

国际关系中的 大数据变革及其挑战*

蔡翠红

【内容提要】 随着智能设备和视频设备等器材的普及,廉价存储设备、高速宽带和云计算的诞生以及社交化和移动化的发展,大数据时代已经来临。在当前具有高度不确定性和快速变化特征的数字化生存环境中,对大数据的高效利用已成为各国获取制信息权的关键,大数据亦将成为国际关系的又一重要研究领域。无论从经济、主权角度还是从安全角度,大数据都在国际关系中占据了重要地位,并将对之产生变革性意义:其作为未来最大的交易商品、经济资源和经济资产的商业价值决定了围绕大数据的新国力竞争;其信息载体的本质奠定了数据主权的基础,数据主权也必将成为各国的博弈对象;其所蕴藏的重要情报信息及预测性功能昭示了对国家安全和战略能力的重要意义。同时,大数据所伴随的新数据鸿沟、数据霸权以及数据跨国安全威胁等也对国际关系的平等性、民主性和合作性提出了严峻挑战。中国应正视大数据时代所带来的机会,从政府顶层设计、一体化建设以及大数据治理等几个方面进行战略布局。

【关键词】 大数据;国际关系;数据主权;数据鸿沟;大数据治理

【作者简介】 蔡翠红,复旦大学美国研究中心副教授,博士。(上海 邮编:200433)

【中图分类号】 D815 G203 **【文献标识码】** A **【文章编号】** 1006-9550
(2014)05-0124-20

* 本文为笔者主持的国家社会科学基金项目《21世纪中美关系中的网络政治研究》(项目批准号 12BGJ018)和上海市哲学社会科学规划课题《互联网与政治文化安全研究》(项目批准号 2009BGJ001)的成果之一。感谢《世界经济与政治》杂志的匿名评审专家提出的建设性修改意见,文中错漏由笔者负责。

随着云技术的发展,“大数据(big data)”吸引了越来越多的关注。大数据不仅迅速成为计算机行业讨论的热门概念,而且也开始渗透到国际关系领域,成为21世纪人类面临的又一政治和经济挑战。2013年五/六月号的《外交事务》杂志刊登的《大数据的崛起》一文正式揭开了国际关系和外交事务中大数据研究的序幕。^①同年六月份开始逐步发酵的“棱镜门”事件进一步使大数据问题白热化,美国利用各大公司进行大数据获取、分析的实质使各国开始广泛关注数据和大数据的安全问题,大数据也因此成为国际关系中一个必要的研究课题。

一 大数据时代的来临

我们已经进入了被数据包围的时代。2010年2月,《经济学人》杂志指出,数据信息已经从稀少发展到了超级富余,数据及其影响无处不在,从商业到科学,从政府到艺术。^②更有学者称我们已进入了“数据工业革命”阶段。^③数据之于信息社会就如燃料之于工业革命,是人们进行创新的力量源泉。^④20世纪初,电报、电话等新的信息流通方式促进了大规模生产的发展,如今,充足的数据资源将可能促进以数据工厂为动力的新一轮工业革命。以云计算为基础的大数据可能引领商业、科学和社会等各个领域的革命性突破。^⑤一种以数据为基础的生存方式开始出现,一种新的“知识基础设施”正在形成,^⑥一个新的大数据时代已经浮现。^⑦

(一) 多途径数据来源催生大数据

2011年5月,麦肯锡全球研究院(McKinsey Global Institute)发布报告《大数据:创

① Kenneth Cukier and Viktor Mayer-Schoenberger, “The Rise of Big Data,” *Foreign Affairs*, Vol. 92, No. 3, 2013.

② “Data, Data Everywhere,” *Economist*, February 25, 2010, <http://www.economist.com/node/15557443>, 登录时间:2013年8月20日。

③ “数据工业革命”一词是计算机科学家乔·赫勒斯坦(Joe Hellerstein)所提出。参见 Abhishek Mehta, “Big Data: Powering the Next Industrial Revolution,” <http://isites.harvard.edu/fs/docs/icb.topic1346257.files/Big%20Data%20Powering%20the%20Next%20Industrial%20Revolution.pdf>, 登录时间:2014年2月16日。

④ 维克托·迈尔-舍恩伯格、肯尼思·库克耶著,盛杨燕、周涛译:《大数据时代》,杭州:浙江人民出版社2013年版,第230页。

⑤ Randal E. Bryant, Randy H. Katz and Edward D. Lazowska, “Big-Data Computing: Creating Revolutionary Breakthroughs in Commerce, Science, and Society,” December 22, 2008, pp. 1-7, <http://www.cra.org/ccc/initiatives>, 登录时间:2013年8月20日。

⑥ David Bollier, *The Promise and Peril of Big Data*, The Aspen Institute, 2010, p. 1, http://www.aspeninstitute.org/sites/default/files/content/docs/pubs/The_Promise_and_Peril_of_Big_Data.pdf, 登录时间:2013年8月20日。

⑦ Steve Lohr, “The Age of Big Data,” *The New York Times*, February 11, 2012, <http://www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html>, 登录时间:2014年2月16日。

新、竞争和生产力的下一个新前沿》^①，大数据开始备受关注。根据维基百科的定义，大数据指无法在一定时间内用通常的软件工具进行捕获、管理和处理的数据集合。^②而研究机构加特纳(Gartner)则给出了这样的定义：“大数据”是需要利用合理成本的新处理模式进行处理才能使其拥有者具有更强的洞察力和决策力的海量、高增长率和多样化的信息资产。^③

大数据首先是一种数据，其数据体量巨大(存量大、计算量大)、数据类型繁多、价值密度低(但价值高)^④、速度快(velocity,增长速度快,处理速度也要求快)。大数据最主要的特点并不是其数据“量”的巨大,而是其数据生态系统相互关联的特征,即数据之间的相互关联。大数据的价值也来源于数据之间的相互关联,例如个体不同信息之间的关联、不同个体相互之间的关联、不同群组之间的关联、甚至是信息本身的结构关联等。所以大数据处理是一种能力,即用关联方法从看似没有价值的数据中挖掘出有用信息的能力。^⑤

大数据并非单纯指互联网世界中人与人的交互信息、物联网世界中的商品和物流信息以及企业内部的经营交易信息等。大数据不仅包含传统数据库收集的结构化数据,还包含许多非结构化数据(unstructured data),例如话语、图片、影像以及电子探头数据等。人、物及其相互之间的接触都会产生大量的数据,越来越多的电子探头和视频监控摄像头每天也产生海量数据。因此,图像、视频、音频文件等基于社交媒体的文件以及在互联网上工作后得到的信息都称为大数据。例如通过天文望远镜拍摄的图像、视频数据;卫星云图数据;工业设备、汽车、电表上的数码传感器测量和传递的有关位置、运动、震动、温度、湿度乃至空气中化学物质变化的数据。随着智能设备和视频器材的普及,大数据的来源将越来越广泛。

(二) 技术与时代背景助推大数据的发展

大数据的产生既有技术背景,也有时代背景。从技术层面看,移动终端的普及、廉

① James Manyika, et al., *Big Data: The Next Frontier for Innovation, Competition, and Productivity*, McKinsey Global Institute, 2011, http://www.mckinsey.com/Insights/MGI/Research/Technology_and_Innovation/Big_data_The_next_frontier_for_innovation, 登录时间:2013年8月20日。

② http://en.wikipedia.org/wiki/Big_data, 登录时间:2013年8月20日。

③ <http://www.gartner.com/it-glossary/big-data/>, 登录时间:2013年8月20日。

④ “价值密度低”的含义是指有价值信息的比例比较低,以视频为例,在连续不间断的监控过程中,有用的数据可能仅仅是长短为一两秒的视频信息。

⑤ Maude Bonenfant, Marc Ménard, André Mondoux and Maxime Ouellet, *Big Data and Governance*, April 2013, p. 2, <http://www.inter-disciplinary.net/critical-issues/wp-content/uploads/2013/04/mondouxcyberpaper.pdf>, 登录时间:2014年2月17日。

