情屡见不鲜，比如利用石油煤炭等作为燃料驱动机器改善生活，却并不明白节制，直到造成了温室效应、臭氧层空洞和大气污染等全球问题，才明白节制使用的意义；塑料袋的出现大大方便了人们日常购物出行等行为，但是直到塑料垃圾变为“白色污染”问题，才明随意丢弃塑料制品对环境的侵害；各种工业化学制品的出现大大方便了企业的生产，但是随之而来的是工业生产中的废水废气的乱排乱放，直至引发大气污染和水资源枯竭等问题才明白应该防止污染。人类的本质是自私利己的，在当今大数据时代的背景下，网络技术和大数据技术放大了主体的利己性，更加暴露了人类主体的复杂性。

#### 7.1.1 网络主体被数字化遮盖了主体的品性

在互联网时代，人们正在上网的同时，已经被数据化了，身高、体重甚至爱



好都被转化成一堆堆的数据，把这些数据整合在一起就构成了我们自身的形象。在网络上，我们认识新朋友也是首先要通过数据，比如 QQ 微信中所设置的个人资料；企业翻看我们投递的电子简历；办理信用卡时，客户经理查询我们是不是黑名单等。这些数据化的信息对我们了解一个人的确有所帮助，但冰冷数据堆砌起来也只是一堆数据，而不能构建一个有血有肉的人，人一个人重要的性格、品质和人际关系等都没有办法通过数据再现出来，这些重要的信息被我们无形中忽略了，因此我们完全通过数据来了解一个人很容易被这些数据所“误导”，对这个人来时其实并不公平，这严重影响人的本质属性，丧失人的主体地位。有学者将主体数据化也视为一种“技术的异化”<sup>1</sup>。

从互联网的普及开始，人类就慢慢被数据化了，而大数据的出现让这一进程大大加快，现如今在大数据技术的作用之下，人类的一切都可以被量化了，大数据让人们以前小数据时代难以想象的海量数据进行采样、存储、共享和分析成为可能，数据种类的类型也几乎涉及所有领域的结构类型，包括文、图、音、像、表等多源、多元数据，一切可以想到的几乎都是可以量化的，可以说是“只有想不到，没有做不到”。网络主体也被大数据量化成了一堆堆的数据信息，比如说人类的性格，可以简单量化成活泼、文静、内向和外向等，但问题是人类性格往往是复杂多变的，一个人可能在某些时候活泼，在某些时候又显得文静，在面对陌生人时面无表情，在面对好朋友时却无话不谈，一个人的性格是难以用数据或者一段信息完全涵盖的，因此仅凭数据信息的只言片语来了解一个人的性格往往是不准确的、片面的，这会让自己的认识产生偏差。同时如果长期通过量化数据来了解他人，很有可能让我们认为仅仅通过量化数据就可以完全了解他人，通过与他人实际相处来认识一个人很可能是没有必要的，这种想法会严重扭曲我们的交友方式，让我们与他人的交往能力变得退化。最后，大数据带来的不仅仅是将过去和现在的数据进行量化，甚至也可以将预测到的未来数据量化表示出来，就跟电影《少数派报告》中的未来很像，只不过把三个“预言家”换成了大数据技术，通过对个人偏好、生活习惯和历史行为等的分析，判断出一个人以后很可能会进行犯罪行为，然后将其逮捕，这无疑是荒谬的，对他人是极端不公正的。

### 7.1.2 企业利用大数据逐利无视他人利益

大数据利益相关者分为大数据生产者，大数据搜集者，大数据使用者三类。大数据生产者是指产生大数据的人群，一般的普通人民群众无疑是其中最主要的代表；大数据搜集者是指采集大数据的群体，大数据使用者是指分析、挖掘和利用大数据的群体，这二者一般是指大数据企业和政府。比如，相对于大数据企业

---

<sup>1</sup> 薛孚,陈红兵.大数据隐私伦理问题探究[J].自然辩证法研究,2015,31(02):44-48.

和政府，我们这些普通人毫无疑问是弱势的一方，我们的隐私数据，个人信息经常被大数据企业所挖掘和利用，大数据企业为了不断追求自身的利润，钻法律的空子，无视普通人的数据权力，这些现象实际上非常普遍。所以大数据企业利用大数据逐利无视他人利益也是导致大数据背景下数字身份伦理问题的原因之一。

在使用大数据技术的过程之中，不同的用户、不同的组织、不同企业之间的着眼点不同，往往都是根据自身的利益为基础，追求自身利益的最大化，无视或者不重视他人利益，这就极有可能会侵犯他人的权益<sup>1</sup>。

大数据搜集者的真正目的并非搜集的大数据本身，而是为了利用大数据技术从海量数据中挖掘出巨大价值，其目的是逐利性的，因此大数据搜集者在搜集数据信息是带有明确目的的行为，所以在搜集的过程中对数据信息的筛选也是必须的，合目的数据就留下，不合目的的数据就抛弃，因此导致了其搜集数据如果用于一开始制定的目标来说的确是全面的完整的，但对于其他目标来说就不一定全面而完整了，因为大数据的价值之一就在于对数据的二次或者多次利用，这是大数据使用者要做的事情，所以大数据搜集者搜集的数据只能保证与特定的目的相关，却不能保证与其他目的相关。大数据使用者是利用大数据技术对已搜集的数据进行分析、挖掘和利用，如果在数据挖掘中正好实现了当初搜集数据目标的价值，那自然皆大欢喜，但是，如果在对数据二次挖掘或者多次挖掘中，产生了新的价值，超出了搜集当初的既定目标，那这些价值该如何分配呢，是大数据使用者独享还是要将其分给大数据搜集者一部分呢？这个问题就给数据的二次利用蒙上了一层阴影，也造成了大数据使用者与大数据搜集者之间的矛盾。

大数据生产者相对于掌握大数据技术的大数据搜集者和大数据使用者明显处于被动的弱势地位，大数据搜集者和大数据使用者在搜集和使用的过程中可能根本就没有告诉大数据生产者，大数据生产者也根本不知道自己的数据信息被收集，被用于数据分析、二次挖掘和预测未来，更加不可能预料这种情况会给自己带来怎样的影响，即使这种情况给大数据生产者带来了一定的影响，但是大数据生产者也会因为影响并不大或者难以维权放弃对大数据搜集者和大数据使用者的追究，放任大数据搜集者和大数据使用者为所欲为。

### 7.1.3 主体隐私观的缺失

大数据已经极大地改变了我们的生活，各种社交软件可以让我们随时随地的分析我们的喜怒哀乐，比如在微博、QQ 和微信朋友圈上传自己的照片、音乐和视频，在斗鱼等直播平台与粉丝和主播互动等等，殊不知这些行为，如果被大数据技术所利用，就有可能暴露我们的位置、爱好、消费习惯等隐私信息和身份信息，

---

<sup>1</sup> 唐凯麟，李诗悦. 大数据隐私伦理问题研究 [J]. 伦理学研究，2016，(06).

