

美国关于大数据技术应用的隐私保护 问题评估及其启示

李修全¹, 宋卫国¹, 李微微²

(1. 中国科学技术发展战略研究院, 北京 100038;

2. 黑龙江省科技资源共享服务中心, 哈尔滨 150028)

摘要: 数据的共享使用与隐私保护是一对矛盾体, 隐私问题已经成为制约大数据技术应用的关键挑战。2014年5月, 美国发布大数据技术评估报告, 其中, 对大数据发展过程中涉及到的隐私保护问题进行了分析。“数字化排放”的信息、“大海捞针”挖掘的海量数据、个性化定制过程等等, 都可能隐含许多个人信息和隐私, 对此, 美国政府的应对措施有: 研究元数据保护措施, 吸引专业技术人才, 强化数据第三方使用管理, 保障隐私保护技术研发经费等。我国大数据发展同样面临隐私保护的挑战, 建议: 组织开展个人隐私保护专项立法研究, 制定完善的个体数据使用规范, 加大隐私保护技术研发投入。

关键词: 美国; 大数据; 信息泄露; 隐私保护

中图分类号: TP309(712) **文献标识码:** A **DOI:** 10.3772/j.issn.1009-8623.2015.02.009

近几年, 大数据技术应用在全世界发展迅猛, 这一方面源于数据采集与处理技术上的能力拓展, 而更重要的动力则是来自于我们已经开始能够利用一部分个人数据, 包括: 电子化交易及各种网络活动所形成的数字化痕迹等^[1]。得益于大数据技术的运用, 越来越多的个人数据开始被发掘并产生商业化价值。然而, 对于应以何种方式使用个人数据, 如何把握数据的充分利用与个人隐私保护之间的平衡, 各国还没有形成统一认识, 法律和制度约束缺失, 隐私保护的技术保障也没有跟上, 已经成为影响大数据技术应用进一步发展的关键挑战。

2014年5月, 美国总统行政办公室发布了题为《大数据: 把握机遇、坚守价值》的评估报告^[2], 对大数据技术的快速发展将会对经济社会产生的影响及面临的挑战做出了全面评价。评估报告认为, 在大数据被用来造福社会的同时, 也会产生社会危害, 可能对保护公众隐私权、确保社会公平、反对

种族歧视等美国核心价值观造成冲击, 必须采取措施以防范这些潜在威胁。其中, 大数据发展对公民隐私权的挑战被认为是这项新技术发展对于传统伦理和价值观念的最大冲击。

1 大数据发展中的隐私挑战

隐私权问题最早是1890年美国哈佛大学的沃伦与布兰德斯在《隐私权》的论文中提出的^[3], 而信息技术变革给人们工作生活带来便捷的同时, 也带来了越来越多新型的个人隐私侵犯事件, 隐私保护在数字化时代已经成为无法回避的问题。这提醒我们, 在大数据发展初期, 应充分认识开展数据隐私立法研究的迫切性。美国政府评估报告中指出了大数据发展中面临的主要隐私挑战。

1.1 “数字化排放”带来的威胁

信息技术正在改变着人们的生活方式, 现代人的生活中, 每天要进行形式各异的电子化活动。传

第一作者简介: 李修全(1976—), 男, 副研究员, 主要研究方向为科技指标、科技政策。

收稿日期: 2014-10-17

感器、公众网络、移动应用程序等不断成为新的数据来源，电子化活动中，我们的一举一动都会留下痕迹，使人们的日常行为产生很多记录信息，或者说，我们每天都在进行着“数字化排放”，而且排放的信息量越来越大，其中，像药店票据、手机地理坐标、IP地址等记录，就会包含个人信息。随着数据分析功能日益强大，众多看似毫无关系的数据碎片，可能会被整合分析，识别或还原出个人行为特征，从而暴露个人隐私。如何保护这些数据及其中挖掘出的信息，并使之得到合法的使用，是当前面临的紧迫难题。