价存储设备、高速宽带和云计算的诞生是大数据时代的前提。^①大数据的相关技术涉及数据产生、数据收集、数据存储和数据分析等方面。各种电子图像自动收集器、化学和生物感知器的普及,数据收集、数据挖掘和数据分析的自动化,数据云存储能力的提升以及各种提取数据技术、语义分析技术、图文转换技术、模式识别技术、行为分析算法的产生都使大数据走向应用阶段。

从时代背景来看,首先,大数据产生于当前具有高度不确定性和快速变化特征的环境中。当前的决定对今后的影响将非常重要,而这些决定又高度依赖不断变化的社会环境数据,数据基础上的定量决策将越来越流行。其次,数字化、网络化、信息化使人的生存方式发生了巨大的变化,并由此带来一种全新的生存方式,即数字化生存。^②某种程度上说,大数据逐渐成为了一种“电子空气”,^③成为了一种支持人类社会的系统资源。数字化生存不仅成为个体的生存方式,即应用数字技术在数字空间工作、生活和学习的全新生存方式,而且还成为一种社会存在状态,即一种新的以数字化形式显现的社会生存现象和文化现象。

(三)用途不断拓展决定了大数据的重要性

“大数据时代的预言家”维克多·迈尔-舍恩伯格(Victor Mayer-Schonberger)指出,大数据开启了一次重大的时代转型。大数据正在改变我们的生活以及理解世界的方式,成为新发明和新服务的源泉,而更多的改变正蓄势待发。^④数据不仅可用来治国,而且“得数据者得天下”。^⑤

从用途来看,大数据早已不是科学家的专利领域,大数据最早用于营销、预测、市场分析、公共政策制定等。目前,大数据的应用已经拓展到人文社会科学研究、政府机构、教育部门甚至有需求的个人。大数据的用途基于这样两点认识:一是所有信息都是具备预测能力的,只是等待何时被解锁;^⑥二是基于更多人群或对象的数据所得出

① David Bollier, *The Promise and Peril of Big Data*, p. vii.

② 数字化生存(being digital),最初是由美国学者尼葛洛庞帝在其1996年出版的《数字化生存》一书中提出的,按照他的解释,人类生存于一个虚拟的、数字化的生存空间,在这个空间里人们应用数字技术从事信息传播、交流、学习、工作等活动,这便是数字化生存。参见尼葛洛庞帝著,胡泳等译:《数字化生存》,海口:海南出版社1997年版。

③ Danah Boyd and Kate Crawford, “Six Provocations for Big Data,” Paper presented at Oxford Internet Institute’s Symposium, September 21, 2011, p. 2, <http://ssrn.com/abstract=1926431>, 登录时间:2013年8月20日。

④ 维克托·迈尔-舍恩伯格、肯尼思·库克耶:《大数据时代》,第1页。

⑤ 维克托·迈尔-舍恩伯格、肯尼思·库克耶:《大数据时代》,第294页。

⑥ Staffan Truvé, “Big Data For the Future: Unlocking the Predictive Power of the Web,” October 19, 2011, <https://www.recordedfuture.com/big-data-future-unlocking-predictive-power-web/>, 登录时间:2014年2月16日。

的分析结果更有价值。如果是来自人群的数据,这一观点被称做群众智慧(wisdom of crowds)。^①在公共卫生、经济发展和经济预测等领域中,大数据的预见能力已经崭露头角。联合国开始的一项新的发展倡议“全球脉动(Global Pulse)”^②和美国谷歌公司(Google)开发出能预警流感传播的“谷歌流感趋势系统”^③便是大数据应用的两个典型案例。^④

随着应用范围的扩大,大数据已经成为一种文化、技术乃至学术现象。^⑤从文化上看,大数据已经成为一种新的社会文化和时代象征。从技术上说,大数据是数据收集、分析的前沿技术,是从各种类型的庞大数据中快速获得有价值信息的能力。从学术领域来看,社会学、政治学等人文社会科学学界也开始运用大数据进行相关研究。对于政治家而言,通过对推特(Twitter)、谷歌、脸谱(Facebook)、微信、微博等信息的分析,可以跟踪大城市的抗议者、发现恐怖主义行径。对于政治学者而言,对微博动态、国会演讲和新闻发布等相关大数据的分析是进行政治思想传播研究的重要手段。

尽管对大数据的应用也同时存在各种各样的质疑,^⑥例如有观点认为大数据很有前景,但数据噪音太多或数据噪音应对技术不成熟会导致数据价值大大降低,无法有效从大数据中挖掘价值;也有观点认为,数据一旦超过了某个量级点,反而会对相关性分析带来很多干扰;还有观点认为,和其他新技术一样,大数据的发展也必然回避不了硅谷臭名昭著的技术成熟度曲线。^⑦但是,不可否认的是,我们生活在大数据时代,数据成为了原始生产资料,成为了一种具有经济和社会价值的新资源。对每天快速产生

① James Surowiecki, *The Wisdom of Crowds*, New York: Random House, 2004.

② 联合国的“全球脉动(Global Pulse)”项目拟借助对自然语言的解译软件对社交网络和文本消息等进行“情绪分析(sentiment analysis)”,来预测某给定地区的失业情况、开支削减以及疾病爆发等,从而将早期预警信号和援助项目相结合,以帮助地区发展或避免其陷入困境。

③ 很多美国人在发现流感症状后首先选择的是上谷歌搜索相关信息,然后做出判断。谷歌据此开发出能预警流感传播的“谷歌流感趋势”系统。“谷歌流感趋势”在测试过程中还显示出反应迅速的优势,甚至可能比疾病控制和预防中心提前1个星期到10天公布流感预报。参见 Miguel Helft, “Google Uses Web Searches to Track Flu's Spread,” *The New York Times*, November 12, 2008, <http://www.nytimes.com/2008/11/12/technology/internet/12flu.html>, 登录时间:2013年8月20日。

④ 更多大数据应用案例,参见郭昕、孟晔编著:《大数据的力量》,北京:机械工业出版社2013年版;冯启思著,曲玉彬译:《数据统治世界》,北京:中国人民大学出版社2013年版。

⑤ Danah Boyd and Kate Crawford, “Critical Questions for Big Data,” *Information, Communication & Society*, Vol. 15, No. 5, 2012, p. 663.

⑥ 大数据的主要问题一般概括为:可扩展性(scalability)、冗余性(redundancy)、偏见(bias)、噪声(noise)、垃圾邮件(spam)和隐私(privacy)。参见 Ricardo Baeza-Yates, “Big Data or Right Data?” <http://www.inf.udec.cl/~loreto/AMW13/paper14.pdf>, 登录时间:2013年8月20日。

⑦ 技术成熟度曲线又称技术循环曲线或炒作周期,是指新技术、新概念在媒体上的曝光度随时间变化的曲线。参见维克托·迈尔-舍恩伯格、肯尼思·库克耶:《大数据时代》,第9页。

的海量数据的高效利用已成为各国获取制信息权的关键。而且“棱镜门”事件等已经使我们无法再对之视而不见,大数据问题应该成为国际关系研究的一个新的重点。

二 大数据在国际关系中的含义

作为一种革命性现象,大数据将影响人类社会发展的方方面面。它标志着人类社会正在从信息时代经由知识时代并快速向智能时代迈进。^①与农业时代的人口和土地、工业时代的钢铁和石油一样,数据已成为未来支撑国家安全和可持续发展的重要战略资源,同时也成为国际关系中需要重点关注的新竞争领域。

