大数据侦查的法律控制

程　雷

摘　要：大数据侦查通过计算机技术对存储于网络与计算机系统中的海量数据进行收集、共享、清洗、对比和挖掘，从而发现犯罪线索、证据信息或者犯罪嫌疑人。大数据侦查主要包括目标驱动型、比对驱动型与事件驱动型三种行为样态，在犯罪预防预测和犯罪侦破领域均有实践应用。大数据侦查对一些基本权利和法律价值构成挑战，有必要对其进行法律控制。然而，传统的法律规范框架存在滞后性，对大数据侦查的法律属性界定模糊，区分数据内容与元数据具有局限性，侦查启动门槛虚置，已然犯罪与未然犯罪界限模糊。对大数据侦查进行法律控制，可采取侦查规范和数据规范的双重路径。在侦查规范方面，应遵循合法性原则、比例原则，加强外部监督和司法监督。在数据规范方面，建议适度引入个人信息保护方面的法律原则和机制，包括确立目的合法与特定原则，赋予信息主体的知悉权与更正权，建立信息安全与数据质量控制机制，以及个人信息使用的监督与救济程序。

关键词：大数据　侦查　大数据侦查　个人信息　技术侦查措施

作者程雷，中国人民大学刑事法律科学研究中心副教授、未来法治研究院研究人员（北京　100872）。

当前，大数据技术已成为影响世界发展格局的大趋势，在自然科学、商业领域、政府管理等社会各个领域产生了直接且深远的影响。2015 年 10 月，党的十八届五中全会明确提出实施国家大数据战略，① 这一技术趋势成为中国社会发展的重要驱动力。社会发展形态的变迁直接决定着作为社会控制机制重要组成部分的犯罪侦查与预防工作，大数据技术在中外警务界的侦查活动中得到愈发广泛的应用。② 然而，

① 《中国共产党第十八届中央委员会第五次全体会议公报》，《中国共产党第十八届中央委员会第五次全体会议文件汇编》，北京：人民出版社，2015 年，第 7 页。

② 为行文方便，笔者将大数据技术在侦查中的各类应用方式统称为“大数据侦查”，用以概括通过计算机技术对数据库进行数据收集、共享、清洗、比对与挖掘，从而发现侦查信息的侦查措施与方法。国内较早使用“大数据侦查”用语的研究，参见王燃：《大数据侦查》，北京：清华大学出版社，2017 年，第 12 页。

在大数据侦查蓬勃发展的背景下，包括中国在内的全球多数国家的刑事司法规范与体系均表现出滞后性，面对陌生的新型技术手段，法律规制滞后于技术发展、法学理论落后于司法实践的现象愈发突出。何为大数据侦查，如何从法律上评价大数据侦查，如何对大数据侦查进行法律控制，这些现实问题亟待在梳理、提炼实践的基础上作出理论回应。

一、大数据侦查的实践类型

大数据是以容量大、类型多、存取速度快、应用价值高为主要特征的数据集合，正快速发展为对数量巨大、来源分散、格式多样的数据进行采集、储存和关联分析，从中发现新知识、创造新价值、提升新能力的新一代信息技术和服务业态。① 大数据技术包括对海量数据的采集、储存、清洗、分析、使用过程，其中最核心的技术为数据挖掘。数据挖掘（data mining），是指通过特定的计算机算法对大量的数据进行自动分析，从而揭示数据之间隐藏的关系、模式和趋势，为决策者提供新的知识。② 数据挖掘技术在大数据趋势成熟之前就有所应用，随着人类社会拥有和可使用数据量的激增，数据挖掘技术的应用广度与深度持续拓展，在大数据时代，数据挖掘技术的作用得到了最佳的展示机会，侦查实践亦成为其深度应用的场域之一。

（一）大数据侦查的行为样态

大数据侦查，是指通过计算机技术对存储于网络与计算机系统中的海量数据进行收集、共享、清洗、比对和挖掘，从而发现犯罪线索、证据信息或者犯罪嫌疑人的侦查措施与方法。其中数据查询、数据比对与数据挖掘是大数据侦查的三种具体行为样态。

刑事侦查工作的核心任务是发现犯罪嫌疑人、收集相关证据，以查明案件事实，③ 其中最为基础性与源头性的工作是锁定犯罪嫌疑人，否则以被追诉人为对象的刑事司法程序根本无从展开。锁定犯罪嫌疑人必然依赖于能够识别犯罪人的个别化信息，而这一任务恰恰是大数据技术的优势所在。大数据技术具备从海量数据中比对、挖掘、识别个人信息的强大功能，这能为侦查机关履行传统侦查职能提供高

① 国务院：《促进大数据发展行动纲要》（国发［2015］50号）2015年8月31日。

② 涂子沛：《大数据：正在到来的数据革命，以及它如何改变政府、商业与我们的生活》，桂林：广西师范大学出版社，2013年，第98页。

③ 《刑事诉讼法》第113条规定了侦查的任务，即公安机关对于已经立案的刑事案件，应当进行侦查，收集、调取犯罪嫌疑人有罪或者无罪、罪轻或者罪重的证据材料。对现行犯或者重大嫌疑分子可以依法先行拘留，对符合逮捕条件的犯罪嫌疑人，应当依法逮捕。

效、简便的智能辅助支持。另一方面，伴随着现代社会的加速发展与剧烈转型，犯罪愈发隐形化、智能化，① 加之恐怖主义犯罪在全球的泛滥，犯罪的严重后果令犯罪预防越来越优先于案发后的侦查与打击，将犯罪消灭在萌芽状态或者在犯罪现场即时破获案件成为了侦查隐形犯罪与恐怖主义犯罪的当务之急。因此大数据预测未来的功能在犯罪预防领域也变得愈发重要。