1.2 “大海捞针”的代价

随着处理器计算能力的增强、存储成本的降低以及算法性能的改进，科学家已经有能力收集并处理海量数据，并从中找出异常数据模式，这就是被称为“大海捞针”的数据挖掘方法。它能够为人们提供通过其他方法不能获得的深度信息，已经成为科学研究甚至大数据行业应用中重要的计算模式。例如，大样本量的个体基因数据集对于识别遗传变异对疾病的影响起到关键作用^[4]。然而，在搜集到的海量数据内，隐含了许多个人隐私信息，这是大数据技术应用面临的另一项关键挑战。

1.3 个性化定制的后果

互联网用户浏览网站痕迹、购物过程等不同类型的记录数据被结合在一起，就能勾画出某个人的行为图谱，并可能分析出他们的个人偏好和行为习惯，从而可以进一步预测消费者的潜在需求，使提前为他们提供所需的信息、产品或者服务成为可能。个性化定制已经成为目前互联网应用的一大特色和亮点，它使人们既能有机会接触到尽可能多的新鲜信息，又避免了被铺天盖地的互联网信息所掩埋。实际上，这种个性化定制的过程同时伴随着个人隐私的发现和暴露。数据挖掘与信息整合技术在使大数据功能日益强大的同时，也对目前个人隐私的保护带来了严重的冲击。

1.4 数据保存的影响

在过去传统的信息传播与保存方式下，人们可以通过信息自然消亡的方式来保护隐私。比如，在经过一定周期后，纸质材料自然毁坏或被销毁，上面的记录信息随之消失。但在数字世界，信息可以被无损地拷贝、传播并且无限期地保存。以前存

储数据的成本昂贵，现在大量数据可以储存在微小的芯片中，简单便捷。技术的进步使得数据一旦产生，就可能保持永恒有效。随着大数据的不断发展，会有越来越多的私人信息产生并在他人的控制下保存，其中，涉及的个人隐私信息保护的问题必须提前考虑。

1.5 大数据时代的儿童隐私保护

青少年是移动应用与网络社交平台等新型互联网应用的最活跃用户。当他们使用这些网络应用时，他们的个人信息也会在不经意间被记录与存储，其中，可能包含个人隐私等敏感信息。相对于成年人，儿童对于个人信息和隐私泄露往往不在意，也不具备必要的自我保护能力。这些无意中被保存的数据有可能在他们成人后对他们的生活产生影响。如今，这类互联网隐私信息侵害行为正在不断发生，如何在给儿童提供互联网探索与学习空间的同时，又确保他们不至于因一时的疏忽而在日后受到挥之不去的侵扰，是世界各国都需要认真思考的问题。

2 美国的应对与计划

面对上述挑战，《大数据：把握机遇、坚守价值》评估报告建议，美国政府应通过制定政策与法律来约束和规范新技术的使用，从而在一定程度上维护个人隐私受保护等基本价值观。

2.1 研究元数据保护措施

“元数据”是关于数据的数据，用于描述数据自身特征。相对数据内容本身，以往元数据被认为透露了较少的信息，利用价值不大，往往不作为隐私信息进行保护。随着大数据技术的兴起，大量分散的数据记录可以重新整合，可能提炼还原成为个人的隐私材料，元数据从此有了新的地位与价值，针对元数据的保护措施也需相应进行调整。未来，元数据作为大数据世界的重要成员，将会被发现有越来越多的利用价值，应该得到更为周密的隐私保护。评估报告认为，随着大数据时代的到来，美国政府应加强对元数据保护政策的研究和制定，对数据服务提供商的行为进行规范并加强监管。