甚至有些大数据企业能够以此推测出我们未来的行为。而做出这些不安全的举措，完全都是因为我们没有认识到某些行为会严重影响我们的隐私安全，即我们的隐私观缺失。

由于社会中某些个体隐私观模糊，并未意识到隐私数据被共享，或者身份信息被泄露。很多人喜欢在朋友圈或者 facebook 等社交平台公开自己的照片视频等信息，但是并不清楚这些行为可能会给自己的隐私和身份信息安全造成影响。这些行为归根到底都是对隐私的认识不够，隐私保护意识不强，大数据时代的隐私虽然容易丧失，但我们却不能像某些人提出的“后隐私运动”一样放弃隐私，“后隐私运动”将隐私视为控制信息被传播共享的一种方法和手段，在如今的大数据时代，由于隐私已经难以被保护，所以我们应该主动放弃隐私。但是隐私是文明人的一项基本权利，放弃隐私必将给一个人的权益带来巨大的损害，因此不但不能因噎废食放弃隐私，还必须得加强手段保护隐私，防止隐私泄露，因此补足完善个人隐私观，是从源头防止隐私的泄露的方式之一。

#### 7.1.4 主体自身的道德素的缺失

我们数字身份信息的泄露也可能是由于某些人的低素质行为，比如恶意的肉肉搜索等。这些人可能出于嫉妒心理甚至是纯粹为了好玩等非常无聊的个人原因，毫无顾忌的将别人的隐私信息公布，这严重的侵害了被曝光者的数据权力，这种经常出现的恶劣行为完全根植于某些人自身道德素质过低。

2013 年 4 月，家住汉口的常小姐不断被陌生人通过电话询问她是不是租房子，并且她发现陌生者对自己的全名等一清二楚，通过搜索引擎的搜索，常小姐发现自己的姓名和手机号码等信息被公布在两家租房网站上，网站上均显示自己有一套位于广州市区的房租需要出租，至今常小姐依然会在一天内接到十几个陌生电话询问租房信息，他迫于无奈不得不将手机关机。常小姐自己推断可能是因为她淘宝网买了一件衣服，拍下并付款之后，卖家不知为何一直没有发货，自己多次通过旺旺与卖家沟通，卖家却迟迟不回，常小姐无奈申请退款，并因为不满意卖家态度向淘宝客服进行了投诉，卖家恶意打击报复造成的，她联系到卖家，卖家却拒不承认。如今的网络技术如此发达，给好评还是差评是用户在网上购物的权利之一，像这种由于卖家被用户差评，故意报复买家的而故意泄露他人身份信息的例子还有很多，这归根结底是由于卖家道德素质缺失。

传统的伦理道德已结无法适应当今大数据作用下，数字身份出现的一系列新问题，并且应用主体们不知道如何合理利用大数据分析数字身份信息所得的知识和结果，容易出现行为失范的现象，所以大数据应用主体自身的道德素质的缺失往往会成为导致伦理失范的主要因素之一。

#### 7.2 客体原因：技术的负面效应

技术发展的不成熟，带来了很多负面效应。大数据技术在给人们带来网络便利的同时，也给数字身份信息的监管提出了难题。政府和企业会更深入的进入我们的生活领域，对我们的隐私安全，数据权利等方面发生不同程度的影响。

### 7.2.1 大数据技术尚不成熟

自从2008年以来，大数据爆发式增长，至今不过只有七八年时间，距离技术成熟的阶段依然十分遥远，认为大数据可以成为科学发现的范式的这一美好憧憬仍然处于理想阶段，人们在大数据发展的道路上仍然有很长的路要走。而人类历史上每每出现一项新的革命性技术，其在发展初期总会出现这样或者那样的负面效应，而其中的伦理问题往往是最直接最突出的一项，比如克隆技术，转基因技术等等。

每一历史时期的科学技术都只是人类认知和实践发展到特定时期的必然产物，任何科学技术的发展都是从不成熟走向成熟的过程，人类利用这些科技的初期根本无法全面预料到这项科技所带来的一系列影响甚至效应，跟无法预知此项科技与其他事物之间盘根错节的相互联系<sup>1</sup>，大数据技术亦是如此。

首先是，大数据技术诞生至今不过短短几个年头，其成熟程度、实用性和访问性方面都还难以赶超传统数据库及其管理套件，即使很多大数据技术都是由大量自由的志愿者进行开发、测试和修改的开源技术，也难以抹平与发展多年业已成熟的传统数据库技术之间的时间差距。然后是大数据技术相较传统数据库学习较为困难，人们多年使用关系型数据库系统已经形成一种惯性，短时间内难以适应完全不同大数据技术的数据存储和处理方式。最后在管理层面上，大数据管理团队的技术与知识完全来自于对过去的小数据时代的关系型数据库的开发与运维，经验虽然丰富但却不一定适用于新兴的扁平化和非结构化的海量数据类型，所以大数据技术尚处于新兴技术阶段，其技术尚未成熟。

### 7.2.2 互联网的固有特性

大数据技术的基础是网络，互联网自身固有的特性让使用者获得了匿名性和自由性，人们容易把在社会生活学习中的不满带入网络之中，在网络上宣泄自己的不满和释放自己的压力，完全不用担心自己在现实生活中的问题，私密聊天、社区交友等应用如雨后春笋般纷纷冒出头来，人们通过这些软件来释放压力、寻找慰藉，这在客观上使得加剧了数字身份伦理问题的发展。

在现实社会中，人们往往承载着来自各方的压力但却无从发泄……而在虚拟的社交网络中，人们则可以通过微博、贴吧、论坛和社区来寻找“志同道合”

---

<sup>1</sup> 安宝洋，翁建定. 大数据时代网络信息的伦理缺失及应对策略[J]. 自然辩证法研究, 2015, (12).