在刑事侦查中根据使用目的的不同，数据挖掘可分为三大类：一是目标驱动型数据挖掘（target-driven data mining），也称适用对象型数据挖掘（subject-based data mining），是指针对特定明确目标进行的搜索其所有记录以获取相关信息的过程；二是比对驱动型数据挖掘（match-driven data mining），此种模式的数据挖掘用于确认某人是否已经被作为“值得关注的人”，即用于确认某人是否为已知的犯罪嫌疑人；三是事件驱动型数据挖掘（event-driven data mining），也称为模式型监控，此种数据挖掘方法并非起始于具体明确的犯罪嫌疑人，而是用于发现过去或者未来事件的违法行为人。② 事件驱动或者模式驱动型数据挖掘用于搜寻反常的或者事先确定的行为模式或关系模式。③ 在各国的刑事侦查实践中，前两类数据挖掘方法早已有之并在侦查实践中得到了广泛的应用，比如查询已知犯罪嫌疑人的全部信息，通过现场遗留的指纹比对出谁是犯罪嫌疑人等，④ 而大数据时代，数据查询与比对方法的革新之处只不过是可供查询或比对的信息库容量发生了巨幅增加，但查询与比对的方式、方法并未发生根本改变。事件驱动型的数据挖掘对于犯罪预防与实时打击意义重大，大数据技术通过对过去一定时期内的犯罪数据进行挖掘后对犯罪热点地区、犯罪人群、犯罪手法等犯罪趋势提出的科学预测，将犯罪预防与防控建立在大数据预测的犯罪规律基础之上，从而能够更为精确、科学地调动警力并实现对犯罪的精确打击。

（二）大数据侦查的应用类型

大数据技术在刑事侦查中的应用前提是收集海量数据并形成各类数据库，这是

---

① 隐形犯罪（invisible offences）最早由美国学者Mark Moore提出，用来指代那些犯罪消息知悉困难、发现犯罪消息渠道不畅的犯罪类型，比如毒品、非法武器交易、贿赂犯罪等。参见Mark H. Moore, “Invisible Offenses: A Challenge to Minimally Instrusive Law Enforcement,” in Gerald M. Caplan, ed., *ABSCAM Ethics: Moral Issues and Deception in Law Enforcement*, Cambridge: Ballinger Pub. Co., 1983, p. 21;程雷：《秘密侦查比较研究》，北京：中国人民公安大学出版社，2008年，第65—75页。

② Christopher Slobogin, “Government Data Mining and the Fourth Amendment,” *The University of Chicago Law Review*, vol. 75, no.1, 2008, pp.322-323.

③ Fred H. Cate, “Government Data Mining: The Need for a Legal Framework,” *Harvard Civil Rights-Civil Liberties Law Review*, vol.43, no.2, 2008, pp.438-439.

④ Fred H. Cate, “Government Data Mining: The Need for a Legal Framework,” pp. 438-439.

数据查询、比对与挖掘技术应用的基础。从我国公安机关使用数据库查询、比对、分析各种记录的发展状况来看，自1998年公安部启动“金盾工程”开始，各类数据库建设与应用就逐步成为重要的侦查手段，2008年起公安部进一步开始了公安大情报系统的建设，当前公安机关内网联网运行的各类信息系统已达7000多个，已建成以全国人口信息库为代表的八大全国公安基础信息库（全国重大案件、在逃人员、出所人员、违法人员、盗抢汽车、未名尸体、失踪人员、杀人案件），存储了数百亿条基础数据。① 此外公安机关还积极运用各类社会管理中建设的数据库，利用各类信息资源开展侦查，包括互联网信息资源、视频监控信息资源、通讯信息资源、银行卡信息资源、各类社会服务中的信息，如保险、民航、工商、税务、邮政、社保、劳务、房产、公路、出租车、二手车交易、物流、出版印刷、房屋交易等。② 上述各类数据库中的海量记录涵盖了信息社会中人们生活、工作、社交等方方面面的信息，对这些大数据进行比对、分析已经成为当前侦查实践中提升破案率的主要驱动力。③

大数据技术在我国侦查实践中的应用方向，既有针对已经发生的刑事案件的回溯性侦查，用以锁定犯罪嫌疑人或查明案件事实，也有防患于未然式的对未来犯罪的预测与预警。在犯罪预测方面，北京市公安局“犯罪数据分析和趋势预测系统”是大数据技术应用的典型例证：2014年5月北京市公安局怀柔分局的上述犯罪预警系统预测提示，近期泉河派出所辖区北斜街发生盗窃案的可能性较高。怀柔公安情报信息中心根据提示，指导泉河派出所对该区域加大巡逻防控，5月7日1时许，泉河派出所巡逻至北斜街南口时，当场抓获一名盗窃汽车内财物的犯罪嫌疑人，经讯问，犯罪嫌疑人李某交代了伙同他人流窜至怀柔区，撬机动车锁并盗窃车内财物作案3起的犯罪事实。④

在针对已然犯罪的刑事侦查过程中，大数据侦查在司法实践中的作用逐渐显现。通过对中国裁判文书网2016年度刑事案件法院裁判文书的检索、查阅，可以得出部

① 参见艾明：《新型监控侦查措施法律规制研究》，北京：法律出版社，2013年，第169—170页。

② 艾明：《新型监控侦查措施法律规制研究》，第171—172页。

③ 记录查询与数据库侦查对于破案率的贡献并无官方统计数据，艾明在G省开展的针对93例个案的小样本实证研究显示，超过三分之二的案件中侦查机关使用了记录监控类的侦查手段进而破获了相应的案件。（参见艾明：《新型监控侦查措施法律规制研究》，第175—179页）也有学者认为，在近年来犯罪形势逐年恶化、诱发犯罪的社会条件逐步加强的背景下，杀人、抢劫等重大恶性案件反而逐年下降，背后的原因恰恰就是公安机关充分利用信息平台开展数据库侦查。参见江涌：《数据库扫描侦查及其制度建构》，《中国人民公安大学学报》2013年第2期。