2.2 吸引专业技术人才

目前，大数据技术应用不仅是一个技术问题，更为重要的是它会带来一系列与技术相关的社会问题。因此，需要在政府部门培养具备大数据技术知

识的人才，建立最适合于技术应用和发展的社会规范体系，以充分发挥大数据发展潜力。根据近期一份关于美国公共与非盈利部门的吸引力及技术人才状况的报告，尽管许多年轻技术人员热衷于公共服务，但私营部门给出的条件和机会却更具吸引力，导致大数据技术人才更多地流向私营企业而非公共部门^[6]。该报告建议，美国政府应致力于营造一个更具吸引力的工作和文化氛围，破除公共部门人才流动障碍，充分利用大数据技术人才的作用，发掘大数据技术在政府公共服务中的发展潜力。

2.3 强化数据第三方使用管理

直接与消费者联系并收集信息的公司通常被称为“第一方”，比如，在线零售商、通信运营商、社交网站等；大众消费者是第二方；“第一方”可能自己使用采集的数据，但有时也会将数据转售他人，这种情况称为“第三方”使用，例如，将消费者购物信息用于信贷审核、保险、就业、医疗等。这些第三方公司与信息对象用户之间并没有发生直接的关系，消费者的信息是被间接收集和使用。在这种情况下，消费者往往无法了解其自身信息被扩散和使用的情况，可能带来巨大的隐私侵犯风险。美国政府计划，通过立法方式，增加大数据第三方使用中消费者隐私保护规定，确保社会弱势群体避免受到不公正的对待。

2.4 保障隐私保护技术研发经费

评估报告称，奥巴马政府一直在优先考虑如何加强对大数据相关隐私技术及个人数据使用的社会问题的研究。网络和信息技术研发（NITRD）计划每年在个人隐私保护方面的研究投入已超过7 000万美元。这些研究主要分为保护隐私安全的技术、企业遵从隐私法的研究、医疗保健隐私的保护和隐私保护的探索性研究等4个部分。

3 我国大数据应用中的隐私保护状况

根据中国社科院发布的《中国城市基本公共服务满意度评价（2012—2013年）》报告^[6]，在全国19 843份问卷中，有21.25%的人表示经常遭遇信息泄露，有50.46%的受访者表示偶尔会遭遇信息泄露，只有28.30%的公众表示信息未被泄露过。这充分反映出我国目前个人信息泄露问题已经较为严重，大数据发展遭遇到隐私问题的严重困扰。我

国隐私保护现状与美国既有共性，又有其自身特点。

3.1 隐私保护立法更为滞后

目前，美国、欧洲等地已经建立了各自的隐私保护法律体系，而我国在这方面发展相对滞后。对互联网信息泄漏违法事件中，数据采集的“第一方”、“第三方”，以及数据泄露者、传播者，分别应当承担何种法律责任，我国仍缺乏相关法律条文的确切规范。实际上，我国已经相继颁布了《关于加强网络信息保护的決定》^[7]及《电信和互联网用户个人信息保护规定》^[8]等指导性文件或部门法规，然而，实际约束力还不足，导致各类个人隐私侵权及个人信息泄露事件频频发生，其影响日益严重。

3.2 数据使用的约束与监管严重缺位

近年来，随着我国互联网快速普及城乡，电子商务蓬勃兴起，2013年1—9月，北京市网络销售额已经首度超过传统百货业^[9]。然而，网上购物后，个人信息被泄露或转卖的事件越来越频繁。消费者往往购买过一次某类商品或服务，随之而来的可能就是纠缠不休的垃圾短信和推销电话。频繁的骚扰电话和垃圾短信，已经成为许多人每天不得不忍受的问题。个人信息被非法收集和滥用，既对人们正常生活工作带来严重影响，也为各类诈骗犯罪提供了便利。而有效监管的缺失，导致违法者受经济利益驱使，严重侵害民众个人隐私权利。我国个人信息采集使用方面，监管力度亟需加强。

