



个人信息安全与隐私保护的实证研究

——基于创新扩散理论的大数据应用视角

匡文波 童文杰

摘要: 随着信息技术、通信技术、互联网技术的快速发展,“大数据”已逐渐渗透到社会的各个领域,深入到人们日常生活的方方面面。近五年来,针对大数据的学术研究成果越来越多,整个社会对大数据的认知度和关注度越来越高。大数据作为一种创新应用,已从其创新扩散的增长期步入到成熟期。然而,互联网与大数据应用引发的个人信息泄露问题日益凸显,个人信息安全与隐私保护逐渐备受关注和重视。大数据应用必须关照人性、关注人的需求满足,才能实现这一创新应用的健康、良性、可持续发展。

关键词: 大数据; 创新扩散理论; 信息安全; 隐私保护

一、大数据应用与研究愈演愈烈

目前,大数据在各行业的应用越来越普及。但是,大数据应用的过程中,是否考虑了其数据资源的贡献者——用户的权益? 是否将用户个人信息安全与隐私保护纳入到商业运营考量的范畴? 个体用户对大数据的应用有怎样的感知? 本研究将从这些方面入手,对当前个人信息安全与隐私保护情况进行调查研究。

作为一种创新应用,以“大数据”为主要研究对象的学术成果,集中在信息技术、计算机科学的技术性探讨和行业应用中的方向性探索,有少量成果从隐私权角度开展研究,到目前为止,鲜有学者从传播领域,尤其从受众视角对大数据这一创新应用的“扩散”过程进行深入研究。基于大数据的、关于个人信息安全与隐私保护方面的相关研究成果,在中国知网数据库中,文献总量不足 200 篇,且大部分是信息安全、信息技术、计算机与通信技术等领域的技术研究。从传播学视角进行的有价值的研究成果相对较少。基本没有公开发表的针对用户对大数据应用的态度方面的实证性调查研究成果。

已有的大数据与个人信息安全、隐私保护方面的研究成果中,比较有代表性的观点有:匡文波在《大数据时代的个人隐私保护》一文中提出,国家应该立法确权——个人网络数据,属于个人所有,互联网企业必须告知用户其个人数据被采集情况,未经授权使用和出售数据企业纳入终身黑名单,泄露用户数据甚至牟利,应视为非法行为^①。陈昌凤、虞鑫在《大数据时代的个人隐私保护问题》一文中提出,大数据环境下保护个人隐私应更着重于数据使用者为其行为承担责任,承担保护和管理个人信息的责任。隐私保护需创新途径,如将数据模糊处理,使之不能精确显示个人的隐私信息^②。袁梦倩在《“被遗忘权”

① 匡文波:《大数据时代的个人隐私保护》,载《中国广播》2015 年第 6 期,第 13~14 页。

② 陈昌凤、虞鑫:《大数据时代的个人隐私保护问题》,载《新闻与写作》2014 年第 6 期,第 45 页。

之争:大数据时代的数字化记忆与隐私边界》一文中提出,大数据打破了数字化记忆和遗忘的平衡,引发隐私边界的动态变化。“被遗忘权”的出现,意在改变数据主体难以“被遗忘”的格局,赋予信息主体对信息进行自决控制的权利,并且有着更深的调节、修复大数据时代数字化记忆伦理的社会意涵^①。吴飞在《大数据与“被遗忘权”》一文中提出,“被遗忘权”的提出为数据隐私的保护问题提供了一个有价值的方向^②。董杨慧、谢友宁在《大数据视野下的数据泄露遇安全管理——基于90个数据泄露事件的分析》一文中,通过对所搜集的90条数据泄露事件的数据特征整理和分析,提出应当从防止大数据企业内部员工泄密、加强技术防护、建立首席信息安全观等方面降低数据泄露事件发生的风险^③。王菲在《互联网精准营销的隐私权保护:法律、市场、技术》一文中提出,各网站均可加入Truste,并遵守其所要求的互联网隐私保护的基本原则,合作网站获得Truste的认证标志粘贴在网站上,向消费者表明自己是对消费者互联网隐私负责的网站,当然,对违反政策的网站给予制裁^④。周丽娜在《大数据背景下的网络隐私法律保护:搜索引擎、社交媒体与被遗忘权》一文以搜索引擎——谷歌和社交媒体——脸谱网为例,依托欧盟立法改革和欧洲人权法院判例,分析了被遗忘权在大数据时代的本质含义,以及在数据搜集处理、数据控制和适用范围等方面面临的瓶颈^⑤。申琦在《我国网站隐私保护政策研究:基于49家网站的内容分析》一文中通过内容分析法得出,92%的网站表示会与第三方共享用户信息,但在信息共享声明中对将会与具体哪些机构共享用户个人信息语焉不详。仅有40%的网站明确表明未经用户允许将不会出售用户个人数据。也有网站声明不会与第三方共享用户的个人信息,但附带有相关的豁免条件^⑥。这些观点、结论对本研究具有一定的启发性。

二、假设:大数据的发展符合创新扩散模型

本文以创新扩散理论为依据,提出研究假设:大数据的发展符合创新扩散理论模型,在其未来的进一步发展中,应当以人为本,努力提升大数据应用过程中的人文关怀,加强对个人信息安全与隐私进行保护,提升对用户个人权益的关照水平。

创新扩散理论是美国学者埃弗雷特·罗杰斯(E. M. Rogers)提出的,关于通过媒介劝服人们接受新观念、新事物、新产品的理论。罗杰斯认为,创新扩散是一个过程,在这个过程中,一项创新经过一段时间,通过特定的渠道,在某一社会系统的成员中传播。创新扩散的四个主要因素是创新、传播渠道、时间以及社会系统^⑦。