美国是全球最早关注大数据的国家。2010年3月,美国政府就要求各部门实施自己的大数据战略,实现从数据到知识、知识到决策、决策到行动的快速转化。2010年12月,美国总统科技顾问委员会发布题为《数字未来设计:联合资助网络和信息技术的研究与发展》的报告,突出强调了数据爆炸式增长所带来的管控和利用困境,并建议美国政府优先发展大数据技术。^②2012年3月,美国政府提出《大数据研究和发发展倡议》(Big Data Research and Development Initiative),^③很快在国际上掀起了大数据的讨论热潮。

这股热潮源于大数据对于国家发展的巨大影响力及其在国际关系中的深刻含义。概括而言,大数据对国际关系的主要范畴产生如下重要的变革性意义:从经济方面看,其商业价值决定了围绕大数据的新国力竞争;从主权方面看,作为信息载体的大数据是数据主权的基础;从安全方面看,大数据挖掘所带来的预测性功能使其成为今后国家战略能力的重要支柱。

(一) 商业价值决定新国力竞争

大数据开发给世界经济带来了突破的曙光,许多经济学家称之为“大数据时代”的到来。首先,数据可能在不久的将来成为最大的交易商品。手中握有数据的公司基

① 徐子沛:《大数据》,桂林:广西师范大学出版社2013年版,第347页。

② *Designing a Digital Future: Federally Funded Research and Development in Networking and Information Technology*, Report to the President and Congress by the President's Council of Advisors on Science and Technology, December 2010, <http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-nitrd-report-2010.pdf>, 登录时间:2013年8月20日。

③ “Obama Administration Unveils ‘Big Data’ Initiative: Announces MYM200 Million in New R&D Investments,” Press Release of the Office of Science and Technology Policy and the Executive Office of the President, March 29, 2012, http://www.whitehouse.gov/sites/default/files/microsites/ostp/big_data_press_release_final_2.pdf, 登录时间:2013年8月20日。

于数据交易即可产生更好的效益。数据的交叉复用将使大数据变成一大产业,如同基础设施建设产业一样,大数据产业将由数据提供者、管理者、监管者等多方构成。数据被誉为“21世纪的石油”,新一轮“淘金潮”已经在大数据领域开始出现。^①

其次,随着数据挖掘技术的提高,大数据将与货币和黄金一样,成为新的经济资源和经济资产。越来越多的企业发现自己通过业务交易产生的“废弃数据”成了宝贵信息。例如,电子商务企业所掌握的用户基本数据和交易信息、物流企业所掌握的产品运输信息和路况信息等。对于新的知识经济而言,大数据是一种不断产生、取之不尽的资源。

再次,大数据将会催生以数据资产为核心的多种商业模式,如大数据内生型、外生型、寄生型、产品型、云计算服务型、混合型等价值模式。^② 定位角度不同的服务模式将出现,如帮企业做内部数据挖掘的侧重数据分析的服务,又如帮企业更精准找到用户、降低营销成本、提高销售率并增加利润的侧重优化的服务等。一个基于云计算-大数据开发-各专门门类应用的倒金字塔形的巨大产业链即将形成,一个新的推动世界经济发展的科技革命将会诞生,国际经济将进入一个新的发展阶段。国际数据公司(IDC)认为,随着世界进入大数据时代,只有那些利用大数据进行实时决策的机构才能繁荣昌盛,而不能适应和利用这一时代转变的机构则可能发现自己越来越处于竞争劣势和面临着消亡。^③

最后,大数据时代的经济和社会将发生融合。我们开始进入了一个“参与式市场社会(participatory market society)”。在参与式市场社会中,消费者同时又是生产者,一个新的词“消费生产者(prosumers)”开始出现。^④ 社会经济竞争方式也将因大数据发生变化,它不仅会使竞争过程发生改观、竞争环境有所不同,而且会促进更多的创新形式的产生。^⑤ 社会经济乃至政府的运行方式也因此不得不随之更新,“管理”将让位于“治理”。^⑥

最早提出大数据时代已经到来的机构是全球知名咨询公司麦肯锡。麦肯锡在其发布的研究报告中指出,数据已经渗透到每一个行业和业务职能领域,逐渐成为重要

^① Dirk Helbing, “Google as God? Opportunities and Risks of the Information Age,” April 11, 2013, p. 1, <http://arxiv.org/ftp/arxiv/papers/1304/1304.3271.pdf>, 登录时间:2013年8月20日。

^② 郭昕、孟晔编著:《大数据的力量》,第186-188页。

^③ Richard L. Villars, Carl W. Olofson and Matthew Eastwood, *Big Data: What It Is and Why You Should Care*, IDC White Paper, June 2011, p. 1.

^④ Dirk Helbing, “Google as God? Opportunities and Risks of the Information Age,” p. 5.

^⑤ Brad Brown, Michael Chui and James Manyika, “Are you Ready for the Era of ‘Big Data’?” *McKinsey Quarterly*, No. 4, 2011, p. 3.

^⑥ Max Wind-Cowie and Rohit Lekhi, *The Data Dividend*, London: Demos, 2012, p. 24.

的生产因素,而人们对于海量数据的运用将预示着新一波生产率增长和消费者盈余浪潮的到来。数据仓库、数据安全、数据分析、数据挖掘等围绕大数据的商业价值利用逐渐成为争相追捧的利润焦点,也成为各国国力发展的新的竞争领域。^①

(二) 信息载体奠定数据主权基础

大数据时代,各国之间的数据主权博弈将加剧。在大数据背景下,数据的概念已不同于信息,数据是信息的原生态,数据所涵盖的范围要远远超出信息的范畴。相比于信息主权,数据主权概念将更加被重视。^② 国家层面的竞争力将部分体现为一国拥有数据的规模、活性以及解释、运用数据的能力。可以预见,数据主权将成为继边防、海防、空防之后,另一个大国博弈的空间,^③各种类型的大数据将是这一博弈的核心。

大数据作为一种数据,具有所有数据的功能和特点。首先,大数据是一种信息载体,具有数据主权的保护价值。某种程度上,作为一种具有全息意义的信息载体,大数据甚至会成为国家基础设施的一部分。^④ 大数据蕴藏着最新科技、社会动态、市场变化、国家安全威胁征兆、战场态势和军事行动等各种政治、经济、文化、安全等信息。作为信息载体的大数据必将是各国主权的争取对象,数据主权也将成为国家整体主权必不可少的一部分。

其次,大数据将成为国家主权实施的权力基础。大数据时代,以地理疆域为基础的二维权力将被三维权力所取代。数据是信息的载体,而信息就是权力。数据成为社会权力的重要来源,也成为各国国力竞争的制高点。信息权力不仅是主权国家控制国内社会的基础,信息权力还是外交力量的倍增器。^⑤ 借助于国际互联网、大众传媒等信息生产、流动平台的辐射,美国霸权在世界范围实现了迅速的扩展。这种不对称的信息流动和数据流动所产生的吸引力和影响力,增强了美国的力量和影响,在美国全球战略中发挥了其硬实力不可替代的作用。如今,美国借助其各大网络数据公司的力量,再次站到了大数据创新的潮头浪尖。各国围绕大数据的权力争夺亦将展开。

最后,对大数据的控制是国家的秩序需求与国家主权的合法性需求。^⑥ 国家对数据都有天然的控制兴趣,并借此维持其权威、合法性以及主权。米歇尔·福柯(Michel

① James Manyika, et al., *Big Data: The Next Frontier for Innovation, Competition, and Productivity*, p.1-13.

② 蔡翠红:《云时代数据主权概念及其运用前景》,载《现代国际关系》,2013年第12期,第58-65页。

③ 刘琼:《数字主权将成大国博弈新空间》,载《中国电子报》,2012年7月17日第3版。

④ David Bollier, *The Promise and Peril of Big Data*, p. 1, http://www.aspeninstitute.org/sites/default/files/content/docs/pubs/The_Promise_and_Peril_of_Big_Data.pdf, 登录时间:2013年8月20日。

⑤ Joseph S. Nye, *Power in a Global Information Age: From Realism to Globalization*, New York: Routledge, 2004, p. 75.