之人，在这些广阔而自由的倾诉空间里，他们肆意发泄着情绪来聊以慰藉<sup>1</sup>。

互联网给了我们自由地发表言论、表达自己心情的新平台，从传统的单一线性传播模式变为双向的反馈互受模式，带来了媒体传播的以此飞跃性的提升。虽然网络崇尚自由，但是不等于网络上就不必讲规矩。如同每个网吧都有网管一样，每个网站都有其负责的网编，每个贴吧都有负责的吧主，每个论坛也都有版主，这些编辑和管理人员的作用就是负责信息在互联网传播的管理者和监控人的作用，这些管理者有权利且有义务去处理不当的网络行为和言论，但是中国近七亿网民中的管理人员却寥寥无几，即使专门有人去监管，但是网络的开放性的自由性让这些管理者很难面面俱到。

### 7.2.3 大数据技术自身的独特性

大数据由于技术方面的硬性原因，大数据技术往往只是掌握在少数企业和政府手中，能够利用大数据技术的也只有他们，我们普通人就算手里握有海量数据也没有办法自己使用，大数据领域由少数企业和政府所垄断，普通民众对大数据技术只有被动的接受能力，无法在大数据技中转为主动模式，这就可能会造成少数企业对普通民众进行的数据独裁。

还有人们利用大数据技术不断发现惊人的商业价值，人们对大数据越来越信任，越来越依赖，导致人们跌下“唯数据主义”的万丈悬崖，对数据的盲目信任，很容易让人们忽略事物的其他属性和本质，引发一连串连锁问题就不足为奇了。如果我们的任何判断、决策都必须完全依赖于数据，就有可能走向唯数据主义，必然导致大数据技术的负面影响扩大化而使大数据技术异化极端化发展<sup>2</sup>。

贵为谷歌公司创始人的拉里·佩奇和谢尔盖·布林一直认为一个人的 SAT 成绩能彰显潜能，大学时期的绩点能够展现成就，故此他们一直要求每一个来谷歌的应聘者携带这两项证明，但可笑的是已经入行十多年并且取得过令人瞩目的成绩的职业经理人在被要求这些时就完全表示不理解谷歌这种顽固不化的行为。

一旦下级向上级汇报时隐瞒上级不喜爱的数据，只汇报上级喜好的数据，报喜不报忧，则很可能造成管理层的决策失误。詹姆斯·斯科特是一位人类学家，他在《国家的视角》中详细的描述了一种政府的状态，不管民众的生活状态，完全只靠图纸来规划重建社区；不懂农业生产，完全依靠农产数据来决定农场的生产方式，鲜明的反映出了如果对数据过于崇拜和以来会造成严重的恶果。

### 7.2.4 数字身份的多元性

数字身份的特性让几乎所有人在网络上都可以拥有不止一个数字身份，这些

<sup>1</sup> 宋振超，黄洁。 大数据背景下网络信息的伦理失范、原因及对策 [J]. 理论与改革，2015，(02).

<sup>2</sup> 陈仕伟。 大数据技术异化的伦理治理 [J]. 自然辩证法研究，2016，(01)

数字身份可以完全不同，各行其是。数字身份越多，主体将之整合起来就越困难，这就使得现实身份与数字身份变得更加分离。从现实身份与数字身份之间是否契合的角度来看，二者标准一致时会给主体带来积极的作用，不一致时，会给主体带来消极影响，“当现实和虚拟世界的标准一致并能带来积极体验时，个体的身份认同整合处于一种和谐的状态，但当现实和虚拟的标准发生冲突并伴随消极体验时，往往带来整合上的困难”。

数字身份的多元性是人类性格的复杂性被网络空间所放大的结果，数字身份构建的过程中并不要求与本人现实相同，不同或者完全不同都是完全可以的，有的人在现实身份是一个女强人，是企业的美女老总，总是不苟言笑，但在网络的某一社区上，可能完全是以一个小女人的身份出现，关心的不是国家大事，经济政治，而是喜欢关注家长里短、娱乐八卦等话题，在另外的某一个社区却是以另外一个中年男人的身份出现，语气严肃但却关心他人，自负看透世情，对谁都怀着三分算计五分防备，让人敬佩之余难以亲近，任谁也不会想到他是一个女性。然而数字身份越多、越复杂、越与现实相背离就越难以与现实身份融合在一起，这些数字身份虽然无法突破网络的桎梏直接显示在生活中的真实身份中，但却可以间接影响个人的思想和行为，让人们难以分清的真实世界的社会身份与虚拟世界中数字身份的界限，造成个人与现实的脱节。

### 7.3 社会背景：外部规约的不足

大数据时代是互联网时代发展的一定阶段的结果，大数据技术的诞生有着复杂的社会背景，所以探讨大数据技术产生的社会背景也要从宏观和全面的角度来进行切入，一般可以分为法律机制不健全、行业内缺乏统一的伦理规范和监督机制不健全三个方面。

#### 7.3.1 法律机制不健全

大数据在我国起步尚早，现有的法律条文，已经无法解决大数据产生的新种伦理问题，而法律的制定总是具有明显的滞后性，法律是人们维护自身权益的最后一种武器，失去了法律这一屏障，人们面对利用大数据技术侵害自身权利的行为显得无计可施，这也从客观上加剧了大数据问题的发展。我国在数据挖掘，数据采集、存储、使用权责方面出现了明显的法律空白也是导致大数据伦理问题的重要原因。

大数据产业尚处于发展阶段，法律中对于数据的所有权、数据交易的对象、数据交易中相关主体的权利义务、数据的价值定义、加工过后数据产生价值的分配问题等仍然是一片空白，并没有明确的法律条文明确规定，这也使得原本就混乱的大数据市场难以变得法制化合理化。

数据交易一直在市场中进行，但是缺乏相关法律规定，到底什么样的数据被交易是合法的，法律法规所保护的到底只是个人数据的敏感部分还是整个数据的全部？毫无疑问数据也分为很多部分，用于交易的部分必须不能对个人产生损害，不能涉及个人隐私，不能被精确定位到具体的个体身上，这样的数据交易才可以被容纳在现有法律的条框之下。因此必须对未交易的数据进行脱敏处理，剥除其敏感部分，保留无害部分，才能够确保不能够通过这些数据定位到具体的个人。但这些能够定位到具体个人的数据到底具有什么样的特征和属性呢？利用大数据技术分析挖掘海量信息源，通过交叉复合运算，完全可以把看似完全无害的非敏感数据转化为个人的隐私信息。利用技术手段披露他人不愿意让别人了解的信息，这样算不散侵犯他人利益？将个人放大到企业和国家，他们所具有的机密信息也可能因为大数据技术而面临泄露的风险。