创新扩散理论认为,创新发展的初期,即引入期,社会系统中通常只有极少数具有冒险精神的创新者能够接受,随着创新者在社会上的活动和影响初步呈现,更多最能把握舆论导向的组织、群体成为早期采纳者,当创新在扩散过程中影响越来越大时,大多数经过深思熟虑的组织、群体成为早期追随者,后期,持怀疑态度的群体逐渐成为创新扩散中的晚期追随者,而墨守成规的落后者最终成为创新的局外人。而创新扩散的整个过程,也因此被划分为创新的引入期、增长期、成熟期和衰退期四个阶段。成熟期创新应用的渗透率达到饱和点,然后开始呈现下降的趋势。如图1。

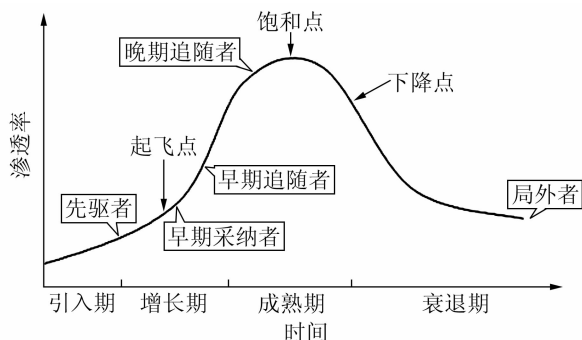


图1 创新扩散理论模型图

①袁梦倩:《“被遗忘权”之争:大数据时代的数字化记忆与隐私边界》,载《学海》2015年第4期,第55~61页。

②吴飞、傅正科:《大数据与“被遗忘权”》,载《浙江大学学报(人文社会科学版)》2015年第2期,第60页。

③董杨慧、谢友宁:《数据视野下的数据泄露遇安全管理——基于90个数据泄露事件的分析》,载《情报杂志》2014年第11期,第157页。

④王菲:《互联网精准营销的隐私权保护:法律、市场、技术》,载《国际新闻界》2011年第12期,第94页。

⑤周丽娜:《大数据背景下的网络隐私法律保护:搜索引擎、社交媒体与被遗忘权》,载《国际新闻界》2015年第8期,第136~148页。

⑥申琦:《我国网站隐私保护政策研究:基于49家网站的内容分析》,载《新闻大学》2015年第4期,第47~49页。

⑦埃弗雷特·M. 罗杰斯:《创新的扩散》,辛欣译,中央编译出版社2002年,第10页。

大数据始于用户上网浏览的各种数据“痕迹”,发展于对用户数据的大量采集、深度挖掘、综合分析 with 个性化应用。先驱者、早期采纳者和早期追随者往往是大数据的技术研发团队、大数据服务提供的企业、社会组织、机构。普通的个体用户可能成为大数据的晚期追随者或者抗拒者。成熟期,晚期追随者的决策尤为重要。大数据时代“得用户者得天下”,然而,当前对大数据的应用,常常是企业的商业化行为,个体用户往往难以直接感知大数据,更难以主动运用大数据。在大数据的整个创新扩散过程中,个体用户往往显得后知后觉,但大数据取之于用户用之于用户,如果用户对大数据持否定、抗拒态度,将其不利于大数据的整个产业性、规模化发展与应用。用户在对大数据持肯定或者否定的决策过程中,个人信息安全与隐私保护状况,最有可能成为个体用户对于大数据应用最大顾虑点和困扰点。

三、方法论证

(一) 数据统计:大数据研究频次呈逐年上升趋势

大数据一词自 2011 年开始进入我国媒体和学界关注、研究的视野。本研究以《人民日报》图文数据库和中国重要报纸全文数据库为样本,用定量的方式,通过对“大数据”一词在传统主流媒体已公开刊登文章标题中出现的频次进行统计、分析,研究近年来大数据在核心纸媒总体报道发展趋势。同时,本研究还选定七大核心学术期刊库进行同类统计分析。七大核心学术期刊库包括:统计源期刊(中信所)、中文核心期刊(北大)、CSCD 中国科学引文库(中科院)、CSSCI 中文社科引文库(南大)、EI 工程索引(美国)、SCI 科学引文索引(美国)、SSCI 社科引文索引(美国)。数据统计的时间跨度从 2011 年 1 月 1 日起至 2015 年 12 月 31 日止。

2011 年以来,大数据一词在传统主流媒体《人民日报》和中国重要报纸全文数据库文章标题中出现的频次越来越高,这一数据在一定程度上体现了大数据创新扩散在传播渠道这一因素中,传统媒体广泛参与了创新扩散活动。其中很多已发表文章为针对企业、社会组织的大数据应用典型探索、案例进行的相关报道。因而,除了媒体传播渠道,创新扩散活动中,社会系统(包括官方组织、企业团体)也逐步作用于大数据的创新扩散活动。五年来,大数据在传统主流媒体中报道频次的的数据走向如图 2。

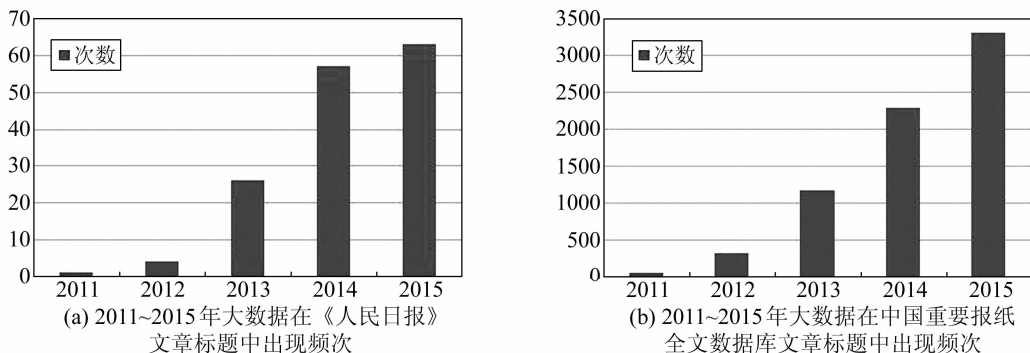


图 2 大数据在传统主流媒体中报道频次的的数据

大数据在各类重要报纸中出现的频次逐年递增,一方面反映了报纸这一传统传播媒介对大数据关注度逐年提升,另一方面,大众媒介报道的驱动力一般来自官方组织和企业团体,因此,官方组织和企业团体对大数据的关注度和贡献力也呈现出逐年递增之势,而 2014 年、2015 年,这种增幅最为明显。