⑥ 蔡翠红:《云时代数据主权概念及其运用前景》,载《现代国际关系》,2013年第12期,第61页。

Foucault)晚年所提的政府性(governmentality)^①概念似乎能够解释国家利用各种策略扩展其权力、控制力和影响力的自然规律。而其中重要的策略之一就是通过对国内各种主体和地域的信息与知识的掌握从而维持其权威。同时,通过对先前没有意识到的、潜在活动的信息的挖掘,国家有将其权力向新领域延展的天性。^②因此,随着应用范围的拓展,大数据将成为国家维护其主权合法性的重要支柱,数据主权也必将成为各国的博弈对象。

(三) 预测性昭示国家战略能力

大数据不仅可用于与公共利益相关的一些社会决策,例如公共卫生、教育等,也可以应用于政治活动,例如调查投票意向等。更重要的是,大数据的功能将从“描述性”向“说明性”乃至“预测性”转变,^③大数据的核心就是预测。^④数据挖掘是指通过特定的算法对大量的数据进行分析,从而揭示数据当中隐藏的历史规律和未来的发展趋势,为决策者提供参考。这一预测性功能对于国家安全特别有意义。海量数据蕴藏着最新科技、社会动态、市场变化、国家安全威胁征兆、战场态势和军事行动等重要情报信息。与国家安全领域相关的大数据方面的发展和必将影响到国家的综合实力和战略能力。^⑤

对于作为社会行为信息库的大数据进行的情报信息处理是加强国家安全的需要。大数据相关技术是在大量信息中寻找威胁国家安全的蛛丝马迹从而达到预警的必要手段。同时,通过对别国的网络空间的监控和大数据的挖掘分析,还可以了解别国社会民情,甚至了解其军事设施和情报,从而对另一国的国家安全形成制衡。

大数据的挖掘和利用还有助于强化对战场大规模态势数据的感知、获取和分析能力,将大规模态势数据转作战需要的情报,为快速形成作战方案提供有效支撑,从而提高作战行动能力,获取战场优势。因此,美军将大数据研发重点同国防发展与国家安全的长远目标、战略需求与现实需求结合起来,并明确将发展数据分析技术和提高分析人员能力作为大数据发展的核心目标。美国国防部在大数据上压下了巨大赌注。2011年4月,美国国防部长签署了一份有关国防部优先发展的科学技术的备忘录,^⑥

① 政府性(governmentality)一词首次出现在1978年3月15日福柯的《安全、领土与人口》课程演讲中,也有“统治性、统治心态、统治术、治理性、治理术”等译法,该词是福柯后期研究中的重要概念。

② Benjamin Forest, "Information Sovereignty and GIS: The Evolution of 'Communities of Interest' in Political Redistricting," *Political Geography*, Vol. 23, No. 3, 2004, p. 430.

③ Maude Bonenfant, Marc Ménard, André Mondoux and Maxime Ouellet, "Big Data and Governance," p. 3.

④ 维克托·迈尔-舍恩伯格、肯尼思·库克耶:《大数据时代》,第16页。

⑤ 汤珊红、许儒红、侯勤:《大数据:信息时代大国技术竞争新领域》,载《国防》,2013年第2期,第73-76页。

⑥ "Secretary of Defense Science and Technology Priority Memo," April 19, 2011, <http://www.acq.osd.mil/chieftechologist/publications/docs/OSD%2002073-11.pdf>, 登录时间:2014年2月16日。

其中有七项科学技术被作为优先的战略投入目标,“数据到决策(Data to Decision)”排在首位。随后美国国防部明确了美军在大数据处理分析方面详细的发展路线图,并在国防部范围内进行整体技术架构设计。^①2012年3月,美国《福布斯》杂志网站发表了题为《重新定义军事情报:战场上的大数据》的文章,^②指出美国的国家与全球安全越来越依赖对快速增长的情报数据的获取、组织和管理能力。

与军方需要全面发展大数据的获取、处理、分析、识别等技术不同,非军方政府部门的大数据计划如果能对本部门长期积累的数据进行处理和分析,则可促进知识创新和科学发现,进而从侧面加强国家的安全能力。比如,美国能源部为响应《大数据研究和发展计划》,将斥巨资建立可扩展数据管理、分析与可视化研究所(The Scalable Data Management, Analysis and Visualization Institute),综合六大国家实验室和七所著名大学的专业知识和经验来开发新工具,以帮助科学家对能源部超级计算机上的数据进行有效管理和可视化处理,这可能对美国未来核武器的发展产生影响。美国国家航空航天局(NASA)则关注在建立未来地球观测的数据信息系统时,如何利用大数据技术降低相关风险和成本,这有可能进一步增强美国的航天能力。因此,无论对于军方还是非军方,大数据的合理应用都将会加强美国国家安全和战略能力。

三 大数据时代国际关系的新挑战

尽管是否会出现以信息和数据为基础的新的世界秩序还不得而知,但是不可否认的是,大数据将成为21世纪的国际关系的新挑战。新的数据鸿沟、数据霸权和数据的跨国安全威胁将对国际关系的平等性、民主性与合作性诉求形成新的压力。

(一) 平等性诉求与新的数据鸿沟

一个公正合理的国际政治经济新秩序是各国人民追求的目标。国际关系的平等性、民主性、合作性则是这一新秩序的基础。

^① “数据到决策”计划的研发内容主要包括:一是以创新方式使用海量数据,通过感知、认知和决策支持的结合,建立真正的、能够独立完成操控并做出决策的自主式系统;二是提高作战人员和分析人员的环境与状况感知能力,增强对任务和流程的支持。美国国防部具体通过国防高级研究计划局的“先期技术与工具开发”计划和“数据扩展”计划以及海、陆、空等方面科研部门的大力协助,开展“数据到决策”计划。其中“先期技术与工具开发”目前在研项目主要有八项,涉及数据处理、数据分析、数据监测、数据安全等方面的具体技术。“数据扩展”旨在开发能够分析海量半结构化数据和非结构化数据的计算技术和软件工具。

^② Chris Young, “Military Intelligence Redefined: Big Data in the Battlefield,” *Forbes*, March 12, 2012, <http://www.forbes.com/sites/techonomy/2012/03/12/military-intelligence-redefined-big-data-in-the-battlefield/>, 登录时间:2014年2月16日。

然而,大数据时代的国际关系平等性诉求面临着巨大挑战。大数据富有者和大数据贫穷者之间存在数字鸿沟(digital divide),^①不具备数据认知能力的大部分主体和少部分数据精英之间也存在数据鸿沟(data divide)。^②这两种关系既可以在一国社会内部产生,如个人用户和企业用户之间,也可以在国家之间产生,并成为国际关系平等性诉求的障碍之一。目前,全球各地区信息产业的分布情况已经导致了地区之间的信息不对称和地区之间的数据鸿沟。而且,数据是被创建的,其本身就嵌入了各种国内和国际特权。有社会特权的主体可能通过数据创建加强其特权,而那些不愿意参加数据创建活动的主体则被边缘化。大数据的开放性和其中隐含的监控功能使个人自觉按照所谓的规范进行自我约束,从而加强这种权力结构中本身存在的不公平。^③因此,数据可能使强者更强。^④

大数据事实上并不对所有主体同等公开,即使公开,很多主体也不具备分析能力或具有不同的分析能力。列夫·蒙纳维奇(Lev Manovich)指出了大数据领域中的三类主体:^⑤第一类是创建数据者(如有意或无意地留下电子脚印的人),第二类是有办法收集数据者,第三类是有能力分析数据者。“大数据热”主要基于对庞大数据的可利用性,因此后两类主体的分化是数据鸿沟产生的原因。而最后一类数量最少也最拥有特权,他们也是决定大数据使用规则和参与规则的主角。^⑥

以社交媒体信息为例,显然,目前只有那些社交媒体公司对于庞大的社交数据才有真正的获取权,如为脸谱工作的人类学家或者为谷歌工作的社会学家可以获取这些数据,但是其他的学者却没有这样的权利。一些公司的研究人员甚至认为学术界没有必要研究社交媒体,因为他们自己的研究人员可以做得更好。^⑦一些公司严格限制对其数据的利用,也有公司高价出售这种利用权,还有些公司对高校研究机构开放少量数据。这就会产生不平等,包括有钱购买使用权者和无钱者之间的不平等,公司内部

① Danah Boyd and Kate Crawford, "Critical Questions for Big Data," p. 674.