### 7.3.2 行业内缺乏统一的伦理规范

大数据行业业内的数据交易、流通规则，业界虽然已有探索，但在大数据的采集、存储、管理和使用方面还没有形成完备的伦理规范，所以大数据企业各行其是，在认定隐私信息、保护数据安全和制定用户权利时有着各种不同的标准，这些五花八门的标准之间难免会有冲突的地方，这就会引起很多伦理问题。伦理道德往往相对滞后，部分互联网运营商在而对商业利益的诱惑下故意侵犯用户数据权力，通过非法渠道出售和倒卖个人的隐私数据，获取利益。

大数据最常见的用途之一就是用来投放精准的定向广告，商家收集我们的浏览记录和购买记录，并以此来分析我们的喜好和习惯，进而迎合我们的口味，投放我们可能需要的广告。事实上，这件事在大数据时代到来之前就一直在做，只不过大数据技术的到来，让这种分析与投放变得更加精准，更加符合我们的喜好罢了。我们日常使用的微信、QQ 和淘宝等等，时不时就会在朋友圈里冒出各种推广广告，我们的电子邮箱里也是塞满了各种各样的购物和服务广告，或许这些广告都充满了诱惑力，这里面的商品和服务可能有些确实是你想要的，但是，我们不得不承认的一点是，这些铺天盖地的宣传广告中，很大一部分是我们不需要的，我们最终也不会去购买它。一些较为友好的企业可以让用户选择退订广告，用户按照说明操作就可以退订该企业的广告推送服务，但是一些不那么友好的企业就会为了广告的效应而不断的向用户发送广告，用户没有选择退出的权利，即使可以选择退出，退出的手段也颇为繁杂，这些退出按钮往往设置在网页上最不起眼的地方，为了点击这些退出按钮我们不得不按照企业说明的方式去做，经常要打开好几个其他的网页，而且在这期间还不断有广告网页弹出。

### 7.3.3 监督机制不健全

大数据的发展不能只靠国家的强制性律法和企业自身的自律来监督，我们应该更多的发挥人民群众的作用，积极利用第三方来监督大数据企业的行为，防止其侵害普通民众的权利。比如美国已经建立了大数据伦理委员会，用以监督企业大数据技术是否违反了伦理原则，侵害公民的合法权益。他律监督机制的不健全，在互联网行业的高速发展和大数据时代到来的背景下也是大数据伦理问题出现的原因之一。

不但搜集和使用大数据要受到监管，通过大数据技术预测他人的行为尤其应该监管。在网上，很多服务都是不需要收费的，但这种看起来“免费”的模式背后却是无处不在的“监控”行为，免费服务的背后蕴藏着对用户数据的“监视”行为，在网页上嵌入后台程序，用户的所有操作都会被一一记录在案，这很容易泄露用户的隐私信息，这种“监视”具有高技术性特点，其行为难以被一般用户所发觉，国家机关的资源有限，难以做到面面俱到，因此效仿其他行业，建立第三方监督机构来缓解政府的压力，约束企业行为，保障用户权利是极为明智的，遗憾的是由于种种原因，第三方机构虽然屡屡有人提及，但并未实际建立起来，这也使得大数据行为缺乏第三方监督，以至于数字身份伦理问题层出不穷。

#### 7.3.4 社会功利主义风气盛行

大数据诞生至今时间已经不短，但是人们所热衷于讨论的仍然是大数据的商业价值，对大数据中所蕴含的人文精神的探讨不过寥寥，而且现今社会功利主义风气盛行，人们往往依据经济利益来判断一项技术是否成功，这更加助长了人们对大数据经济价值的挖掘，忽略了其中的人文价值，这也在客观上助长了大数据技术的伦理问题。现实中的个人主义会导致个人的自我孤立，享乐主义会让我们只使用大数据技术来满足自己的感官享受，极端的功利主义片面的追求利益最大会会让我们忽视大数据技术所潜藏的负面作用<sup>1</sup>。

功利主义风气容易让我们关注集中于大数据短期的眼前利益，忽视长远利益，例如某企业的一项大数据行为从长远看有益于公益事业的发展，但短期内却难以取得经济效益，在功利主义风气的作祟下，很多人认为该公司的此项行为是愚蠢的表现，而公司的管理人员也因为功利原因判断这是一项失败的投资，所以取消了这种行为，这无疑是一种极其短视的功利主义。

是美国的两大 IT 巨头谷歌和微软公司就较为注重大数据的公益行为，谷歌公司创建了一个国际热线网络收集关于贩卖人口事件的数据，而微软则一直积极利用数据科学来帮助打击侵害儿童。从短期看，谷歌和微软的行为确实无法为企业自身赢得经济收入，但是从长远来看，谷歌和微软的举动不但有益于公共事业的

---

<sup>1</sup> 陈仕伟. 大数据技术异化的伦理治理 [J]. 自然辩证法研究, 2016, (01)



发展，而且在还在群众中获得了良好的口碑，其效果并不亚于花重金打造电视广告的效果。

## 第八章 大数据背景下数字身份伦理问题的对策性研究

大数据是一门复杂性的科学，其内容可以说涉及当今社会的各个角落，由于大数据伦理问题的根源十分复杂，所以解决这些伦理问题也需要宏观性和全面性的考量，所以绝大部分学者都是社会环境全面分析，通盘考虑大数据伦理问题的解决之道。

对策性研究分为两个部分，第一部分是对大数据伦理问题进行伦理治理，也就是我们处理大数据问题，发展大数据技术所要遵循的伦理原则；第二部分则是我们应对大数据问题是所要采取的具体措施。

### 8.1 大数据背景下数字身份伦理问题的伦理治理

无规矩不成方圆，有了伦理准则之后才能相应的制定应对措施，因此为了规范大数据行为，为大数据行为制定对应的伦理标准是必要的，否则个人信息将会遭到滥用，隐私会被曝光，尊严与名誉也会荡然无存。伦理问题当然要用伦理本身来进行治理，坚持必要的伦理原则是解决大数据背景下的数字身份伦理问题的必由之路。