④ 金江军、郭英楼：《智慧城市：大数据、互联网时代的城市治理》，北京：电子工业出版社，2016年，第112页。

分定量分析结论。① 在检索到的 570 件明确表明适用过技术侦查措施的刑事案件中，使用大数据技术锁定犯罪嫌疑人的案件为 113 件，涵盖的罪名根据出现频率的高低排序依次为盗窃（79 件）、抢劫（13 件）、抢夺（6 件）、交通肇事（5 件）、故意杀人（4 件）以及故意伤害、绑架、非法制造买卖枪支弹药爆炸物、信用卡诈骗、诈骗、强奸案各 1 件。

总结上述 113 件个案中大数据技术的应用情况，可以发现如下应用趋势。

首先，大数据侦查的应用对象主要为作案工具或作案对象为摩托车、电动车或汽车等车辆或手机的侵财类案件及相关刑事案件，这些案件中的犯罪嫌疑人在作案中通常会产生公共场所的视频监控与手机移动轨迹的数据，两类以上的数据库信息为数据比对提供了条件。大数据技术适用的案件范围是基于适用案件的客观情状经由侦查人员自发选择加以适用的，并未受到案件严重与否、罪名范围等条件的限制，恰恰相反，样本案件显示多数案件都是轻微犯罪，适用大数据侦查主要是基于侦查便利考量。

其次，大数据侦查的主要目的是发现并锁定犯罪嫌疑人，在上述 113 件样本案件中绝大多数案件都是陌生人之间发生的偶发性流动犯罪，基本上无法通过犯罪现场提取有效的痕迹物证，受害人基本上无从指认出相应的犯罪嫌疑人，因此锁定犯罪嫌疑人就成为了上述案件中侦破犯罪的基本前提，也是最为关键的侦查步骤。数据来源主要集中于公共场所的视频监控与手机通讯的基站数据这两类数据库，数据比对与挖掘的内容与对象较为单一。

最后，大数据技术尽管在锁定犯罪嫌疑人这一过程中发挥了重大作用，但在完成侦查工作的第二项重要任务即收集证据材料方面作用十分有限。在 570 件适用技术侦查措施的裁判文书中，虽然有 52 例案件中技术侦查措施用作了诉讼证据，但上述 113 件适用大数据技术的案件均未涵括在内。换言之，大数据技术锁定犯罪嫌疑人过程中的相关材料与信息在诉讼过程中基本上无从发挥证明作用。大数据技术的应用结果在案件材料与诉讼过程中至多作为“抓获经过”、“到案经过”、“破获经过”等辅助性说明材料出现。由于这些说明性材料在内容上的模糊与缩略，② 一方面对于锁定犯罪嫌疑人的方式并未进行详尽、如实的说明，导致大数据技术的应用过程

① 检索对象为中国裁判文书网（wenshu.court.gov.cn），访问与检索时间为 2017 年 5 月 5 日。检索范围为选择“刑事案件”，检索之日共有刑事案件的裁判文书 1461530 件，检索关键词为“技术侦查”，年份选择“2016”。

② 司法实践中此类关于犯罪嫌疑人、被告人如何到案或者如何被抓捕归案的过程描述性材料几乎存在于所有刑事案卷当中，但关于其证据属性与证明作用有无，理论界与实务界长期聚讼不一，相关讨论可参见陈为明：《〈案发经过〉不应当作证据使用》，《中国刑事法杂志》2004 年第 4 期；李继华：《浅谈“抓获经过”》，《公安研究》2000 年第 1 期。

被极大地忽略；另一方面，也导致这些情况说明材料不属于法定的证据种类，不是证据，只能作为加强法官内心确信的辅助材料使用。①

（三）大数据侦查的实践特征

与传统侦查行为相比，大数据侦查在实践运行中呈现出以下四项特征，恰恰是这些独特属性显示出对其进行法律规制的极大必要性，也从根本上影响着相应的规范工具与立场。

首先，大数据侦查具有权利干预的普遍性与深刻性。大数据侦查通过大数据技术对海量存储信息加以充分挖掘利用，对公民个人信息乃至隐私权的干预都具有史无先例的广泛性与深刻性，公民对于大数据侦查中侦查机关收集与使用公民个人信息的过程，既不知情亦无法抗拒。

其次，大数据侦查的出现改变了侦查权的权力分布格局，侦查权逐步社会化与弥散化。大数据侦查改变了传统的侦查参与主体结构，由于大数据主要是掌握在社会机构、商业机构手中，在大数据侦查过程中，侦查机关对社会机构、商业企业机构收集的公民个人信息进行数据比对与挖掘，形成了国家—社会—个人三方参与的新型侦查主体分布模式，社会力量而非侦查机关在侦查权行使过程中的作用愈发重要。

再次，大数据侦查在应用时间节点上呈现出前瞻性与主动性。此类侦查行为主要发生在立案之前发现犯罪嫌疑的早期阶段，具有典型的“无中生有”的特点，这与传统侦查行为系针对具体的犯罪嫌疑进行的回应性侦查模式明显不同。