② Peter G. West, "The Data Divide: A Society Divided by Open Data," http://www.peter-west.co.uk/media/pdf/original/irp_poster.pdf, 登录时间:2014年2月16日。

③ Jeffery Alan Johnson, "From Open Data to Information Justice," Presented at the Annual Conference of the Midwest Political Science Association, Chicago, April 13, 2013, pp. 1-20.

④ Michael Gurstein, "Open Data: Empowering the Empowered or Effective Data Use for Everyone?" *First Monday*, Vol. 16, No. 2, 2011, p. 1.

⑤ Lev Manovich, "Trending: The Promises and the Challenges of Big Social Data," in Matthew K. Gold, ed., *Debates in the Digital Humanities*, Minneapolis: The University of Minnesota Press, 2011.

⑥ Danah Boyd and Kate Crawford, "Critical Questions for Big Data," p. 675.

⑦ 2011年7月19日在巴塞罗那举行的一次关于博客和社交媒体的国际会议(ICWSM)上,推特公司的专职研究人员林志颖(Jimmy Lin)在主题演讲中指出,推特自己的研究人员拥有对其数据的优先获取权,因而相关研究也更有优势,外部研究人员不必去探究。

和公司外部的不平等,甚至不同国家之间的不平等。

各国能够应对大数据挑战的能力不尽相同,新的数据鸿沟即将出现。按照克雷格·蒙迪(Craig Mundie)的分类法,国家政府在大数据浪潮中有三种类型:已经能够驾驭这一浪潮的、永远没有能力驾驭的以及有学习能力并有可能驾驭的。^①

(二)民主性诉求与数据霸权问题

国际关系民主化有利于推动建立与时代发展相适应的新型国际关系,而霸权主义则与国际关系民主化针锋相对。冷战后,以对抗-力量均势为主要特征的国际关系体系正在向以合作-利益均衡为主要特征的国际关系体系转变。国际关系民主化的主旨就是各国的事情要由各国人民做主,国际上的事情要由各国平等协商,对于事关世界和地区和平的重大问题,应该按照《联合国宪章》的宗旨和原则以及公认的国际关系基本准则,通过协商谈判的方式和平解决,全球性的挑战要由各国合作应对。^②

如果说数据鸿沟可能是由客观因素造成的全局状态,那么数据霸权则包含了体系强国更多的主观意图。大数据是继超级计算机和互联网之后美国信息技术发展的新领域。除美国外,欧洲主要国家、日本、俄罗斯等国对海量数据的处理也非常重视,但还没有上升到国家战略层面。通过《大数据研究和发展倡议》的推出,美国再一次走在世界其他国家之前,并试图占据大数据技术的主导地位,保持和扩大在大数据技术领域的持续优势。

尽管美国在一些网络指数如渗透度(penetration)排名上并不很靠前,但是有一个指数却遥遥领先,即和大数据战略息息相关的互联网2.0(Web2.0)产业的发展。美国的互联网2.0产业非常强健,几乎在全球都被广泛使用。如谷歌、脸谱等公司的存在,使美国几乎可以获得全球任何地方的数据。^③美国开发的三个搜索引擎:谷歌、雅虎(Yahoo)、必应(Bing),不仅主宰了美国的国内市场,也是全球最受欢迎的搜索引擎。^④也就是说,数据在全球范围内产生,却大部分被美国公司所收集。

国际数据公司于2012年年底首次进行了一项“数字地球分布”统计。通过对数据和信息的产生、捕获或者利用的地点进行分析,国际数据公司描绘了一个地理分布图。根据2012年12月的分析结果,美国在数字地球中占据32%,西欧诸国共同占

① Clifton Martin and Laura Jagla, *Integrating Diplomacy and Social Media*, p. 14.

② 《江泽民文选》第三卷,北京:人民出版社2006年版,第474页。

③ Stephane Grumbach, "The Stakes of Big Data in the IT Industry: China as the Next Global Challenger?" The 18th International Euro-Asia Research Conference, Venezia, January 31-February 1, 2013, p. 5.

④ Stephane Grumbach, "The Stakes of Big Data in the IT Industry: China as the Next Global Challenger?" pp. 6-7.

32% ,中国占 13% ,印度占 4% ,其余国家占 19% 。虽然总体趋势是不断从美国、西欧等发达世界向新兴市场国家转移,但根据国际数据公司的预测,中国 2020 年所占比例也仅能达到数字地球 21% 的份额。^① 所以,数字地球短期内还是西方发达国家的霸权阵地。

如果考察全球前 50 名网址的各国占比情况,则有如下结果:美国占了 2/3 以上,中国虽仅次于美国,但大约只占 16% 。其他国家所占比例更低,如俄罗斯占 6% ,英国占 2% (参见表 1)。^②

表 1 全球网址前 50 名各国占比^③

国家	所占比例	部分网址举例及排名
美国	72%	Facebook(1), Google(2), YouTube(3), Yahoo(4)
中国	16%	Baidu(5), QQ(8), Taobao(13), Sina(17), 163.com(28), Soso(29), Sina Weibo(31), Sohu(43)
俄罗斯	6%	Yandex(21), Kontakte(30), Mail.ru(33)
以色列	2%	Babylon(22)
英国	2%	BBC(46)
荷兰	2%	AVG(47)

大数据和云计算的结合某种程度上更推进了数据霸权。通过将大量的电子数据存储在少数云端,云计算推动了一种新的权威结构的变化。正如工业革命将工人与生产方式逐步分离一样,云计算使硬件、软件、内容和数据等网络生产越来越集中到少数互联网服务商手里。^④ 大数据最重要的一些服务由少数几个公司所提供,而且这些公司多数在美国。这些服务使其收集数据非常便利。

“知识就是力量”在大数据时代似乎可以解读成“全知就是全能”。谁能收集更多、更全的全球数据谁就可能更强大,特别是那些同时拥有分析工具的对象,如美国国

① John Gantz and David Reinsel, “The Digital Universe in 2020: Big Data, Bigger Digital Shadows, and Biggest Growth in Far East,” *IDC iView*, Vol. 3, No. 1414, December 2012, p. 6, <http://idcdocserv.com/1414>, 登录时间:2014年2月16日。

② Stephane Grumbach, “The Stakes of Big Data in the IT Industry: China as the Next Global Challenger?” p. 6.

③ Stephane Grumbach, “The Stakes of Big Data in the IT Industry: China as the Next Global Challenger?” p. 6.

④ Primavera De Filippi and Smari McCarthy, “Cloud Computing: Centralization and Data Sovereignty,” *European Journal of Law and Technology*, Vol. 3, No. 2, 2012, pp. 1-2.