#### 8.1.1 人类应遵循的伦理原则

人类应遵循以人为本的原则，以无害为前提，向善为目的。从人技术发展的目的来看，必须坚持以人为本的原则，任何技术都是为人类服务的，都必须以人为目的，都是为了人类的全面发展，而不是为了技术、经济或者其他任何东西，大数据技术也不例外。对大数据技术的开发、使用和创新必须旨在促进人类健康的发展和幸福的生活。以人为本内在包含两种含义，一是前提，技术的发展不能伤害人类；二是目的，技术的发展要促进人类的幸福，追求高层次的“善”。

1999年，斯皮内洛就在《世纪道德》一书中将信息技术的伦理标准之一定义为“无害性”，即“人们不能使用计算机和信息技术给其他网络主体和网络空间造成直接的或间接的伤害”<sup>1</sup>，大数据技术依然属于信息技术的一种，当然首先也应该以不危害认为为前提，在这个前提之下才能够引导人们向善。无害是技术发展的前提，

人类借由量子物理理论的不断创新发展发明了核能发电技术，但同时也藉由该理论创造出了原子弹，给日本的广岛和长崎带来了毁灭性的灾难，造成了人类

<sup>1</sup> 斯皮内洛. 世纪道德——信息技术的伦理方面 [M] 北京:中央编译出版社, 1999.

社会数十年的“恐怖和平”。我们通过研究化学问题获得了能够开山劈石的炸药和火药，但是讽刺的是无数生命也死于这些炸药之下。所以我们要尽最大努力避免对他人的工作和生活产生任何不必要的损害，也就是所谓的“无害”。

不仅仅只是需要无害，我们发展大数据的目的就是为了让人类生活的质量变得更高，人类能够更好的适应这个世界，更快的进行发展，所以我们还应要求“向善”。每个人都有要对大数据的发展做出一定的努力，维护他人，尤其是弱势群体的权益。

对大数据技术的开发、使用和创新的目的是让社会更加和谐健康的发展，让人们生活的更加幸福快乐，所以任何行动均应遵循不伤害人和有益于人的伦理原则<sup>1</sup>。

### 8.1.2 服务商应遵循的伦理原则

服务商应遵循权责统一的原则，确保用户的知情同意和隐私安全。从权利与义务的角度来看，任何人在使用大数据技术时，必须始终如一的坚持责任心，在享受大数据技术带来的好处时必须承担起大数据技术带来不良后果，这些责任不仅仅该由使用数据的一方来承担，采集这些数据的一方同样也必须承担相应责任。就如同环境治理中的“谁污染谁治理”的原则一样，大数据也要坚持“谁记录谁负责”“谁使用谁负责”的原则，坚持责任与权利相统一。大数据技术的开发、使用和创新必须在承诺保护个人权益的条件下进行，但对不端行为坚持诚信也是大数据技术使用的标准之一，必须坚持权利结合责任，实现谁搜集利用谁负责<sup>2</sup>。

在企业利用大数据技术获益时，必须要负起对等的责任，第一步就是要确保用户的知情同意。确保用户的知情同意是对用户最起码最底线的尊重，在某些时候无法一一征得所有用户同意的情况下，告知用户和进行公示就成了对最起码的尊重。知情同意原则意味着对他人的尊重，要求尊重人的自主性和自我决定权，必须坚持知情同意或知情选择原则。收集个人信息、将个人信息再使用于另一目的时，必须获得同意。政府机构和互联网服务商在提取用户数据时，应该对利害关系人充分讲解数据的存储内容、使用情况以及潜在风险，让利害关系人自主决定是否授权。

无论在大数据时代中拥有何种数字身份，我们都要保护自己和他人的隐私权，不能随意披露他人的个人信息，也要防止他人的隐私泄露。在海量数据关联分析触及揭示个人真实资料的敏感信息时，数据预测方应自动向数据权人发送通知并立即停止挖掘，直至获得对方明示同意。

### 8.1.3 市场行业应遵循的伦理原则

<sup>1</sup> 张燕南，赵中建. 大数据教育应用的伦理思考 [J]. 全球教育展望，2016，(01).

<sup>2</sup> 陈仕伟. 大数据技术异化的伦理治理 [J]. 自然辩证法研究，2016，(01)

整个行业市场应该坚持公平正义原则，平衡大数据利益相关者的利益。从公平正义的角度来看，公正的分配有限的信息资源是必要的，我们要避免在大数据时代产生形成新的弱势群体，减少甚至避免数字鸿沟的马太效应。我们应维护每个人享有大数据技术及应用利益的权利，关注社会中的信息弱势群体，努力缩小数字鸿沟，防止因不适当地泄露个人信息而产生污名和歧视。

大数据涉及到各方面相关者的利益，应对公司与用户，团体与个人，国家与民众之间的利益进行合适的处理，尽量避免不必要的伤害，尤其是对弱势群体的伤害，绝不能让公众利益因为公司和企业的利益而受到损害。如果不得不损害某一方或者某几方的利益，也要做到伤害最小化。

在大数据技术应用中，利益相关者的利益冲突应作出合理处置。任何情况下人民(尤其是脆弱人群)的利益不能因追求专业人员或公司的利益而受到损害，必须要着眼于集体视角，实现大数据利益相关者的利益最大化，同时也要实现伤害最小化。

如果在使用大数据过程中不得不侵害数据生产者的利益，需要及时对其进行必要的补偿，依据情况不同，可以分为精神补偿和物质补偿两种。如大数据使用者对大数据生产者确实产生了伤害，那么就必须对我们给予相应的补偿.....这也是大数据搜集者和大数据使用者必须承担的责任。

#### 8.1.4 政府应遵循的伦理原则

政府应遵循开放共享的原则，实现过程的透明与公开，同舟共济。从大数据的特性角度来说，大数据意味着开放与共享，大数据时代也是开放和共享的时代，开放和共享原则是大数据行为所必须坚持的原则之一。数据的开放和共享、数据间的相互作用很可能会让数据产生新的价值，因此开放共享是未来大数据发展的大势所趋。面对大数据时代的来临，保持积极开放的态度，确立奉献和分享的精神，才会让大数据发挥最大的价值。

大数据的使用应对公众保持透明，避免黑箱操作，这样才能解除公众对大数据的神秘感，值得注意的是，此处的透明指的是各种意义上的透明，包括数据采集、分析和应用的各个数据生命周期的一切环节，无论是对象是国家还是个人，企业还是用户都必须坚持透明原则。