最后，大数据侦查的实现过程具有智能化、低风险性和常规化趋势。大数据侦查主要依赖于大数据挖掘与比对技术通过计算机自动进行，机器学习、人工智能的应用使得发现犯罪线索的工作过程逐步实现了由机器替代人工，极大提高了识别特定目标与特定事项的效率，降低了侦查过程中侦查人员人身安全的风险，正逐步成为逢案必用的常规化侦查手段。

二、大数据侦查法律控制的必要性

大数据侦查是侦查机关顺应信息社会发展潮流的明智选择，其深度应用既有助于提高犯罪预防的精确性，提升警力配置效率，也有助于增强侦查取证的科学性，提高破案效率与破案能力，带来用信息换安全的社会效果。从权利保障的角度看，大数据技术的应用将无所不在的记录与数据经过分析、挖掘得出更为客观、

① 江必新主编：《〈最高人民法院关于适用《中华人民共和国刑事诉讼法》的解释〉理解与适用》，北京：中国法制出版社，2013年，第124页。

精确的证明犯罪过程的材料，客观上有助于改变长期以来侦查机关对口供的严重依赖，可以降低对严重干预公民隐私权的技术侦查的依赖，带来用信息换权利的法律效果。对于大数据侦查带来的侦查效能提升与侦查模式转型的积极效果，应当充分肯定。

然而，大数据侦查尚处于初始应用阶段，其双刃剑效应亦同时凸显，执法司法实践中已然暴露出一些问题，对一些基本权利和法律价值形成挑战，及时对其进行法律控制具有必要性。

首先，大数据侦查的推广适用标志着隐私逐渐受到限制，甚至有消亡的危险。边沁在1787年设计的圆形监狱概念（panopticon），在大数据时代应验成真。① 在大数据时代，公民个人的所有活动实现了全部数据化存储，生活、学习、人际交往间的所有活动均留存下各种类型的数据记录，当这些记录藉由大数据技术进行自动化的分析、比对之后，所有公民的一举一动甚至所思所想都被纳入系统的、广泛的监控当中，从而形成了边沁所言的每个人随时可能受到监视，但每个人却不知道何时受到监视的类似圆形监狱的效果。信息社会的特点决定了只要公民个人需要参与正常的生产、生活，就必须选择交出个人隐私，留存下各类个人信息。从这个角度来看，公民个人隐私的消亡是不可阻挡的历史发展趋势，渺小的个人在信息社会发展的大潮面前显得如此弱小与无力。当隐私消亡时，不仅仅是个人尊严、人格自治等固有的人类价值会受到威胁，从社会发展的整体角度观之，隐私权保护缺失的国度必然导致民主制度受损，也必然会威胁到公民个体创造力的发挥进而导致整个社会缺乏创造力与发展活力，国家的发展动能与样态令人堪忧。

大数据侦查植根于公民为参与信息社会生活而不得不交出并汇集的海量信息，必然带来大规模监控（mass surveillance）的效果，即全体公民的各种信息都成为了其分析对象，这是一种不以犯罪嫌疑为前提的广泛监控，全体国民甚至全球民众都可成为潜在的侦查对象，大量无罪公民的个人信息在大数据侦查的过程中被储存、比对、挖掘。

大数据侦查的广泛应用促使侦查权干预权利的类型发生转换与升级，侦查行为的对象由传统上的人身权、财产权转向平等权、隐私权、人格尊严、精神自由等基本权利和自由，权利干预的类型更加无形化、抽象化，在权利体系中的地位更接近权利构造的顶端。权利本身的无形性、抽象性令干预权利的侦查行为更难识别与感知进而导致权利的救济困难；权利位阶具备更强的政治性，则意味着与国家权力的冲突会更为剧烈。在这个意义上，对大数据侦查进行法律控制是维系

① 吉隆·奥哈拉、奈杰尔·沙德博尔特：《咖啡机中的间谍：个人隐私的终结》，毕小青译，北京：三联书店，2011年，第192—193页。

国家治理体系正当性的必然要求。如果任其发展，此类侦查方法将会加剧社会不平等的裂痕，抑制社会的活力与创造力。

同时，大数据也可能犯错，错误原因主要源于两个方面：一方面大数据技术通过机器学习与人工智能，根据侦查人员设计的各种模型对数据进行挖掘，而各种算法与分析模型的来源只能是侦查经验的人为积累。人类侦查经验的局限性会照搬给机器算法，大数据的预测功能同样会产生错误。在模型建构过程中，侦查人员的自由裁量权乃至偏见会融入大数据侦查当中，形成选择性执法、执法偏见与歧视。比如，基于过于某类手法的诈骗犯罪具有较强的地域性，侦查人员会将该地区的户籍所在地设为模型要素，将其作为重点监控对象，这显然属于违反法律面前人人平等原则的选择性执法与执法歧视。类似问题在大数据应用程度较强的美国普遍存在，大数据侦查过程中，对于社会底层人群特别是有色人种的执法歧视被进一步放大，比如在大麻毒品犯罪打击过程中，尽管白人与黑人具有相同的吸食比例，但黑人犯罪嫌疑人的犯罪数据更多地被收集并存入数据库，其结果是更多的黑人犯罪嫌疑人被警察抓捕。①