家安全局(NSA)。^①因为在知道了所有信息后,总能找到阿基里斯之踵,这也正是全球对“棱镜门”事件反应如此强烈的原因。

(三) 合作性诉求与数据的跨国安全威胁

在全球化时代,合作将是未来国际关系的主流趋势。大数据的价值在于其数据的全面性和大量性,因此各国之间如果能够在大数据方面进行合作则可能产生更大的经济和社会效益。然而,由于国家间利益的竞争性乃至冲突性,各国的数据政策不同,数据在跨境流通中面临着许多安全威胁。

“跨边界数据流通(transborder data flow)”是网络安全和信息安全的重要威胁,其对于国家主权和安全的影响在20世纪70年代末就引起了各国政府的注意。1978年召开的由78个国家的代表团参加的政府间信息局国际会议就曾发表报告,认为跨边界数据流通“将国家置于危险境地”。^②而美国的“棱镜”计划在全世界所激起的安全恐慌表明,跨边界数据流通问题在云计算和大数据的背景下更加突出。

大数据与云计算几乎是孪生姐妹,云计算的出现将数据存储和数据分析变成了一个可以更加方便获得的网络服务。然而云时代的数据跨国安全威胁更为明显。首先,云计算本身的特点决定了数据的跨国安全威胁。虽然很多国家可能对一些重要数据的境外存储和跨国传输都有规定,然而云计算却使这些规定难以在技术上执行。在云计算中,数据存取都是在网络在线状态,用户无法控制自己调取数据的方式,无法保证第三方是否滥用该数据。云的物理位置的分散性和国际性使安全问题更加严重。为了满足用户的不同需求,云服务提供商可能基于技术限制、网络效率、法律因素、经济成本(如硬件成本、电价水平、税收因素等)的考虑,利用不同地点的云来完成任。③数据可能会在不同国家的云之间进行传输,数据存储的地点往往也难以确定。不仅国家难以知晓数据在被跨境传输,就连信息数据传输者本人可能也并不知晓,云端数据泄露问题普遍存在。云计算的发展使国内数据(domestic data)的定义越来越模糊。这可能严重威胁个体的隐私权、企业的商业秘密以及国家的主权和安全。

其次,存储在不同国家的云数据的处理和传输受不同国家的法律管制,而各国对于数据往往有不同的政策,这是国家安全所面临的重要威胁之一。网络数据的传输无法预知传送路径,同一起点和终点的数据可能经过不同的中间路线。除非数据在传输途径中的任何流经国家有一致的数据保护政策,否则数据就面临各种安全威胁。然

① Dirk Helbing, “Google as God? Opportunities and Risks of the Information Age,” p. 3.

② 任孟山:《全球化背景下的信息传播与国家主权》,载《中国传媒报告》,2007年第3期,第56页。

③ Primavera De Filippi and Smari McCarthy, “Cloud Computing: Centralization and Data Sovereignty,” p. 14.

而,各国的数据政策不尽相同,对数据的保护力度也有所区别。为此,欧盟曾规定过《数据保护指令》(Data Protection Directive),规定个人数据不得传输至欧盟以外的国家,除非它们有足够的保护水平。^①而许多欧洲国家的数据保护法律也规定所有数据都不得传输至国外,除非数据流经的国家都采取了和数据输出国同等的保护法律。^②

如果说上述因素属于数据跨境流通的被动安全风险,那么大数据还带来了另一种主动安全风险,大数据的巨大潜在价值刺激了跨国数据获取的热潮。有些国家甚至立法授权跨国数据获取,其执法、安全等公共部门拥有可以根据本国的法律获取数据的权利。事实上,大数据时代国家安全最大的挑战是来自第三方的有法律根据的数据获取。最典型的案例便是“棱镜门”事件。根据美国的《爱国者法案》(USA PATRIOT Act),政府可以获取任何存储在美国数据中心的信息或者是美国公司所存储的信息,这不仅不需要数据主体(data subject)的事先同意,有时数据主体可能根本没有意识到政府获取了信息。^③而美国《对外情报监控法案》同样赋予了美国政府对美国云服务提供商所保存或处理的数据的获取权利。^④这些法案正是“棱镜门”事件发生的法律依据。所以,越来越多的外国公司和政府机构在将数据传输到云端时都更加迟疑,同时,越来越多的国家开始意识到与美国云服务提供商合作的安全风险,因为它们担心美国的云服务提供商一旦获得数据或者数据传输经过美国,美国政府可能依据法律获取数据。因此,正如德国总理默克尔支持有关让欧洲建立与美国隔离的互联网服务的呼吁一样,^⑤各国都可能在不久的将来致力于自身的大数据应用和互联网服务系统的研发,尽管这并不能完全阻止数据的跨国安全威胁。

四 大数据时代的战略布局建议

无论我们欢迎与否,大数据时代已经到来。现在的关键是如何应对。在部分以不

^① Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, http://europa.eu/legislation_summaries/information_society/data_protection/114012_en.htm, 登录时间:2014年2月16日。

^② 例如英国就有相关数据保护法律(Data Protection Act: United Kingdom Law), <http://www.dpalaw.info/uklaw.htm>, 登录时间:2014年2月16日。

^③ Primavera De Filippi and Smari McCarthy, "Cloud Computing: Centralization and Data Sovereignty," p. 15.

^④ Judith Rauhofer, "Protecting Their Own: Fundamental Rights Implications for EU Data Sovereignty in the Cloud," *University of Edinburgh School of Law Research Paper Series*, No. 28, 2013, pp. 1-29.

^⑤ "German Chancellor Proposes a Europe-Only Network," February 18, 2014, <http://www.npr.org/2014/02/18/279216370/german-chancellor-proposes-a-europe-only-network>, 登录时间:2014年3月20日。

确定性和快速变化为特征的时代背景中,当前的战略布局对于未来的发展至关重要。对于大数据的挑战,政府别无选择,只有积极应对。科技乃立国之本,各国以新一代信息产业为核心的新一轮产业革命正在兴起。因此,中国要正视世界正在经历的大数据时代,把握新技术革命非线性突破的特征,^①审时度势,抓住机会,尽早进行应对大数据时代的战略布局。具体而言,本文有如下几点建议:

(一) 政府顶层设计与行业示范效应

政府力量对大数据行业的推动将起到示范效应。目前,美国、日本等国已初步启动了大数据的顶层设计,而中国政府目前尚未直接对“大数据”这一较新的概念给予政策支持。2012年3月,美国政府投资2亿美元启动《大数据研究和发展计划》。2013年5月,日本政府启动新ICT(Information、Communication、Technology)战略研究,重点关注大数据应用所需的云技术、传感器、社会化媒体等智能技术的开发。

出于对科技发展前景和成果的判断,美国奥巴马政府认识到大数据对于未来国家发展的重要意义,将其定义为“未来的新石油”,赋予了发展大数据非同一般的战略意义,将其升级为一项借助信息技术提升国家竞争力的宏观战略。通过发布《大数据研究和发展倡议》,美国政府把大数据研究上升为国家意志,这是大数据从商业行为上升为国家战略的分水岭,表明大数据正式被提升到战略层面,大数据也因此在经济社会的各个层面都受到重视。

中国政府在大数据领域尚未形成顶层设计,没有提出相关战略指导文件,也还没有国家层面的部门主导对大数据的研究和发展工作。目前,中国的大数据发展还主要依赖商业行为的研发。不过,在2011年12月8日工业和信息化部发布的物联网“十二五”规划中,^②信息处理技术作为四项关键技术创新工程之一被提出来,其中包括海量数据存储、数据挖掘、图像视频智能分析等技术,这都是大数据发展的重要组成部分。而另外三项关键技术创新工程,包括信息感知技术、信息传输技术、信息安全技术,也都与大数据密切相关。

伴随着大数据时代的来临,中国不能落后于先发国家,要在新科学革命和新技术革命初见端倪的时候就抓住机遇,进行相应的战略规划与布局。拥有海量用户的中国门户及社交网站和移动信息终端的迅速增长,使中国成为最大的大数据市场之一。然而,中国数据中心市场呈现客户大量外流及国外数据中心强势进入的态势。特别是在