接着,本研究以“大数据”为关键词检索了 2011—2015 年以下七大重要期刊库——统计源期刊(中信所)、中文核心期刊(北大)、CSCD 中国科学引文库(中科院)、CSSCI 中文社科引文库(南大)、EI 工程索引(美)、SCI 科学引文索引(美)、SSCI 社科引文索引(美)的所有学术文章,将所有标题中含“大数据”一词的文章进行了统计,分析大数据在学术领域的发展趋势。

“大数据”在中信所的统计源期刊数据库、北京大学的中文核心期刊数据库两大数据库中研究成果最多,其次是中科院的中国科学引文库和南京大学的中文社科引文索引数据库。

笔者以“大数据”为关键词检索七大数据库的学术论文,发现大数据研究自 2011 年以来呈上升趋

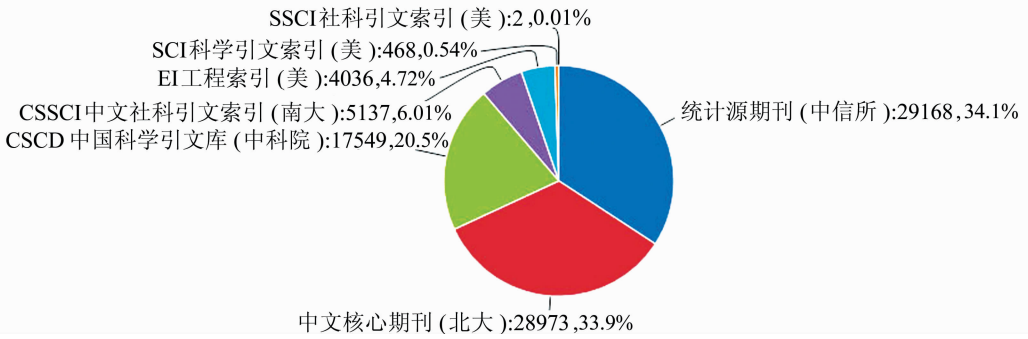


图 3 (2011—2015)大数据在 7 大期刊库分布情况表

势,2013 年增速更为明显。各类型学术成果中,期刊类型的研究数量持续快速上升,会议论文和学位论文类型的研究总量上升到 2014 年的最高点后,开始呈下降趋势。图书类成果的数量较少,上升趋势缓慢。图 4 也从学术研究的角度,初步勾勒出了大数据创新扩散的发展趋势,这一趋势与创新扩散理论模型的图形趋势走向较为一致,因此,大数据发展基本符合创新扩散理论模型。当前,大数据发展已逐渐进入创新扩散的成熟期。

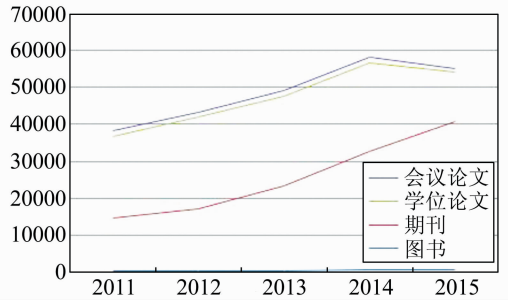


图 4 (2011—2015)大数据在各类型学术发展趋势曲线

七大数据库中国内外主要刊物对大数据研究的文章,大量集中在计算机、通信、信息管理等工业技术领域、金融等经济领域(如图 5、图 6)。大数据创新扩散的过程中,计算机技术领域的人员通常最先了解、认知、分析、接受大数据,然后是商业、金融类大数据应用领域的研究人员,他们作为大数据相关技术的先导,引领相关研究和应用尝试。较之其它领

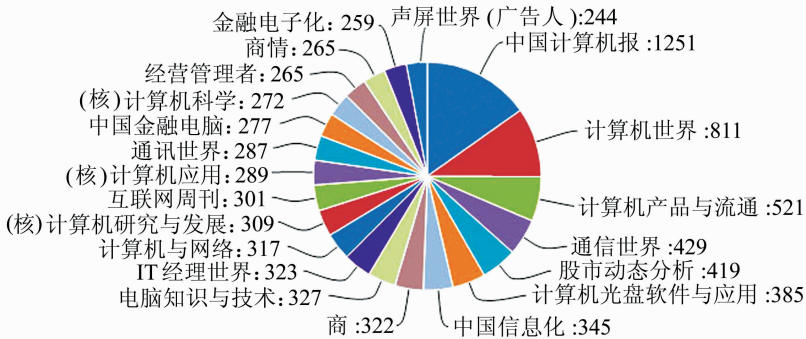


图 5 (2011—2015)大数据在刊物种类统计情况

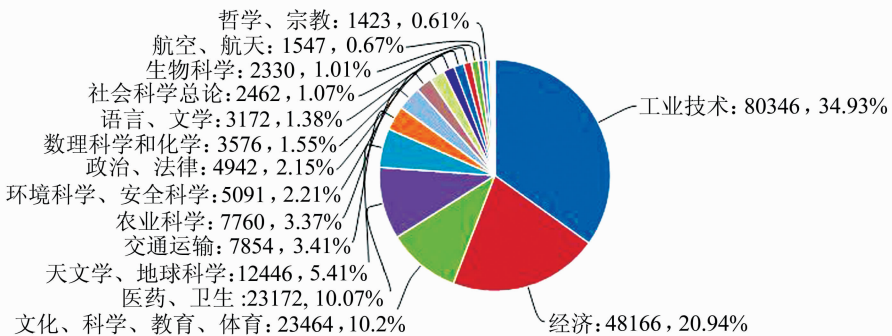


图 6 (2011—2015)大数据在中文学科分类统计情况

域,这些领域的研究人员更早涉足大数据的研究与应用。社科类等以用户为主要研究对象的相关领域,目前形成的大数据方面的研究成果比例相对较低。

以上数据基本证实,大数据的发展与应用,已逐渐进入其创新扩散的成熟期。在引入期和增长期,通信、计算机等技术领域的技术开发、应用人员、研究者成为这项创新应用的先驱者、早期采纳者和早期追随者。成熟期,普通的个体用户可能成为这一时期的晚期追随者或者抗拒者,晚期追随者的决策尤为重要。