另一方面，大数据的挖掘或预测结果取决于数据的质量，作为源头的数据质量瑕疵将直接导致误导性甚至根本性错误。与商业领域不同，刑事司法领域的容错率相当有限，毕竟刑事司法事关公民的生命与自由。② 数据质量上的瑕疵将导致公民权利受到错误干预，大数据侦查的基础是正确、客观的数据库，而基于未经核实的甚至是错误的数据开展的大数据侦查将直接得出错误的推理结论，并误导着侦查机关错误干预公民权利甚至错误剥夺公民自由。无论是警方自建的各类数据库，还是利用社会第三方机构的数据库，司法实践已经反复证明，数据瑕疵与质量低下的数据经常导致错误的关联，甚至对公民自由带来直接损害。③ 数据质量瑕疵导致无辜公民被错误抓捕的事例在国内近年来的执法实践中也屡次出现，多名无辜公民由于身份证被冒用或重名、重号等原因而被警方错误羁押。④ 由于侦查机关相关数据库对公民信息的错录以及对数据质量管控的失责、失察，还会导致无辜公民的声誉、出行自由、参军招考、经济交往等基本权利受到侵犯。⑤

① Andrew Guthrie Ferguson, "Big Data and Predictive Reasonable Suspicion," *University of Pennsylvania Law Review*, vol. 163, no. 2, 2015, p. 402.

② 关于大数据侦查在美国司法实践中暴露出来的弊端及部分实际危害，参见 Andrew Guthrie Ferguson, "Big Data and Predictive Reasonable Suspicion," pp. 398-403.

③ Andrew Guthrie Ferguson, "Big Data and Predictive Reasonable Suspicion," p. 399.

④ 类似事例的报道可参见杨涛：《错误拘留频现亟待建立有效防范机制》，《北京青年报》2013 年 12 月 11 日，第 5 版。

⑤ 林崇寿、洪双敏：《错录公民违法犯罪身份信息引发问题的思考》，《河北公安警察职业学院学报》2017 年第 2 期。

三、传统法律规范框架的问题

人类社会迈入信息社会的发展态势与大数据侦查广泛应用的司法实践，超越了传统法律规范与法学理论所提供的规范框架。国际范围内形成于二战后的刑事诉讼法传统规范工具表现出滞后性，法律控制机制的阙如形成法律的真空或者稀薄状态，与大数据侦查的勃兴及其挑战形成鲜明对比。这一判断放在中国法的语境下依然适用。2012年刑事诉讼法修改对技术侦查措施设置了全新的规范程序，大数据侦查对刑事程序权利的干预深度与广度超过了技术侦查措施，但却处于无法可依的状态。基于比例原则的精神，干预公民基本权利的剧烈程度应当与其法律控制程序的正当性成比例，当前各国大数据侦查的法律控制强度均低于技术侦查的已有法律程序，法律控制体系严重失衡。

（一）大数据侦查的法律属性模糊

对一项全新科学技术在刑事侦查中的应用进行法律规范首先应当明晰其法律属性，大数据侦查属于何种侦查措施是对其进行规范的前置性问题。对于这一问题，有两种不同的解决方案：如果能够将大数据侦查归为传统侦查行为当中，就可以依照既有的法律规范遵照实施；如果传统的法律框架无法容纳下这一新技术，则需要修改法律创设全新的法律规范框架。面对这一新技术浪潮，不同法治传统的国家基于各自不同国情在上述两种解决方案上选择各不相同。

美国和德国作为两大法系的代表性国家，其各自规范大数据侦查的进路颇具代表性。美国联邦宪法第四修正案关于搜查及隐私权保障的判例法一直以来都被奉为规范政府各类获取信息行为的圭臬。1967年美国联邦最高法院裁决的Katz案是美国隐私权保障的标杆性判决，在该案中，美国联邦最高法院将联邦宪法第四修正案中搜查的界定标准由物理侵入说改为隐私保护说，隐私权保护的标准被确定为对隐私的合理期待。① 大数据侦查涉及对各类公民数字记录的应用，能否被视为搜查从而被纳入宪法规范视野，取决于大数据侦查是否构成干预公民对隐私的合理期待。根据美国联邦最高法院1976年Miller案和1979年Smith案确立的自愿交与第三方规则，即公民对自愿交给第三方机构保存的各类信息记录无隐私的合理期待，使用这些信息的政府行为当然不被视为搜查行为，联邦宪法第四修正案无从适用。② 即使经过几十年的时代变迁，面对大数据时代的来临，第三方

① Katz v. United States，389 U.S.387(1967).

② Miller v. United States，425 U.S. 435(1976)；Smith v. Maryland，442 U.S. 735(1979).

理论仍然主导着美国的隐私权保护规则。虽然在2012年的United States v. Jones案中，美国联邦最高法院在协同意见中提出，在现代电子化时代第三方理论应加以反思，但该案并未推翻Miller案与Smith案的基本结论。① 总体上看，美国联邦宪法第四修正案关注的焦点在于政府执法机构未经个人同意而获取信息的搜查行为，② 只关心数据的获取过程，对于获取数据后的使用过程并非第四修正案的规范旨趣。③ 数据挖掘与数据比对等大数据技术是对已经留存于社会各领域的海量数据进行后续深度应用的过程，只规范收集不规范使用的第四修正案及搜查法规范，导致在美国数据挖掘式的侦查行为基本上不受规范。④

德国基本法及联邦宪法法院规范政府干预公民个人信息的工具主要是人格尊严与信息自决权，并将其视为一种积极性权利，在宪法位阶之下的德国刑事诉讼法典也详尽规定了干预公民个人信息自决权的各类侦查行为。⑤《德国刑事诉讼法典》第98条a、b和第98条c分别规定了计算机排查侦缉和数据比对，⑥ 计算机排查侦缉与英语中的数据筛查（data screening）语义相同，是指通过计算机的数据模型对数字化的信息进行挖掘、比对以确定犯罪嫌疑人或者排除犯罪嫌疑人。从工作原理上看，德国法中的计算机排查侦缉与美国法中的数据挖掘是相同的信息技术应用过程。德国法典中规定的数据比对，是指刑事诉讼中获取的个人数据与政府已经掌握的执法司法数据库进行机器比对，以查明犯罪事实或者定位被侦缉人员所在地。⑦ 德国法典对计算机排查侦缉或者数据挖掘规定了严格的法定程序，而数据比对的规范密度要低得多，主要原因在于前者实施过程中可以对刑事追诉机关之外的其他部门保存的数据进行海量数据挖掘，⑧ 而后者比对的数据库仅为刑事司法部门管理的数据