3.3 新的隐私侵害出现国际化趋势

某些国际大型跨国公司在智能数码产品市场上占有优势，也会对我国个人隐私信息的保护带来威胁。2014年7月，某世界著名手机制造商承认，其预先嵌入智能手机中的数据提取技术，使该公司技术人员能够提取手机用户的个人数据，包括，短信、联系人信息及照片等。这再次引起我们对于个人隐私保护甚至信息安全的普遍担忧，数字化时代的个人隐私侵犯已经出现国际化特征。

可见，大数据技术是一把双刃剑。发展大数据技术应用就要最大限度地开发利用数据，而数据用得越多，个人隐私泄露或权益侵犯的问题就可能越严重。如果不保护或处理不当，就有可能使公众个人隐私信息被不适当泄露而带来社会问题，有可能为此付出巨大代价。因此，我国要想跟上甚至引领世界大数据产业发展，必须对技术应用中所面

临的关键性挑战有清醒的认识，并提前采取技术和政策措施，才能在未来国际大数据产业竞争中占据主动。

4 启示与建议

目前，阻碍大数据技术应用进一步发展的并不是技术本身，而是数据成为改变商业模式的关键要素后给社会所带来的冲击。隐私保护问题需要在大数据发展初期就引起国家相关部门高度重视，并通过立法、行业规范、技术措施等多种途径，予以应对，为大数据技术应用的健康发展奠定良好基础并铺平道路。

4.1 组织开展个人隐私保护专项立法研究

大数据技术的全球化应用，使任何国家都必须认真面对隐私保护问题。2014年5月，欧洲最高法院关于因“被遗忘权”起诉谷歌公司的一项裁决，已经将网络时代的被遗忘权问题提到法律界人士面前^[10]。我国应尽快组织开展隐私保护立法研究，以适应大数据技术发展带来的严峻挑战。通过建立完备的、适应于新的技术条件的隐私保护法律体系，使政府执法机构能够更加有效地打击个人隐私侵权行为，并使个人诉讼和维权有法可依。同时，大数据发展的全球化特征，对研究和完善我国隐私保护立法与全球各国隐私保护框架之间的兼容性提出了新的要求。

4.2 制定完善的个体数据使用规范

数据作为资产已经为企业带来新的商业价值，而对个体数据商业化使用中的规范与管理也应提高到新的高度。随着大数据技术应用向各行各业加速渗透，对于个体数据使用的隐私处理规范也已迫在眉睫。我国应严格规范用户信息使用告知，明确标注所采集数据的目的、使用方法，强化数据第三方使用的监管；同时，研究建立有利于全程数据隐私处理的数据标准与采集规范，使不同类型的个体数据记录可以在受到安全控制的条件下方便访问，并将企业违反个体数据隐私处理规范行为纳入社会征信体系。

4.3 加大隐私保护技术研发投入

自21世纪初“大数据”在全球兴起之后，我国紧跟此次技术变革，在学术界迅速掀起了“大数据”研究热潮。但与此同时，科研项目申请中的盲

目性也随之出现，各种与大数据沾边的科研项目一哄而上。建议在大数据技术研发的项目支持中，加强技术发展路径研究，明确首先要解决的关键挑战。数据的共享使用与保护隐私是一对矛盾体，隐私保护已经成为阻碍大数据技术应用的迫切需要解决的问题。在关注体系结构、挖掘算法和处理能力提升的同时，应优先保障信息加密、去身份化等隐私保护技术研发，以加强大数据技术发展背景下个人隐私保护及国家安全保障能力。