笔者在上述“大数据”为关键词检索结果的基础上,选取识别度高、用户最为关注的“隐私”作为关键词进一步检索,得到“大数据”与“隐私”在学术领域的研究情况与发展趋势(如图 7,图 8)。

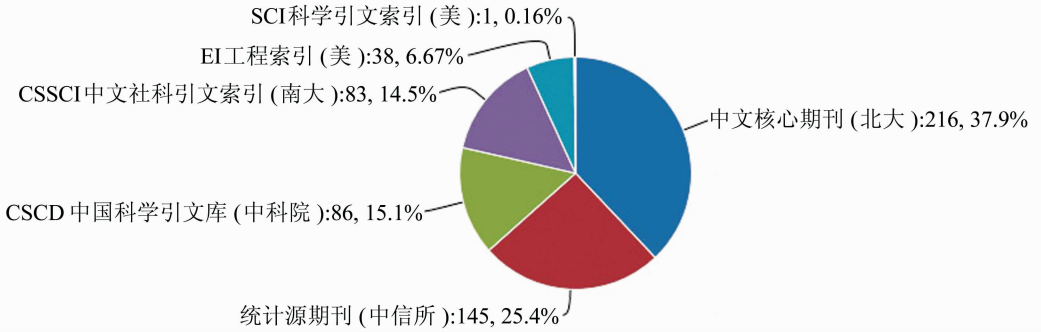


图 7 (2011—2015)大数据、隐私在 7 大期刊库分布情况表

对比图 3 与 7 可知,针对“隐私”这类个体用户(晚期追随者)的研究成果数量在大数据研究成果总量中占比较低,不足 1%。图 8 表明,对大数据与个人隐私方面的研究整体呈上升趋势,基于大数据应用的个人隐私保护越来越受关注。

依据创新扩散理论,作为晚期追随者的个体用户对于大数据应用的感知和决策在成熟期显得尤为重要,而与隐私相关的个人信息安全与隐私保护,则是用户关注的重要方面,且这种关注程度总体呈上升趋势。

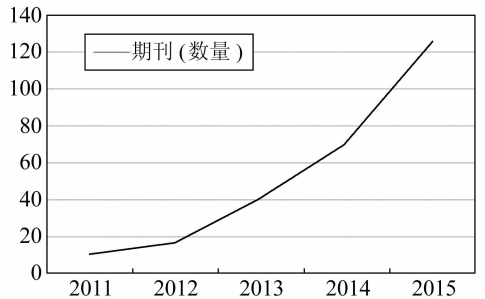


图 8 (2011—2015)大数据、隐私-各类型学术发展趋势曲线

(二) 深度访谈:个人数据安全与隐私保护最受关注

大数据的应用是一个复杂的过程,且往往是企业和组织的商业化行为,个人对于大数据应用的真实感知并不明显,对大数据应用存在的问题与后果的感知也各不相同。本文认为,深度访谈法适合了解复杂、抽象的现状,通过深度交流,获取主题相关的各类信息,才能进行综合概括与分析汇总。因此,为了进行深入研究,尽可能了解个人对于大数据应用的态度、对个人隐私安全的现状认识,本文采取深度访谈法,对每个被访者进行 15 分钟左右的半开放式访谈。

为了消除被访者被访时可能受到的心理压力,本文采取无结构的一对一深度访谈。其优点在于消除被访者的各种顾虑,便于受访者更自由、真实地表达自己的想法。另一方面,以一种客观的、对等的访问态度,与受访者进行平等的交流、充分的情感互动,这样能保障访谈所获取信息的真实性和深刻性。对比问卷调查和小组座谈法,深度访谈能更好地帮受访者做到所想与所表达信息的一致性。

本文选择了中国比较有代表性的四个城市:北京、上海、广州、武汉,选取了 30 名不同年龄、行业、学历、专业、性别的受访者。访谈时间为 2016 年 2 月 18 日~3 月 18 日。

访谈过程中,大部分受访者对大数据有所了解,少数人不知大数据为何物。对大数据有一定了解的受访者中,少数人对大数据应用存在的问题和后果并不了解,从行业和社会发展方面支持大数据应用,但大部分受访者都表示曾经有个人数据和隐私信息遭泄露的经历,并对个人信息安全与隐私保护表现出严重的担忧,希望政府等相关部门能制定有效政策、措施,加大对大数据规范性、科学性应用的管控力