① United States v. Jones, 132 S. Ct. at 957 (2012)，参见Sotomayor大法官的协同意见部分。

② Russell D. Covey, "Pervasive Surveillance and the Future of the Fourth Amendment," *Mississippi Law Journal*, vol. 80, no. 4, 2010, pp. 1289, 1294-1295.

③ Elizabeth E. Joh, "Policing by Numbers: Big Data and the Fourth Amendment," *Washington Law Review*, vol. 89, no. 1, 2014, p. 63.

④ Paul M. Schwartz, "Regulating Governmental Data Mining in the United States and Germany: Constitutional Courts, the State, and New Technology," *William and Mary Law Review*, vol. 53, no. 1, 2011, p. 354; Department of Defence, *Safeguarding Privacy in the Fight Against Terrorism: The Report of the Technology and Privacy Advisory Committee*, Create Space Independent Publishing Platform, 2004, pp. viii-x.

⑤ Paul M. Schwartz, "Regulating Governmental Data Mining in the United States and Germany: Constitutional Courts, the State, and New Technology," p. 354.

⑥ 参见《德国刑事诉讼法典》第98条。本文引用的中译本均为《德国刑事诉讼法典》，宗玉琨译注，北京：知识产权出版社，2013年。

⑦ 参见《德国刑事诉讼法典》第98条c。

⑧ 参见《德国刑事诉讼法典》，第54页。

库，二者涉及的公民个人信息自决权的干涉范围不同。对于计算机排查侦缉，《德国刑事诉讼法典》第98条a、b设置了与电话监听相当的严格程序，须遵循一系列干预公民权利的传统法律原则，比如法官令状原则、重罪原则、比例原则与最后手段原则等，同时还应遵循个人信息保护的基本法律原理，比如数据的有限使用原则、及时删除原则以及接受数据保护部门的监督。①

美德两国之间对于大数据侦查法律属性的差异化处理，根源于对此类侦查措施干预权利类型的不同认识与判断。德国法认为，大数据侦查是对公民个人信息自决权与人格尊严的干预，进而应遵循干预基本权利的基本要求，在刑事诉讼法典设置严格而详尽的法定程序；美国法坚持在联邦宪法第四修正案关于搜查与隐私权保障的框架内审视数据比对与数据挖掘，其结果是无法对大数据侦查施加有效控制。两国的共同之处是，从权利干预的角度出发来界定大数据侦查的法律属性。从规制思路的社会背景看，美国法仅关注个人信息保护中的核心区域，即隐私权保护，对其他大量个人信息保护问题持放任态度，这与美国信息产业蓬勃发展并维系其信息世界领导地位的社会发展需要直接相关；而欧洲大陆国家基于二战后形成的重视人格尊严、个人自治的法治传统，对公民个人信息保护强调严格的法律控制政策，当然这也在一定程度上限制了欧洲信息产业的发展。

对中国而言，隐私权与个人信息权两种规范路径的选择各有利弊，兼顾二者并适度调试两种规范路径在未来制度体系中的权重是更妥当的选择。整体上看，中国刑事司法中对隐私权的保护有待完善，同时也面临信息社会信息使用与保护的需求，这种迭代发展的现实状况要求在刑事司法制度的设计安排上应当通盘考量两种权利路径的兼容。当然，两大法系国家的出发点都是基于权利保障的视角对待大数据侦查，这一基本出发点尤其值得我们认真对待。

在中国的制度语境中，《刑事诉讼法》第二编第二章“侦查”共规定八种法定的侦查措施；“证据”章第48条在规定证据种类时，间接确认了辨认这种侦查行为。《刑事诉讼法》第113条还概括性授权侦查机关对已经立案的刑事案件，应当进行侦查并收集、调取相关证据材料。《刑事诉讼法》第52条规定公安机关有权向有关单位和个人收集、调取证据，有关单位和个人应当如实提供证据。公安部《公安机关办理刑事案件程序规定》第59条将《刑事诉讼法》第52条规定的调取证据视为一类侦查行为，并规定相应的调取程序与法律文书。②

在上述法定侦查行为中，有三项侦查行为可与大数据侦查产生关联，即搜查、调取与技术侦查。但笔者认为，上述三种侦查行为都难以作为大数据侦查的规范依

① 参见《德国刑事诉讼法典》第98条a、b。

② 《调取证据通知书》是公安机关进行调取证据时的制式法律文书，关于其内容、制作要求及样式参见孙茂利主编：《公安机关刑事法律文书（2012版）制作与范例》，北京：中国公安大学出版社，2013年，第297—302页。

据。换言之，大数据侦查的法律属性既不是搜查，也不是调取，亦不能被视为技术侦查。

首先，我国《刑事诉讼法》第134—138条规定的搜查与美国法中的搜查存在重大差异，前者仅指在被搜查人与见证人在场的情形下，对人的身体、物品、住处和其他地方等有形物或地点进行的搜索过程。① 大数据侦查的对象是数字化的信息，且获取、使用相关数字信息时信息主体并不知情。将大数据侦查比照为搜查进行规范，不符合我国刑事诉讼法的既有规范框架。