组织在使用要在不同数据过程中提高其负载价值的透明度，将选择权回归个人。大数据时代是一个开放的时代，任何数据只要不涉及个人隐私、组织秘密或国家安全，都应该对公众开放。

参与必须建立在透明的基础之上，大数据不仅仅是一家企业，一个团体，一个国家自己的事情，而是事关所有人的重大问题。所以，要让受试者参与到实验中来，要让公众参与到大数据中来，充分发挥公众的主观能动性，避免闭门造车

的困局。需要采取措施促使公众对大数据技术的了解，并引导所有利益攸关者或其代表在上游就参与大数据技术的研发及其应用的决策过程。

相对于企业与政府，我们普通民众无疑是大数据中的弱势群体，同舟共济原则就是要求强势一方有义务帮助弱势一方，全社会也应多关注弱势群体而非强势群体。我们要维护每个人享有从大数据技术研发及其应用受益的权利，特别关注社会中的脆弱人群。在现代信息活动的知识权力体系下，受益最多的互联网服务商和相关从业人员有责任帮助那些处境最不利者在大数据技术的应用中获益。

### 8.1.5 社会应遵循的伦理原则

整个社会需要坚持道德与法律相结合的原则，完善法律制度，形成道德自律。从法律与道德的关系来看，我们必须要坚持伦理治理与法律治理相结合的原则，不但要加强立法，还要道德自律。不但要坚持伦理准则，坚决对其贯彻执行，还要不断对其进行修订。不能仅仅依靠外界的他律来规范使用大数据的相关行为，在他律的基础上，必须以自律原则相结合来防范社会道德失范现象。

法律是治理社会的必要手段，法律天然具有强制性能够调整社会中盘根错节的利益纠葛，具有不同价值观和道德观善恶利害相关者正是出于对法律这种强制性的恐惧和信任才有了共同的行为规范，他们清楚地明白，自己一方面害怕法律，一旦自己破坏法律必然招致国家机器对自己的制裁，同时又信任法律，一旦有他人不顾自己的利益破坏法律也必然会遭到惩罚，所以守法是所有人必须的行为。虽然道德并不具有这种天然的强制力，但它在社会中依然发挥着举足轻重的作用，法律通常来以道德为基础，法律的最高线就是道德的最低线，又反过来对道德起到强化与促进的作用，二者互为补充，缺一不可，所以道德同样是一种治理社会的必要手段，永远不会退出历史的舞台。

法律规范和道德约束是解决数字身份伦理问题必不可少的两种手段，法律以其强制性从外在对人们的行为进行规范和约束，是构建数字身份伦理规范的有力保障，道德则从内在对人们的行为进行规范和引导，如果法制不利，人们必然会毫无顾忌，肆意妄为，治理数字身份伦理问题必然艰难无比，如果不重视道德规范，人们在法制的重压下必然会产生恐惧和反抗，社会矛盾会不断凸显，数字身份伦理问题的冲突也会不断加剧。

### 8.2 大数据背景下数字身份伦理问题的具体措施

制定强大的应对措施来保护相关者的利益来是必不可少的，保障我们的隐私，安全，自由和正义，从而充分实现大数据研究潜力就必须要有具体的应对措施。但是，旨在保护数据对象的干涉手段，不应该阻止那些愿意参与大数据研究的人们，更不能屈服与人类在生物学上的见解和行为，来阻碍大数据研究的发展空间。

### 8.2.1 制定和完善法律法规，加强政府监管力度

目前我国大数据发展十分迅速，陈旧的法律法规已经无法适应大数据的各种新型伦理问题，为了适应大数据时代日新月异的变化，修改完善现有的和制定新的法律法规是刻不容缓的。同时，由于大数据企业在大数据领域内相对于普通人而言处于绝对的强势地位，大多数企业都是以盈利为目的的，所以我们必须加强政府机关在大数据活动中的监管作用，维护普通人民群众的数据权利。要让我国在解决大数据伦理问题的方面处于领先地位就必须由国家 and 政府为大数据发展提供法律保障和政策监督<sup>1</sup>。

我们可以参考欧美等国颁布数据保护法案，制定大数据法律法规，并加强从业人员的法制概念，来威慑大数据搜集、利用、处理和保护中的不道德行为。比如，欧盟通新颁布的《数据保护法》禁止与他国的数据交换，保证国内的数据流通，保障了欧盟成员的隐私问题；德国为防止个人隐私信息的泄露修改了《联邦数据保护法》；俄罗斯禁止国内互联网企业将公民个人身份信息数据存储于境外服务器上；新加坡新颁布的《个人资料保护法》也对国内数据流到国外加以限制。国际社会的情况告诉我们，我国大数据立法和相关设施建设以及刻不容缓，当务之急是必须加紧立法速度。但考虑到国内实际情况，大数据技术具有很强的专业性而且涉及面极广，目前立法可能有一定困难，但是可以缓和折中出台相应条例，保证大数据产业的健康发展，经过一段时间的修改和完善之后再提升为法律。

### 8.2.2 提倡行业自律机制，建立社会监督模式

相对于伦理道德来说，法律的底线较低，而且法律具有一定的滞后性，所以提倡企业自律，规范大数据从业者的行为，坚持他律与自律相结合也是解决大数据伦理问题的必要手段。在大数据的急速发展模式遍及各行各业，大数据具有的广阔前景和无限商机已经被企业广泛的予以关注，如何充分释放大数据技术所蕴含的能力就成了各行各业所必须思考的问题。在大数据技术爆发惊人的能量时，也在不断地创在新的伦理问题，这些新问题的出现必然会破坏大数据良好发展的稳定性环境，这是很多企业和机构所不愿意看到的，因此这些企业和机构也自发的开始制定相应的标准以规范自身的行为，比如建立数字认证技术，双方订立授权条款等等，但是个别企业和机构的单打独斗很难彻底规范整个行业的行为，因此行业内部有组织有纪律的自发性合作是行之有效的方式。行业自律就是行业内部自行约定、自行规范自行管理自己行业内部的行为，必须首先制定好约定，哪些行为是可以的，哪些行为是不可以的，这些约定的前提是必须要遵守法律法规，这样不但可以保证自己约定的正当性也可以警告其他违反约定的行为。这种体现

---

<sup>1</sup> 唐凯麟，李诗悦. 大数据隐私伦理问题研究 [J]. 伦理学研究，2016，(06).