^① 许正中:《“大数据”时代下的战略突围构想——重新审视中国经济社会独特性》,载《人民论坛》,2012年第14期,第15页。

^② 《我国物联网“十二五”发展规划公布》,2011年12月8日,<http://mobile.people.com.cn/h/2011/1208/c227890-919477941.html>,登录时间:2014年2月16日。

大数据整体解决方案领域,甲骨文、EMC 等公司都在全力占领中国大数据和云计算服务市场。因此,在新一轮科技革命中中国需从顶层设计和战略性高度着力应对大数据挑战,推动大数据行业的科技发展,培养高水平相关人才,并不断开拓创新,引领大数据发展前沿。

(二)“一体化”建设与信息割裂问题应对

“一体化”建设是各国大数据发展的重点方向。美国的情报体系固然庞大,然而即使如此庞大的体系,也越来越难以及时处理浩如烟海的互联网信息以及无人机、侦察机等各种平台传送回来的海量信息。因此,规范数据质量和标准、整合情报网络体系、提高情报信息综合处理能力的的需求日益迫切。美国国会早在2001年通过的《数据质量法》(Data Quality Act)可以算是朝“一体化”建设方向较早迈出的一步。^①

虽然中国政府在大数据上尚未形成顶层设计,但是有些部门和领域已经开始采取行动,国内互联网公司在大数据的商业应用方面也在向前推进,中国各地也制定或公布了云计算、物联网等产业规划。但是,中国如想在大数据领域占据先机,也需同时进行“一体化”建设,以形成合理配置的研发布局,并应对全国的信息割裂问题,比如各地区之间、各领域之间和各行业之间的信息割裂。

当下中国大数据的“一体化”建设可着力发展以下四个方面:一是设立一套规范的大数据建设标准。数据治理的重点是数据定义的一致性和数据质量。大数据时代,不同系统之间的数据进行整合的前提就是要有统一的元数据定义。这不仅是中国而且是全世界都面临的挑战。^②因此,应建立面向不同主题、覆盖各个领域、不断动态更新的大数据建设标准,为实现各级各类信息系统的网络互连、信息互通、资源共享奠定基础。二是建立一套运行机制。大数据建设是一项有序的、动态的、可持续发展的系统工程,必须建立良好的运行机制,以促进建设过程中各个环节的正规有序。三是搭建一个共享平台。数据增值的关键在于整合,而自由整合的前提是数据的开放和共享。应在各专用数据库建设的基础上,通过数据集成,实现各级各类信息系统的数据交换和数据共享。四是在全社会弘扬数据文化的同时培养一支专业队伍。大数据建设的每个环节都需要依靠专业人员完成,因此,必须培养和打造一支懂技术、懂管理的大数据建设专业队伍,才能指导全国的大数据“一体化”建设。^③

在“一体化”建设基础上,中国也可针对性地从一些易见效果的应用领域开始进

^① http://en.wikipedia.org/wiki/Data_Quality_Act, 登录时间:2014年2月16日。

^② 徐子沛:《大数据》,第350页。

^③ 彭默馨、张璐:《大数据时代要有大方略》,载《学习时报》,2012年7月9日第7版。

行相关大数据的研发。例如,在安防领域应用大数据技术提高应急处理能力和安全防范能力;在民生领域如医疗、卫生、教育等部门,应用大数据技术提升服务能力和运作效率;在金融、电信领域,提升有针对性的用户行为分析和信用评级水平并提供针对性服务等。

在上述统一标准和共享平台的基础上,中国还应从多个角度推进大数据的相关研发,综合公司企业、科研高校、政府部门的力量进行相关研究。由于潜在的巨大经济效益的驱动,美国的一些大型公司不仅赞助大数据相关项目的竞赛,并且为高等院校的大数据研究提供资金。正是这些企业行为的出现,大数据才受到美国政府的高度重视。因此,奥巴马也强调联邦政府必须动员全社会力量,和公司、大学等相互合作,来应对大数据挑战。^①

(三) 大数据治理与平衡关系度量

美国总统科学顾问约翰·侯德然(John Holdren)说,“大数据确实是个大问题”。^②每年新产生的海量电子数据来源多种多样,从地下物理实验数据到天空望远镜探测数据,再到零售交易数据、社交网站的发帖等,这些都需要我们去进行管理,也因此使“大数据治理(big data governance)”成为今后的一大重要挑战,也是中国在大数据战略规划中需要考虑的问题。

劳伦斯·莱斯格(Lawrence Lessig)在其著作《代码》中曾指出系统的四个规制因素,即市场、法律、社会规则、建构(architecture)(针对技术而言,建构也就是编码或者技术标准)。^③对于大数据而言,这四个规制因素也适用,它们决定了大数据治理需要平衡很多关系。

技术建构决定了数据共享与数据安全的关系。某些特殊行业的数据应用,比如金融数据、医疗信息以及政府情报等都有自己的安全标准和保密性需求。但大数据分析往往需要多类数据相互参考,而在过去没有这种数据混合访问的情况,因此,大数据应用催生出一些新的、需要考虑的数据共享中的安全性问题,如云端数据泄露等。

市场因素决定了数据价值与投入成本的平衡关系。市场认为大数据绝对是个机会:可以应用于目标广告设计和市场营销、产品设计、保险对象优化分析、市场变化分析等。但是市场也有其自身规律的约束,最主要的一条就是大数据的经济回报率,即

^① Michael Goldberg, “Obama Administration Big Data Initiative Calls All Hands on Deck,” April 25, 2012, <http://data-informed.com/obama-administration-big-data-initiative-calls-all-hands-on-deck/#sthash.xurnjQbu.dpuf>, 登录时间:2014年2月16。

^② Jeffrey Mervis, “Agencies Rally to Tackle Big Data,” *Science*, Vol. 336, No. 6077, 2012, p. 22.

^③ Danah Boyd and Kate Crawford, “Six Provocations for Big Data,” p. 2.

大数据的价值和投入成本的比率。考虑到大数据的庞大数量和超快速增长速度,搜索出有用数据确实非常困难。甚至有专家说,“美国自己碰到问题了,就抛出来让全球帮它解决,我们必须重视‘大数据’研究过程中的安全性”。^①因此,大数据治理需要遵循信息生命周期的规律,制定合理的存档、研发和利用策略,保持数据价值和投入成本的合理配比。

法律因素决定了公开性和隐私性的平衡关系。大数据的应用前提是数据在一定程度上的共享性和公开性,但是法律则对数据的收集和保留加以限制,要求大数据首先不能侵犯国家安全,其次不能侵犯个人隐私。大数据尤其对隐私权提出了严峻挑战,^②例如在很多情况下,一个特定用户可能希望自己的过往网络行为和相关数据被忘记,特别是社交媒体信息。但是在大数据时代,这一点似乎很难做到。又例如,个人在搜索引擎上的搜索内容也许能够部分反映其生活、兴趣、个性等并不想与人分享的信息。大数据的挖掘技术还使关于我们个人的信息远远超过我们自己创造的信息,这些信息就成为我们的“数字影子(digital shadow)”。^③随着各种云技术的发展,数字影子与日俱增,其中很多是我们本人所不知晓的,而且也是我们不愿公开的信息。所以立法机构应从隐私权等因素出发,立法限制数据的收集和保留。美国2011年就提出了《禁止追踪网上行为的法案》(Do Not Track Online Act of 2011)。还有学者提出了“信息审计”的概念,认为为了防止数据滥用,首先应该限制数据收集,比如公司等主体只应收集当下需要的数据。^④

社会规则决定了大数据的技术性应用与社会伦理的平衡关系。社会规则和社会个性化趋势对相关信息的获取提出了更高的技术要求,但同时产生了许多伦理难题,并将公众进一步分裂。^⑤而且,网上实时动态似乎使我们的世界从道德约束型变成了动态冲动型,也就是所谓的“网络驱动行为主义(cybernetic behaviourism)”。^⑥道德冷漠、人际情感疏远、网络不平等使用、窥秘心理等原因使网络伦理问题不断涌现,例如人肉搜索、隐私侵犯、人权自由等。大数据的滥用可能会加剧这些网络伦理问题。虽

① 双瑞:《“大数据”:信息时代新挑战》,载《经济参考报》,2012年7月12日第5版。

② Omer Tene and Jules Polonetsky, “Privacy in the Age of Big Data: A Time for Big Decision,” *Stanford Law Review Online*, Vol. 64, Issue 2, 2012, pp. 63-69, http://www.stanfordlawreview.org/sites/default/files/online/topics/64-SLR0-63_1.pdf, 登录时间:2014年2月16日。

③ John Gantz and David Reinsel, “Extracting Value from Chaos,” *IDC iView*, No. 1142, June 2011, pp. 9-10, <http://www.emc.com/collateral/analyst-reports/idc-extracting-value-from-chaos-ar.pdf>, 登录时间:2014年2月16日。

④ David Bollier, *The Promise and Peril of Big Data*, p. 36.