其次，调取并非刑事诉讼法明文规定的侦查行为，刑事诉讼法只是在“证据”章第52条第1款规定，公安机关有权向有关单位和个人收集、调取证据，有关单位和个人应当如实提供证据。根据《公安机关办理刑事案件程序规定》第57—59条以及《公安机关执法细则（第三版）》（以下简称《执法细则》）的相关规范，侦查实践中，当侦查机关发现有关单位或者个人持有与案件有关的证据时，即可予以调取，调取行为的对象是作为证据使用的实物证据，主要是物证、书证、视听资料。② 调取首先要表明调取的对象是与证明案件事实相关的证据材料，其次应当制作清单详细写明物品或文件的名称、编号、数量、特征等，被调取的单位和个人应签字确认调取的内容。③ 通过上述规范内容可知，调取行为根本无法作为大数据时代对海量记录进行比对与挖掘的规范依据，大数据侦查获取的全数据样本中必然包含大量与案件无关的信息，更谈不上满足“与犯罪事实有关的证据”这一调取行为的前提条件，如果让侦查机关逐一告知海量数据的持有人，则调取行为根本不具有可行性。调取行为的本质是小数据时代针对已有一定根据表明具体的持有人持有与案件事实证明有关的证据材料，进而要求其提供的一种非强制性侦查行为，在大数据时代，调取行为的本质功能如不进行拓展，根本无法作为获取海量数据的正当化手段。

最后，大数据侦查与技术侦查措施之间也存在本质的不同。2012年刑事诉讼法修改过程中新增技术侦查措施一节以及后续公安部制定《公安机关办理刑事案件程序规定》过程中，对于技术侦查措施的内涵与外延都采取了回避态度，导致技术侦查措施包括哪些具体的措施与手段十分模糊。《公安机关办理刑事案件程序规定》第255条将技术侦查措施的范围概括为记录监控、行踪监控、通信监控、场所监控等措施，在侦查机关看来，技术侦查措施的本质是监控，上条规定中的“记录监控”虽未进一步明确，但从名称上看与大数据对海量数据、记录的比对、挖掘的过程最为相关。对这一问题的讨论，应当回归技术侦查措施的本质问题。笔者主张技术侦

① 参见《刑事诉讼法》第134条。

② 参见孙茂利主编：《公安机关执法细则（第三版）释义》，北京：中国民主法制出版社，2016年，第295—296页。

③ 参见《公安机关执法细则（第三版）》第21—01条。

查措施的各类监控手段不仅应具有秘密性与技术性的特征，还应兼具同步即时性的本质要求。① 从立法者对已有技术侦查手段的部分列举中可以归纳出同步即时性的特征，技术侦查措施通常包括的电子侦听、电话监听、电子监控、秘密拍照或者秘密录像、秘密获取某些物证、邮件检查等专门技术手段，② 毫无例外均属在违法犯罪行为实施过程中同步展开的侦查行为。这与调取通讯记录或话单、查询财产等针对已储存信息的各类侦查行为在刑事诉讼法规范上存在明显区别。

正是由于现行刑事诉讼法及法律解释中侦查行为的分类无法容纳大数据侦查这一新兴侦查措施，公安部在《执法细则》中将“查询、检索、比对数据”单列为了一种侦查措施，规定进行下列侦查活动时，应当利用有关信息数据库查询、检索、比对有关数据：（1）核查犯罪嫌疑人身份的；（2）核查犯罪嫌疑人前科信息的；（3）查找无名尸体、失踪人员的；（4）查找犯罪、犯罪嫌疑人线索的；（5）查找被盗抢的机动车、枪支、违禁品以及其他物品的；（6）分析案情和犯罪规律，串并案件，确定下步侦查方向的。③ 这一规定凸显出数据比对、挖掘等大数据侦查技术的独立性，侦查部门也认识到此类侦查措施与刑事诉讼法已经规定的传统侦查行为之间的差异以及单独予以规范的必要性。当然，由于《执法细则》本身属于内部规范，仅限公安机关内部适用，不得在法律文书中引用，不向外部单位、个人公开，④ 这些特点导致《执法细则》欠缺法律文件的基本属性，相应的大数据侦查依然处于无法可依的状态。

迄今为止，人类社会经历了从农业社会到工业社会、再到信息社会的演进，刑事诉讼法对权利的保护重点也相应经历着由关注人身自由权到财产权、再到公民个人信息隐私权的变迁。大数据侦查在为侦查机关提供更高效的犯罪控制工具的同时，对公民个人信息隐私的干预程度超出传统侦查措施。信息社会发展至今，超过98％的信息都已转化为数字化记录，大数据技术得以对全数据进行分析、挖掘与应用。在迅速扩散的信息技术面前，规范隐私权的工具不能适应大数据时代的发展需要，因为社会与个人都需要依赖于海量个人信息的共享获得发展动力。传统法律规范缺失与滞后的主要原因在于其仅仅关注信息搜集过程，而对大数据背景下的如下核心问题完全忽略：当公民基于适应现代信息社会的必然要求而留存在社会各个机构的数字记录，侦查机关将这些记录改变最初留存目的用于侦查工作时，法律应当如何评价侦查机关的行为以及设定何种法定程序。我国刑事诉讼法的相关规定比较抽象，且法律解释工作相对滞后，对大数据侦查的本质和法