度,加强个人信息安全与隐私保护,维护个人隐私权益和生命财产安全。不了解大数据的受访者表示日常工作、生活中有没有大数据应用都无所谓。

表1 受访者基本信息表及受访内容回复简述一览表

编号	受访者	地域	年龄	职业	回答要点	对大数据应用的困惑与顾虑概括
1	冯女士	北京	35	传媒企业普通职员	主要从新闻上了解到大数据,用大数据分析国家、行业的趋势很好,个人对大数据应用没有太多切身体会,总体上不抵制也不支持大数据发展	个人隐私信息泄露和个人网络账户安全
2	佟女士	北京	30	部门助理	经常网购,喜欢部分浏览器利用大数据分析提供便捷的个性化产品推送服务,希望大数据给日常生活带来更多方便,但由于多次遭遇各类网络安全威胁,对目前的大数据应用有一定的隐忧	隐私容易泄露且被泄露后很难追踪泄露的根源和渠道
3	王先生	北京	44	商务经理	大数据这些年被媒体和一些公司炒得很火,但很多数据的真实性值得考量,很多企业、组织为了商业价值最大化,对数据进行刻意加工处理	大数据的真实性监督机制和大数据服务公司的诚信管理
4	张女士	北京	26	广告设计	个人十分看好大数据应用,媒体报道中用大数据采集的新闻非常有说服力且可信度很高,在日常生活中感受不到个人使用大数据,希望能开放一些大数据对个人的应用	没有
5	付女士	北京	37	国企经理	目前很多应用只能算是数据分析,还算不上是大数据,大数据在于海量的信息和深度挖掘,但目前只有极少的组织、企业获得了能堪称海量的大数据。个人觉得大数据分析非常利于官方机构、大型机也的顶层决策,比如行业性的资源、产能配置、流通布局等,如果可能,希望以后管理的项目中利用大数据指导决策的制定	从个人角度考虑,大数据就是将个人的所有信息记录在册,必要是进行分析、挖掘,挖掘的深度代表着隐私的泄露程度,个人角度不看好
6	何先生	北京	39	地产经理	大数据对于分析地区性市场、需求、供给有一定的帮助,能够更清楚地挖掘潜在客户,帮助提升营销推广活动的效率和成功率。但大数据应用也存在很多冗余信息,还无法做到供需匹配的更精准,希望未来大数据应用更加精准有效。	加强信息安全
7	黄先生	北京	33	媒体主编	中国缺少国家级的大数据,所以很多大数据实际掌握在企业手中,企业为了盈利必然让其数据有所倾向,隐私,个人往往成为大数据应用的无辜受害者。中国应当建立国家级大数据资源库,利用官方的权威性建立公平可信的大数据平台,并对现有大数据应用过程中的各种侵权、数据泄露以及由此带来的诈骗等恶性发展形成有效遏制	大数据应用缺乏有效管控,个人隐私没有保护机制
8	索女士	北京	55	事业单位职员	除了简单的微信聊天外很少上网,不了解大数据,只看过电视里关于大数据的新闻报道,基本没有遭遇过网络安全危害,但经常会收到基金、房产销售电话、短信	普遍化统计分析尚可,希望不要侵犯个人隐私
9	朱女士	上海	35	宣传主管	利用大数据分析进行适当的宣传推送可以,但现在资料被到处卖,特别是房产和保险推销,很讨人厌	优化个人隐私保护,减少各类骚扰
10	陈先生	上海	36	研究员	大数据应用目前大部分是企业的商业行为,确实利于一些企业产品的推广、市场与生产的布局,个人对于大数据的感知并不明显,除了不知来源的各类骚扰邮件、电话、短信	大数据市场化发展的同时如何实现人性化关怀
11	尚先生	上海	33	IT工程师	很多IT公司、互联网公司都有自己的数据平台,在企业内部的应用还是合理的,比如决策市场取向、产品策略制定等。当外部化使用时,就需要考虑是否影响了一些群体的自身利益。此外,每个数据采集机构和相关工作人员都应当加强数据安全保护意识并自律	缺乏数据泄露的根源追踪机制、渠道、途径,以加强大数据应用的规范性

12	贾女士	上海	28	企业职工	大数据应用目前有一定的市场,也有各种各样的问题,用得好是创新,用不好是祸害,尤其对普通群众而言	大数据创新的可持续发展
13	王女士	上海	23	在校研究生	个人学习生活中并不需要大数据服务,不喜欢网上搜索了一次后,各种相关产品扑面而来的感觉,我会自己去找我想要的东西,而不是网站你告诉我应该买什么	侵犯个人隐私
14	夏女士	上海	29	法律顾问	不反对大数据的应用与发展,但网路无隐私,我们只要上网,就会在网络上留下痕迹,少则可能被人肉,多则遭受账户密码及财产丢失,个人应加强隐私保护	隐私安全和个人权益保护
15	许女士	上海	36	资深记者	大数据利于新闻采写,为我的工作提供很多便利和很多素材,趋势性的数据统计、分析还是相对必要的,过于个体性的数据,应当隐去部分细致的、敏感的信息,对个人的数据分析应该征得用户个人的同意,这样可避免很多对个人的各类骚扰行为发生	大数据整个产业的健康发展
16	杨先生	上海	37	传媒公司副总	只要通过有效的数据统计、分析,你可以得到细致到惊人的商业数据。只要进行数据挖掘,你将毫无隐私可言,大数据往往比你更懂得你。	大数据应用管理规范化与大数据实际应用时对个人隐私方面相关保护的考量
17	廖先生	深圳	38	技术开发工程师	大数据是一种发展趋势,在整个大环境下,个人应加强网络安全意识,减少在网络上各种平台输入各种私人信息,尤其不正规的网站,降低信息泄露的几率,此外,很多网站利用读取本地缓存来窃取用户行为、账户等信息,建议大家多清理缓存	个人权益和隐私的保护
18	王女士	深圳	23	工厂工人	不了解大数据的应用,只要对不侵犯个人权利就好	无
19	余先生	深圳	31	企业职工	大数据并没有给我带来什么好处,并不想贡献自己的资料、数据成就所谓的大数据,但现在很多 APP、网站都要求填写个人信息才能使用,所以个人很被动	很多时候不得不贡献甚至丢失个人信息
20	吴女士	深圳	55	早餐摊贩	不知道大数据,认为与我无关	不知道
21	肖女士	深圳	40	家庭主妇	日常跟互联网的数据有关的,大概就是通过交易记录、评价条数和信息,来选定饭店或者购买商品的商家,对大数据应用、了解不太多,但如果各种骚扰短信、电话是源于大数据应用,个人还是非常厌烦	希望大数据应用有些底线,不要过度骚扰个人
22	袁先生	深圳	36	互联网金融 COO	大数据对我们行业还是比较有用的,创业初期需要招揽一定的用户,一方面我们自己平台积累用户,另一方面通过一些合作方拿到潜在客户的一些资料,或者将我们的产品推广给潜在客户,对产品创业发展初期非常有帮助,相信很多创业公司都会选择类似的方式积累客户	相关应用对创业公司来讲成本不低
23	冉女士	深圳	28	护士	听说了一些,感觉身边每个人上淘宝网,都会有不同的商品推荐,就是大数据分析的结果,在其他网站也会有一些推荐,都是近期曾查找过得类似商品,有时候很方便,有时候的场合比较尴尬,相比之下,觉得可以用需要这样的服务	应该更加考虑个人的感受
24	童女士	武汉	37	医生	大数据应该有用,比如国家对一些大病的监控、治理和规定制定,但目前具体的日常工作生活中很少涉及	病人的隐私应多加保护
25	郭女士	武汉	31	列车员	支持国家对大数据的发展,反对部分商家、黑客为了自身利益侵害他人利益,平常会用手机上网,遭遇过账号密码被盗等风险,希望国家和相关部门加强管控。	网络安全