⑤ Eli Pariser, *The Filter Bubble: What the Internet Is Hiding from You*, New York: Penguin Press, 2011.

⑥ Maude Bonenfant, Marc Ménard, André Mondoux and Maxime Ouellet, “Big Data and Governance,” p. 7.

然大数据的应用能够为社会秩序和国家安全提供有意义的预测性建议,如防范恐怖主义行为,但其过分滥用也可能导致“假想敌情景(dystopian scenarios)”,即根据大量数据分析构建出事实上并不存在或尚没有发生的敌情或犯罪行为等。^① 过分依赖数据分析还可能会产生“数据真理妄想(apophenia)”,^②即在没有任何意义的一堆数据中可以找到任何方向的“规律”。因此,大数据治理还需要综合考虑其技术性应用和社会伦理的平衡关系。

五 结论

综上所述,大数据时代已经来临,大数据将成为21世纪国际关系的新挑战和新竞争领域。无论从经济、主权还是安全角度,大数据都在国际关系中占据了重要地位:其商业价值决定了围绕大数据的新国力竞争;其信息载体的本质奠定了数据主权的基础;其预测性功能昭示了对国家安全和战略能力的重要意义。同时,与大数据相伴随的新数据鸿沟、数据霸权以及数据跨国安全威胁等也对国际关系的平等性、民主性和合作性提出了严峻挑战。

美国目前占据着大数据领域的霸权地位。但是大数据也赋予了中国以千载难逢的机遇。在20世纪80年代和90年代的两次IT浪潮中,中国都处于被动跟随状态。但在当前以云计算为代表的第三次IT浪潮中,中国的大数据产业基本与全球处于一起跑线上。^③ 中国目前信息技术渗透率为37%,已经超过30%的世界平均水平并位于金砖国家之首。^④ 中国在传统的信息技术方面很难摆脱现有网络空间权力结构的布局。然而,在云计算的结构之下,中国有最大的手机用户、最大的互联网用户、最大的宽带用户和网络第二大语言,完全能支撑住以中国为主的技术标准。^⑤ 在这样的客观条件下,政府从国家层面对发展大数据予以扶持并制定专项规划,对中国的综合国力发展是非常重要的。因此,中国应该抓住当前这一时代赋予的机会,从政府顶层设计、“一体化”建设以及大数据治理等几个方面进行战略布局。

(截稿:2014年3月 实习编辑:冷鸿基)

① David Bollier, *The Promise and Peril of Big Data*, p. 33.

② Danah Boyd and Kate Crawford, “Critical Questions for Big Data,” p. 668.

③ 郭昕、孟晔编著:《大数据的力量》,第189页。

④ Stephane Grumbach, “The Stakes of Big Data in the IT Industry: China as the Next Global Challenger?” p. 2.

⑤ 田溯宁:《中国互联网“时势造英雄”》,载《互联网天地》,2011年第8期,第31页。

choring practices. The CICA and the ARF cases are illustrated to support the argument. In the end, the paper gives a brief analysis of the preconditions and limitations for lesser agents leading international institutions.

[Key Words] CICA, ARF, institutions, networking, process, practice

[Author] Wei Ling, Professor of International Studies, Director of the Institute of Asian Studies, China Foreign Affairs University.

Abstention or Veto: How Does Chinese Choose to Express Opposition in the Security Council

Qi Haixia Zhang Zuoli (101)

[Abstract] As a permanent member of the Security Council, China sometimes chose to abstain from voting, sometimes chose to veto it, to express her disapproval. To analyze what is the determinants of China's decision to choose abstention or veto, this research collects data of China's vote from 1971 to 2012. After statistical study of the related data, the author builds a regression model and finds that recently China tended to use veto to clearly express her disagreement while before 2000 China inclined to use abstention. Firstly, this paper tries to explain China's change by loss aversion. China would use veto when she thought East Asia was facing risk and China was facing losses if she did nothing. However, as to the Middle East issues, China may not think these issues belonged to her most important field. Secondly, we can also explain China's change by the effect of international socialization. After the interaction with other member states in U. N. and the criticism suffered when China chose abstention in 1990s, China chose to veto recently when she disagreed with the proposal. At last, we can find that together with China's rise, her confidence and responsibility can be observed in her vote in the U. N.

[Key Words] China, UN Security Council, abstention, veto, United Nations

[Authors] Qi Haixia, Lecturer, International Relations Department of Tsinghua University; Zhang Zuoli, Master Candidate, International Relations Department of Tsinghua University.

Connections and Challenges of Big Data in International Relations Cai Cuihong (124)

[Abstract] The era of Big Data has arrived, along with the popularity of smart devices and video equipment, the birth of cheap storage, high speed broadband and cloud computing, as well as the development of social networking and mobile trend. In the current digital environment, with a high degree of uncertainty and rapidly changing features, the effective and efficient use of Big Data has become a key to a country's information power. Big Data will thus become another important area of research in in-

ternational relations. Whether from an economic, sovereignty or security point of view, big data occupies an important position and may have revolutionary significance in international relations; its commercial value as the potential goods, resources and assets entails the national power competition around Big Data; its nature of information carrier lays the foundation of data sovereignty which is to become a target of the international political games; its underlying abundant information and corresponding predictability function implies its significance to national security and strategic capabilities. Meanwhile, the digital divide, the data hegemony and the transnational security threats associated with Big Data present challenges for the equality, democracy and cooperative pursuit in international relations. China should seize the Big Data opportunities and make strategic layout through the government top-level design, the nationwide coordination and integration as well as the Big Data governance.

[Key Words] big data, international relations, data sovereignty, digital divide, big data governance

[Author] Cai Cuihong, Ph. D., Associate Professor at the Center for American Studies of Fudan University.

The Governance of Global Cyber Space in a Post-Snowden World **Shen Yi (144)**

[Abstract] The asymmetric distribution of resources and capabilities among different actors is one of the most important characteristics of global cyberspace. The competition based on the capability to protect the data sovereignty has already become the front of the competition among great powers. The main goals of the competition are to ensure the national cybersecurity and shape the code of conduct in the global cyberspace. The main driving force of the competition comes from the system level which is the deep evolution of global productivity. It could be a better choice to understand the core task of the governance of global cyberspace in a post-Snowden world in such a framework of analysis, and it would be proper to shape a grand strategy in the same way. During that process, the governance system of global cyberspace which covers the governing of key infrastructure, the technologies and standards and the key resources also needs to be readjusted accordingly, while it will occupy the center position of world stage.

[Key Words] data sovereignty, global cyberspace, cyber security, governance

[Author] Shen Yi, Associate Professor of School of International Relations and Public Affairs, Deputy Director of the Center for Brics Studies, Fudan University.