行业内部道德水平的行业自律也体现出了一种社会性的关怀。当今关于大数据的伦理标准和法律法规相对滞后，如果不抓紧进行自律机制建设，大数据可能会引发各种难以收拾的恶果，所以提倡行业自律机制必须从现在开始。

同时，我们还可以效仿欧美国家，建立第三方监督机构“大数据伦理委员会”，通过社会来监督企业的大数据活动。可以效仿美国，成立第三方的监管机构——伦理委员会，负责对行业道德难题进行伦理评价、对信息活动进行全程监督、制定行业伦理标准、规约个体信息行为等。此外，还可以制定奖惩机制。对于那些自律性较差和缺乏社会责任感的大数据从业者来说，严格的利益奖惩机制是规范其网络行为的一条必由之路。通过对互联网企业进行各种形式的利益奖惩，可以有效地促使其为了自我利益而主动践行信息伦理规范，并最终发展为主体的一种道德自觉。

### 8.2.3 调整隐私观念，强化自我保护意识

大数据时代不断更新和挑战我们的隐私观念，所以必须调整现有的隐私观念，使之与大数据时代的发展相适应，人们在享受着数据便利的福利时，也应该越来越注重隐私伦理道德观的形成。调整隐私观念是解决问题的必由之路。.....虽然目前个人的选择权还相对弱势，但调整我们的隐私观与时代相适应，不断寻求更能保护自身隐私的行为方式还是十分必要的。

由于在大数据时代，我们的隐私安全遭受着前所未有的威胁，所以必须从自身做起，强化自身的保护意识，积极维护自身的隐私安全，不能对侵害自身隐私安全的行为无动于衷。要懂得利用各种道德和法律武器来维护自身的数据权利。

在当前大数据时代的背景下，强化自我认同意识，积极构建一个健康、积极和全面发展的自我。.....最后要增强数据维权意识，一旦发现个人信息遭到泄漏，要及时向有关部门进行申诉和维权。如果像小数据时代一样，对自己的数据采取放任态度，那么这些数据就很可能对自身的隐私、机密、透明、身份和自由选择等等产生威胁。我们的合法权利遭受到不合法的侵犯时，我们要有勇气拿起法律武器于这种侵犯行为相抗争，为自己索取精神方面和物质方面的补偿。

### 8.2.4 努力进行技术创新，降低隐私泄露风险

大数据技术依托于互联网技术，其问题的解决必须也要依靠技术的创新和发展，因此我们必须鼓励对大数据技术进行创新，从技术手段上对大数据技术加以防护，处理脱敏数据以防止利用敏感信息追溯的行为，降低隐私泄露的风险。

我们必须不断升级入侵检测技术、数据加密技术和认证技术、防火墙技术、

病毒防护等技术，从技术层面对大数据伦理加以规范<sup>1</sup>。消除大数据技术的负面效应必须依靠网络安全防护技术的不断创新和发展.....在最大限度上确保数据库的系统稳定和安全运行。首先要完善计算机安全技术，建立一个安全的计算机系统，利用各种手段测试系统安全，找出漏洞并进行及时修补。其次，要引进新的数据安全技术，对敏感字段进行脱敏处理，防止隐私泄露，净化网络环境。

从国家角度来看，要大力支持技术创新，设立专门的大数据基金会，推进与高校研究所的全面合作，实现产学研相结合的发展模式；从技术角度来看，努力发展数据信息安保技术

比如数据信息脱敏技术、防火墙、杀毒软件等，从传播渠道上保障全部网络环境的安全，从技术上对数字身份伦理问题加以控制；从管理角度来看，必须加强数据管理制度，涉及到核心机密的数据必须要采取严格周全的保护措施，不仅要从技术上进行保密工作，也要从人员方面进行考虑，采取多人制共通保管核心机密是一种较为稳妥的措施。

#### 8.2.5 全民提高社会公民素养，消除数据信息不公平现象

大数据涉及整个社会所有公民，虽然现今大数据已经成为热词和关键词，但是公民的大数据素养却依然落后于大数据的迅猛发展，因此对民众大力开展大数据伦理的宣传和教育是十分必要的，授民众以渔，这才能从根本上较弱甚至消除数据信息不公平现象。

大数据应用组织应该对员工进行大数据隐私伦理的思想道德教育，有针对性地开展一些相关道德教育的栏目，以此来引起大数据应用主体们的高度关注，让他们从心理上受到熏陶。专家学者要从多方面向企业、政府和公众开展大数据讲座，帮助群众提升大数据素养。

大数据时代，仅仅知网法律的强制措施来消除数字身份的伦理问题是不可能的，要想消解数字身份的伦理问题必须从根本上提高社会公民的数据信息素养，因此加强伦理教育，开展道德学习就是必不可少的一环，让人们发自内心的形成一种正确的隐私观和价值观。网络道德教育的缺失必须加以弥补，商企政府的通力合作可以更快的解决这一问题，商企政府可以利用资源定期组织一些较有针对性的道德教育讲座，还可以将其录制成视频的形式，方便人们在网络上观看学习，从心理上和思想上对人们进行引导和规范。

---

<sup>1</sup> 陆伟华. 大数据时代的信息伦理研究 [J]. 现代情报, 2014, (10).