26	贾先生	武汉	26	电商运营策划	熟悉电商运营尤其是大数据在电商领域的应用,大数据利于企业、商家进行更为精准的市场推广、销售等商业行为,但近年来,随着数据价值的挖掘,将用户数据私下买卖几乎成为圈内潜规则,比如电商、地产、保险、快递行业。目前用户数据买卖价格少则2~3元/个,多则5~10元/个。在用户不知情的情况下对其个人数据进行交易一定程度上侵害了用户的权益,长此以往不利于大数据的发展。只有国家出台相应的管理办法或法规才能解决现在的问题	最大的困惑是解决大数据应用过程中用户个人隐私信息泄露和非法买卖,保障大数据应用合理合法的发展
27	柳女士	武汉	24	创业网店店主	经常有推荐购买用户数据,也有同行通过数据刷信用,抢客户,个人觉得适当的应用可以帮助产品销售,过度使用反而影响小店信誉,造成客户丢失	数据买卖、刷信誉这些情况是否应该杜绝,还电商竞争一个干净的环境和生态
28	易先生	武汉	45	初中教师	不了解大数据的应用	无
29	李女士	武汉	20	在校学生	听老师们说大数据感觉很高深、很复杂,实际接触不多,没有直观印象。不过阿里巴巴的马云说过一句话,阿里巴巴对一个人的了解程度超过你自己,电脑比你更了解你。然而我并不要让电脑比我更了解我自己	如果想要通过大数据了解我,是否应该先征得我的同意
30	何先生	武汉	52	大学教授	大数据非常值得研究,目前的大数据取得很大进步,较之以前的研究方法更加科学,但也不尽完善,部分社交、电商平台大数据资源多,实际运作中经常进行有目的的挖掘、应用,官方需加强管控,减少各类由此引发的社会问题	大数据企业化发展与个人利益、社会安全稳定的协调

创新的宗旨往往是优化现状,增强便捷性,它们或以更方便的方式满足人,或将人们从低成本、低效率的运行现状中解放出来。大数据的发展,代表着科技与社会的进步,但大数据的发展,必须关照人性,以满足人的需求为核心,应充分体现其人文关怀。大数据的创新扩散与应用过程,离不开对人性的尊重和人的各种需求的满足。

(三) 抽样调查:网络安全威胁与个人隐私保护不容忽视

前面的数据分析和深度访谈研究,初步验证了本研究的假设。为了弥补深度访谈样本量低可能存在的不足与误差,进一步验证研究假设的科学性和有效性,本研究通过互联网渠道进行了随机的抽样调查。抽样调查的样本规模设计为500人,调查时间为2015年3月10—15日,有效回收样本为431人。调查对象按照从事的年龄、性别、职业、学历等进行条件划分。问卷设计的目的是分析用户对于大数据这一创新应用的态度、困惑与建议。调查问卷中,针对以下具体问题进行了调查统计:1. 上网过程中最大的担忧,2. 曾遇到过的网络安全威胁,3. 对网站采集用户数据的行为支持与否,4. 对大数据应用的态度,5. 安全性最差的互联网应用,6. 对加强互联网个人隐私保护的建议。随机抽样调查的结果如图9、图10。

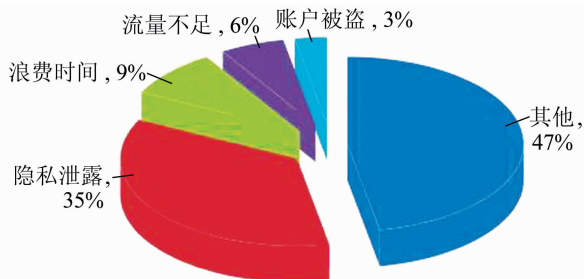


图9 用户上网过程中最大的担忧

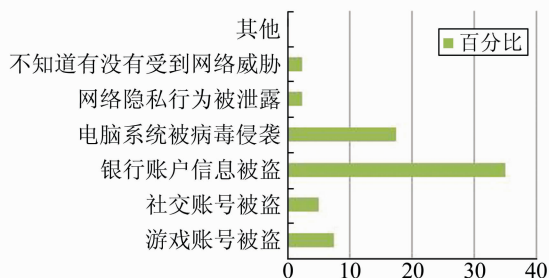


图10 用户曾遭受的网络安全威胁

被调查者在上网过程中,最担心账户被盗和隐私泄露,其次是浪费时间、流量不足等问题。

在被调查者遭受的网络安全威胁中,电脑系统被病毒侵袭的情况发生最多,其次是社交账号被盗、

网络隐私行为被泄露,然后是游戏账号被盗、银行账户信息被盗等,部分被调查者由于账号被盗,曾遭受过一定程度的经济损失。

被随机调查的人群中,约 68%的人反对在上网过程中采集用户数据信息,但大部分人表示不知道个人数据信息何时何地、通过何种渠道采集等。



图 11 对互联网用户数据采集行为的态度

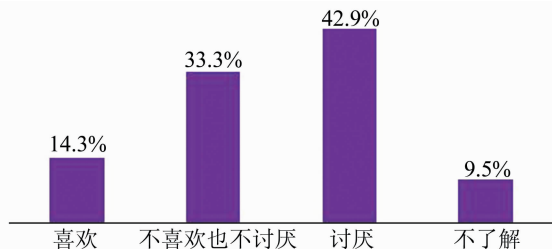


图 12 对大数据应用的态度

对于大数据应用,在被随机调查的人群中,42.9%表示讨厌,33.3%觉得无所谓,14.3%很享受当前大数据带来的个性化服务,9.5%表示并不了解大数据。对大数据应用的态度表示讨厌的被调查者大都认为个人信息安全存在严重隐患,表示无所谓的人则觉得与我无关或与我无害,表示很享受当前个性化服务者,大都是电商消费的忠实者,喜欢“千人千面”这种个性化的信息推送、产品推介。