参考文献：

- [1] Gantz J, Reinsel D. Extracting Value from Chaos, IDC[R/OL]. (2011-06)[2014-05-26]. <http://www.emc.com/collateral/analyst-reports/idc-extracting-value-from-chaos-ar.pdf>.
- [2] Executive Office of the President of the White House. Big Data: Seizing Opportunities, Preserving Values[R/OL]. (2014-05)[2014-06-05]. http://www.whitehouse.gov/sites/default/files/docs/big_data_privacy_report_may_1_2014.pdf.
- [3] Warren S, Brandeis L. The Right to Privacy[J]. Harvard Law Review, 1890, 4(5): 193-220.
- [4] Kellis M. Importance of Access to Large Populations[R]// Big Data Privacy Workshop: Advancing the State of the Art in Technology and Practice. (2014-03)[2014-06-05]. http://web.mit.edu/bigdatapriv/ppt/ManolisKellis_PrivacyBigData_CSAIL-WH.pptx.
- [5] Ford and MacArthur Foundation. A Future or Failure?: The Flow of Technology Talent into Government and Civil Society[R/OL]. (2013-12)[2014-06-20]. <http://www.fordfoundation.org/pdfs/news/afutureoffailure.pdf>.
- [6] 中国社会科学院, 等. 中国城市基本公共服务满意度评价(2012—2013)[R]. 北京: 中国社会科学院, 2013-12.
- [7] 中国人大网. 全国人民代表大会常务委员会关于加强网络信息保护的決定[EB/OL]. (2012-12)[2014-06-20]. http://www.npc.gov.cn/npc/xinwen/2012-12/29/content_1749526.htm.
- [8] 中华人民共和国工业和信息化部. 电信和互联网用户个人信息保护规定[EB/OL]. (2013-07-19)[2014-06-20]. <http://www.miit.gov.cn/n11293472/n11294912/n11296542/15514014.html>.
- [9] 新华网. 北京网购零售额破 637 亿元 超传统百货业[EB/OL]. (2013-11-14)[2014-06-20]. http://news.xinhuanet.com/tech/2013-11/14/c_125699555.htm. (下转第 76 页)

Mapping the Research Collaboration Between BRICS Countries During 2001–2012 Using Co-authorships

LIU Ya

(Institute of Scientific and Technical Information of China, Beijing 100038)

Abstract: Based on 16,807 articles published during 2001–2012 and collected by the SCI-EXPANDED database and the SSCI database, this paper analyzes the research collaboration between 5 golden brick countries. The empirical results show that collaboration over the past 12 years presented an upward trend; Internal collaboration within BRICS countries wasn't the major mode of collaboration, most collaborative work was established via non-BRICS countries; Collaboration encompassed a broad spectrum of research disciplines with common focuses and also individual country interests; Collaboration, in general, was playing an important role in improving the research impacts of BRICS countries; and China was the leading country among five countries and should consolidate its advantage by enhancing self-directed orientation and focusing on its priority fields.

Key words: BRICS countries ;SCI article ;research collaboration ;collaboration mode ;research impact

(上接第 66 页)

- [10] 中国网. 欧洲法院力挺“被遗忘权” 谷歌需更改其搜索结果[EB/OL]. (2014-05-15)[2014-06-20]. http://www.china.com.cn/news/world/2014-05/15/content_32396026.htm.

Assessment of US Federal Government on the Privacy Protection in Big Data Development

LI Xiu-quan¹, SONG Wei-guo¹, LI Wei-wei²

(1. Chinese Academy of Science and Technology for Development, Beijing 100038;

2. Heilongjiang Science and Technology Resources Shared Service Center, Harbin 150028)

Abstract: Data sharing and privacy protection are contradictory to each other in most circumstances. Safeguarding privacy has emerged as one of the key challenges that restrict the development of big data. The U.S. government released an assessment report on the big data technology in May 2014, and the privacy protection was analyzed in the report. “Digital exhaust”, massive data needed by “finding a needle in a haystack”, customized online services etc., are all likely to involve personal information and personal privacy. The measures taken by U.S. government include: taking metadata protection measures, attracting specialized talents, enforcing “third-party” data management, and ensuring research funding on privacy protection technology. In China, privacy protection is also a challenging problem in the process of big data technology application. Suggested measures include: starting legislation study on personal privacy protection, improving regulations of using personal data, and increasing funding support for privacy protection technology research.

Key words: the United States ;big data; information leakage ;privacy protection