## 参考文献

- [1] 杨山. 网络身份盗用类型及防范机制研究[D]. 西南交通大学, 2008.
- [2] 杨诚. 身份犯罪立法的国际动态及其启示[J]. 政治与法律, 2008(08): 35-39.
- [3] 林荫茂. 从信用卡犯罪看身份信息犯罪[J]. 政治与法律, 2008(09): 83-89.
- [4] 薛孚, 陈红兵. 大数据隐私伦理问题探究[J]. 自然辩证法研究, 2015, 31(02): 44-48.
- [5] 安宝洋. 大数据时代的网络信息伦理治理研究[J]. 科学学研究, 2015, 33(05): 641-646.
- [6] 黄欣荣. 大数据技术的伦理反思[J]. 新疆师范大学学报(哲学社会科学版), 2015, 36(03): 46-53+2.
- [7] 张欧阳. 网络民主的要素分析[J]. 兰州学刊, 2012(12): 176-180.
- [8] 文贤庆. 信息理论视域下身份认同的再审视[J]. 云南社会科学, 2017(03): 41-47.
- [9] 宋辰婷. 网络时代的感性意识形态传播和社会认同建构[J]. 安徽大学学报(哲学社会科学版), 2015, 39(01): 149-156.
- [10] L. Jean. "Digital Identity". IEEE Technology and Society Magazine. IEEE.
- [11] 约翰·帕克. 《全民监控》[M]. 金城出版社
- [12] McRae, Shannon (1997). "Flesh Made Word: Sex, Text, and the Virtual Body". In Porter, David. Internet Culture. New York: Routledge. p. 75.
- [13] López, Néstor; Opertti, Renato; Vargas Tamez, Carlos (2017). Youth and changing realities: Rethinking secondary education in Latin America (PDF). UNESCO. pp. 44 - 45. ISBN 978-92-31 00204-5
- [14] 方环非. 大数据: 历史、范式与认识论伦理[J]. 浙江社会科学, 2015(09): 113-120+160+2.
- [15] 赵翔. 数字签名综述[J]. 计算机工程与设计, 2006(02): 195-197.
- [16] 石贵珠. 利用 RSA 密码系统进行数字签名[J]. 太原科技, 2007(03): 82-83.
- [17] 李丽新, 袁焯. 网络安全通讯中的数字签名[J]. 现代情报, 2007(08): 53-56.
- [18] 谢刚, 冯纛, 李治文, 李文鹤. 大数据时代电子公共服务领域的个人数字身份及保护措施[J]. 中国科技论坛, 2015(10): 34-38.
- [19] 黄雯, 翟晓梅. 数字身份与数字裂沟的伦理分析和管理研究[J]. 中国医学伦理学, 2014, 27(01): 15-17.
- [20] 肖俊芳. 从郭美美伪造网络身份谈开去[J]. 上海信息化, 2011(09): 18-20.
- [21] 钟瑛, 刘海贵. 网络身份的意义探析[J]. 复旦学报(社会科学版), 2003(06): 78-8



2.

- [22] 邱仁宗, 黄雯, 翟晓梅. 大数据技术的伦理问题[J]. 科学与社会, 2014, 4(01):36-48.
- [23] 钟光荣. 网络“自我”与现实“自我”的比较[J]. 边疆经济与文化, 2005(03):68-70.
- [24] 王贵斌. 网络参与的身份建构[J]. 新闻知识, 2011(04):49-50.
- [25] 杨晓楠. 互联网实名制管理与公民个人信息的保护[J]. 情报科学, 2012, 30(11):1613-1616.
- [26] 黄雯. 大数据时代信息通信技术应用: 伦理管理和政策研究[D]. 北京协和医学院, 2014.
- [27] 刘卫红. 社会性别与身份认同的语言建构[J]. 江西社会科学, 2014, 34(07):238-242.
- [28] 马云驰. 匿名、流动性与道德需求[J]. 伦理学研究, 2007(03):38-41+67.
- [29] 祝菲. 信息时代不能忽视的文化身份[J]. 理论月刊, 2008(08):79-81.
- [30] 鲍跃华. 从身份认同到角色认同: 党外代表人士代表性的现代转换[J]. 中共浙江省委党校学报, 2012, 28(05):76-81.
- [31] 张毅菁. 数据开放环境下个人数据权保护的研究[J]. 情报杂志, 2016, 35(06):35-39.
- [32] 肖冬梅, 文禹衡. 数据权谱系论纲[J]. 湘潭大学学报(哲学社会科学版), 2015, 39(06):69-75.
- [33] 宋振超, 黄洁. 大数据背景下网络信息的伦理失范、原因及对策[J]. 理论与改革, 2015, (02).
- [34] 王灵芝, 郝明. 医疗大数据的特征及应用中的伦理思考[J]. 医学与哲学(A), 2017, (04).
- [35] 蒋洁, 陈芳, 何亮亮. 大数据预测的伦理困境与出路[J]. 图书与情报, 2014, (05).
- [36] 唐凯麟, 李诗悦. 大数据隐私伦理问题研究[J]. 伦理学研究, 2016, (06).
- [37] 陈仕伟, 黄欣荣. 大数据时代隐私保护的伦理治理[J]. 学术界, 2016, (01).
- [38] 冯登国, 张敏, 李昊. 大数据安全与隐私保护[J]. 计算机学报, 2014, (01).
- [39] 罗弦. 网络新闻生产中大数据运用的伦理问题及编辑对策[J]. 科技与出版, 2015, (01).
- [40] 谭峻楠, 廉小伟, 刘晓韬. 2012 年度数据泄露调查报告摘述与分析[J]. 保密科学技术, 2012(9).
- [41] 黄国彬, 郑琳. 大数据信息安全风险框架及应对策略研究[J]. 图书馆学研究,

2015, (13).

- [42] 胡鞍钢, 周绍杰. 新的全球贫富差距: 日益扩大的“数字鸿沟”[J]. 中国社会科学, 2002, (03).
- [43] 薛伟贤, 刘骏. 数字鸿沟的本质解析[J]. 情报理论与实践, 2010, 33(12): 41-46.
- [44] Tafel H, Turner J C. The social identity theory of intergroup behavior. In: Worhel S, Austin W (eds). Psychology of Inter group Relations. Chicago: Nelson Hall, 1986. 7- 24.
- [45] Turkle, S. (1995). Life on the screen: Identity in the age of the internet. New York: Touchstone, p184.
- [46] 安宝洋, 翁建定. 大数据时代网络信息的伦理缺失及应对策略 [J]. 自然辩证法研究, 2015, (12).
- [47] 陈仕伟. 大数据技术异化的伦理治理 [J]. 自然辩证法研究, 2016, (01)
- [48] 斯皮内洛. 世纪道德——信息技术的伦理方面 [M] 北京: 中央编译出版社, 1999.
- [49] 张燕南, 赵中建. 大数据教育应用的伦理思考 [J]. 全球教育展望, 2016, (01).
- [50] 陆伟华. 大数据时代的信息伦理研究 [J]. 现代情报, 2014, (10).

## 攻读硕士学位期间的学术活动及成果情况

### 1) 参加的学术交流与科研项目

- (1) 生命之善——生物中心主义环境伦理思想研究，中央高校基本科研业务基金，2016年9月-2017年8月
- (2) “原理”课程培育和践行社会主义核心价值观实践教育基地，省部级，2016年9月-2017年12月

### 2) 发表的学术论文（含专利和软件著作权）

- (1) 董军, 程昊. 大数据技术的伦理风险及其控制——基于国内大数据伦理问题研究的分析[J]. 自然辩证法研究, 2017, 33(11):80-85.