在各类互联网应用中,大部分被调查者认为社交应用安全性最差,其次是购物、娱乐、互联网金融及游戏类应用。

在被调查的人群中,36.5%的人群认为应该由政府加强相关法律法规、制度建设与监管,24.2%的人群认为应该由行业加强网络服务提供商数据采集、应用的监督与道德自律,22.6%的人群认为互联网服务平台应通过技术和管理手段加强安全性,16.1%的人群认为应该由个人努力提高自身信息意识和相关技能。

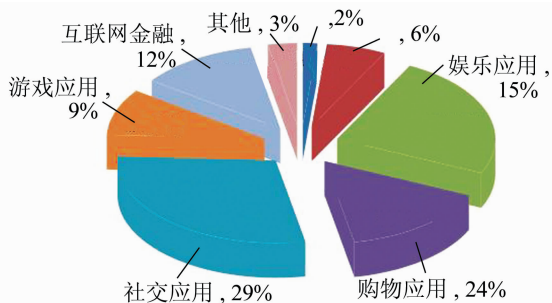


图 13 个人信息安全性最差的互联网应用

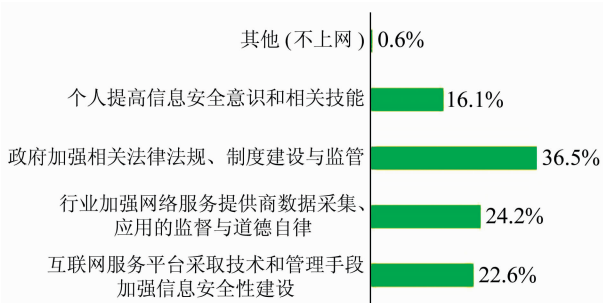


图 14 对加强互联网个人信息的建议

上述数据支持了研究假设,即:大数据的扩散与应用过程中,应提升对社会系统中用户这一群体个人权益的关照水平,尤其体现在对个人数据信息安全和隐私的保护上。只有充分关注人性的需求,正视并解决发展中存在的缺陷与不足,才能得到受众的支持,推进大数据这一创新应用的继续健康、可持续发展。

四、研究结果

本研究通过数据统计分析和深度访谈,基本验证了大数据的发展符合创新扩散理论模型。大数据经过多年的发展,已经达到创新扩散的成熟期,其追随者即将接近最高点。这一时期,用户对于大数据的倾向性尤其重要,用户的选择直接表现为对大数据应用的肯定与否定。

经深度访谈、随机抽样问卷调查可知,从行业趋势应用分析、市场定向精准推广、用户定制化服务等角度来说,大数据应用基本受到社会的认可,但从个人数据安全和隐私保护权益的角度来讲,大部分被调查者对当前的大数据应用持怀疑、甚至否定的态度,主要原因有以下几点:

首先,对个人来说,大数据应用带来的弊远远大于利。用户对于大数据带来的便捷性感知较弱,比起频繁被各类销售、诈骗电话和短信骚扰,以及账户被盗、各类隐私数据泄露等情况,他们更倾向于放弃当前大数据为个人提供的各种服务。调查中对大数据应用持无所谓态度的用户,大都是年龄偏大、学历偏低或者基本不上网、不太了解大数据、对数据采集无感知的人群。

其次,用户在使用某项互联网应用时,无论该应用服务为有偿提供或无偿提供,用户都无法跳过被应用平台采集个人数据信息的环节。通常,在进行用户登录或 APP 应用安装时,读取用户信息作为一项霸王条款,让用户不得不同意,否则无法进入服务或者安装的下一步。在深度访谈中,多名用户表示,安装 APP 过程中不同意被读取个人数据的结果只能放弃使用。也有很多网站以网络协议、用户手册等形式,要求用户在使用网站相关服务之前必须接收数据采集相应条款,部分网站甚至将此类条款用多级页面方式隐藏很深,用户很难发现。再次,大部分网站使用 Cookie 收集用户信息,而绝大部分用户不会自行禁用 Cookie,更不知道使用 Cookie 的后果,当数据被采集时,用户毫无感知,更无法感知某次、某时的上网行为,其是否遭遇隐私数据信息泄露、在什么平台遭遇了数据泄漏。当前也没有任何有效机制、手段、渠道帮助用户进行追查,用户处于十分被动、弱势的地位。

最后,部分企业、组织以提供各种应用、服务为名,采集用户信息并进行交易,并未提前征得用户的同意,也并未将数据买卖的收益分享给用户,用户信息在买卖交易中被作为交易的商品,用户在整个交易过程中只有权利的失去,没有利益的回报。部分黑客也以盗取用户数据信息为目的,并借助各类用户信息,窃取用户信任,进而开展诈骗等非法活动。本研究通过深度访谈和问卷调研还得出以下几个方面的结论:大数据处于快速发展阶段,而针对大数据采集应用的法律法规严重缺失,用户除了缺乏数据泄露的追踪渠道、手段之外,还缺乏相关的法律知识保护。计算机、通信行业的用户比其他行业更了解这项新技术,因而在实际应用中会关注或者考虑通过技术手段进行隐私信息保护,然而对大部分其他行业的用户来说,信息安全意识淡薄。社交、游戏、娱乐、金融是普通用户最可能使用的应用服务,然而通过调查发现,这些类别应用的个人信息安全性最差。在寻求加强对个人信息安全和隐私进行保护的方面,用户对政府管理诉求最高,其次是行业监管以及企业自律。

五、结 论

技术决定论者认为技术引起整个社会的创新发展,但技术不是独立于社会之外,一项新的创新,其扩散效果取决于这项创新对于社会的构筑。技术是社会发展的产物,一项新技术的创新扩散活动,必然受到社会准则和价值观的影响。经济因素,比如利润,引发了创新,但这项创新最终要面对消费者,面对个体用户。创新扩散的最后一个阶段是创新的结果,即接收或者拒绝。因而,为了让创新的结果得到接受,创新扩散过程中,对人性关照、对人的需求的满足,这些人文关怀非常重要。

然而,当前的大数据时代,个人信息安全保护现状堪忧。2016年8月21日,山东临沂准大学生徐玉玉因被诈骗电话骗走上大学费用9900元,伤心欲绝,郁结于心,最终导致心脏骤停,不幸离世。19日,徐玉玉接到一通陌生电话,对方声称有一笔2600元助学金要发放给他,在这通陌生电话之前的18日,徐玉玉曾接到过教育部门发放助学金的通知,让她办理了助学金的相关手续,并被告知助学金几天就能发下来。骗子掌握的信息如此精准,因此完全骗去了徐玉玉的信任。而徐玉玉母亲的电话号码、考大学信息、助学金信息是如何被泄漏的?新闻晨报的记者通过调查后发现,网络上公然售卖考生信息者大有人在,而这些信息包括考生姓名、学校、电话、住址等,考生信息成为黑客利用安全漏洞从事地下黑色产业的关键内容。个人数据信息轻易地被获取并非法买卖,而买卖后引发的结果,就是夺走了花期少女的生命。徐玉玉之后,时隔八日,清华教授被电信诈骗1760万,这起诈骗的背后,更凸显了大数据时代个人信息安全与隐私保护面临的空前威胁。诈骗者掌握数据之及时、准确,正反映了大数据时代个人信息泄漏的严重性、加强信息安全与隐私保护工作的紧迫性。个人信息安全与隐私保护已经成为一种社会性的问题,如处理不当,轻则损害人们的个人隐私权,重则伤及人们的生命财产安全。大数据时代,科技让人们的生活更加方便,同时,也为人们的生命、财产安全威胁埋下了隐患,个人信息的泄漏事件频

频发,必将影响这一创新应用的进一步发展。

大数据作为一项社会性的创新应用,在其成熟期的扩散、发展过程中,应重点凸显社会系统的人文关怀,积极关照整个社会准则和价值观,强化对用户这一社会群体的需求满足、权益维护,保障个人数据信息安全,做好用户隐私保护。首先,我国目前关于个人信息保护的法规尚不健全,立法机关要加强个人隐私保护法规的立法工作,严格规范政府部门、企事业单位对公民个人信息的采集活动。其次,司法机关要对未经公民知情与允许,出售、泄露公民个人信息、恶意或过度商业化使用公民个人信息等非法行为依法予以严惩。遗憾的是,迄今为止,我国尚没有对出售、泄露公民个人信息等行为追究刑事责任的案例。第三,政府部门要依法严格规范并减少企事业单位不必要的公民个人信息采集活动。例如:不少旅游点购买门票需要收集用户身份证信息,张家界等景区甚至还要采集用户指纹,等等。这种公民个人隐私信息的采集不仅毫无必要,而且这些景区管理部门事实上很难安全保管公民的隐私信息。第四,即使是政府部门要采集公民信息,也需要有合法的授权及法律依据。如果相关政府管理部门、电信公司、互联网公司过度地推行实名制,将加大大数据时代个人信息保护的难度。作为政府管理部门、电信公司固然不会有意泄露个人隐私信息,但是很难确保个别的职员不被利益诱惑去出卖公民个人隐私信息。此外,黑客的攻击也严重威胁个人信息安全。美国互联网公司雅虎 2016 年 9 月 23 日发表声明,证实黑客在 2014 年对其发动大规模网络袭击,窃取至少 5 亿雅虎用户的信息,占了雅虎 10 亿用户的一半以上。这应是有史以来规模最大单一网站信息泄露事件。

大数据的应用与良性健康发展,必须直面当前商业活动运行背后的各类用户数据安全与隐私泄露的危害,应用新的技术手段和方法,积极调动全社会力量,共同致力于加强个人的数据信息安全与隐私保护,建立起完善的用户信任机制,只有这样才能营造大数据发展良好、有序的新型市场环境,推进大数据全行业的健康发展。

An Empirical Study on Personal Information Security and Privacy Protection:

From the Perspective of Big Data Applications Based on Innovations Diffusion Theory

Kuang Wenbo (Renmin University of China)

Tong Wenjie (Renmin University of China)

Abstract: With the rapid development of information technology, communication technology and Internet technology, big data has gradually penetrated into all areas of the society, and also, all aspects of people's daily life. In recent five years, there are more and more academic research results about big data. The awareness of big data of the whole society is rising and more and more attention is being paid to it. As a kind of innovative application, big data is transitioning from its diffusion growth period to maturity period. However, people are facing the increasingly serious privacy leak problems, which caused by the application of Internet and big data. More and more people are concerned on personal information security and privacy protection. People who carry out the applications of big data should pay more attention to humanistic care, and try best to meet the needs of people, so can it lead this innovative application to a good future with healthy and sustainable development.

Key words: big data; innovations diffusion theory; information security; privacy protection

●作者地址:匡文波,中国人民大学新闻学院,中国人民大学新闻与社会发展研究中心;北京 100872。

Email:kuangwb@ruc.edu.cn。

童文杰,中国人民大学新闻学院;北京 100872。

●基金项目:国家社会科学基金一般项目(11BXW037)

●责任编辑:何坤翁