① 程雷：《检察机关技术侦查措施相关问题研究》，《中国刑事法杂志》2012年第10期。

② 郎胜主编：《〈中华人民共和国刑事诉讼法〉修改与适用》，北京：新华出版社，2012年，第277页。

③ 《公安机关执法细则（第三版）》第29—02条。

④ 《公安机关执法细则（第三版）》第1—02条。

律属性的认识与处理落后于大数据时代的发展步伐，司法实践中对大数据法律属性的认识盲区导致多层级的侦查部门大数据侦查技术的应用处于无序的发展状态，同时囿于法律授权的阙如，侦查机关的数据共享与合理利用也面临瓶颈。

（二）数据内容与元数据区别化处理的局限性

传统侦查过程对信息内容的重视程度远超过信息的形式，因为信息的内容可以直接作为证明犯罪的证据使用，而信息的形式主要是辅助证明信息的来源，其重要性不如信息的内容。数据信息的形式，即元数据是关于数据的数据或关于信息的信息，其表示的是数据的存在形式与产生过程，只要人们使用任何一种电子产品或者电子服务，都会产生元数据，以电子通讯为例，其主要包括通讯的时间、地点、时长、通讯双方的地址或号码，使用的电子设备及其唯一识别码。①

由于数字化时代对隐私权的干预方式主要是通过收集电子通讯的形式要素，再通过大数据的挖掘、分析技术深描出个人的完整信息，在传统观点下，这些通讯形式方面的信息与通讯内容不同，不是隐私权保障的对象。大多数国家对通讯形式的法律保护力度远低于通讯内容，在我国刑事诉讼中亦是如此，虽然法律文本上并未区分通讯内容与通讯形式，但司法实践中调取通话记录的适用频率远远高于对通讯内容的监控。② 这一传统观点在大数据时代的局限性愈发明显，因为大数据的本质就是对多样化的海量记录进行集成、碰撞以产生预见性的知识，从某种意义上讲，通讯的形式包括位置信息、通话时长、通话对象等比通讯内容更有价值。2014 年联合国人权事务委员会在其提交给联合国大会的专题报告中呼吁各成员国与时俱进地摒弃上述传统思维，在新信息技术背景下树立全新的信息保护理念，区分通讯形式与内容从保护隐私权的角度来看是不具有说服力的，因为信息的合成，通常称之为元数据（metadata），能够显示个人行为、社会关系、私人嗜好、身份等方方面面的信息，甚至比通讯内容更能全面地揭示一个人。③

---

① Bryce Clayton Newell，“The Massive Metadata Machine：Liberty，Power and Mass Surveillance in the U. S. and Europe，” *A Journal of Law and Policy for the Information Society*，vol.10，no.2，2014，pp.487-488.

② 这一结论通过对中国裁判文书网上的刑事裁判文书的关键词检索能够得到充分印证，以笔者 2017 年 7 月 1 日的检索结果为例，以“通话记录”为关键词可以检索到 105385 件刑事案件的裁判文书，而以“通话内容”为关键词检索，只能检索到 693 个裁判文书样本，二者之间的差异巨大。

③ The Right to Privacy in the Digital Age，Report of the Office of the United Nations High Commissioner for Human Rights，30 June 2014，p.3，http://www.ohchr.org/EN/HRBodies/HRC/RegularSessions/Session27/Documents/A.HRC.27.37_en.pdf，2017 年 4 月 6 日。

（三）侦查启动门槛虚置、已然犯罪与未然犯罪界限模糊

为防范警察权的滥用，两大法系国家都通过设置警察权启动的事实条件为刑事程序启动设置了限制条件。大陆法系国家的传统理论在对警察权控制机制上坚持区分犯罪预防与犯罪打击两个领域，二者的界限在于只有出现具体的犯罪嫌疑或者犯罪将要发生的即刻危险时，警察才能干预公民权利，此种警察的行动方式被界定为回应型警务模式（reactive policing）。① 这种警察职权启动模式可将警察权严格限制在不得以方可使用的必要范围内，有助于防止警察权的滥用。这种古典自由主义思想下的警察权控制模式在20世纪60年代起逐渐发生变化，警察不再仅仅被视为“执行工具”，而应成为智能化、主动型的犯罪抗制机构。基于这种理念变化，警察的调查方法发生了很大变化，一些“预防性犯罪控制手段”的侦查方法开始在侦查实践中推广，包括计算机数据库检索、拉网缉捕、电子监控等。② 通过这些大数据技术新型侦查手段的运用，警察可以发现用以确定初步怀疑的各种信息，从而正式启动侦查程序，如此一来，警察所承担的预防犯罪与打击犯罪两大截然不同的功能开始混合。

在英美法系的代表国家美国，法律规范警察执法权的起点是警察权对公民自由的干预，始于警察对公民的截停，自此刻起联邦宪法第四修正案为警察权启动设置的事实要件为合理怀疑（probable suspicion）。③ 对于警察针对某人截停前的发现、判断犯罪嫌疑的过程，美国联邦宪法基本上不予评价，委诸警察根据自己的经验以及具体案件、对象的个案情况进行自由裁量。④ 大数据技术在侦查初期的应用增强了警察发现犯罪嫌疑人信息的能力，凸显出合理怀疑标准的固有漏洞，也暴露出该标准的脆弱性。⑤ 大数据对潜在犯罪嫌疑人的强大识别功能，令原本设置在警察权启动之初的门槛性条件流于形式。

为防止侦查权的恣意启动、任意干预公民权利，我国刑事诉讼法将立案程序设

① Funk, A. Polizei and Rechatsstaat, 转引自 Cyrille Fijnaut and Gary T. Marx, eds., *Undercover: Police Surveillance in Comparative Perspective*, The Hague: Kluwer Law International, 1995, p.58.

② Funk, A. Polizei and Rechatsstaat, 转引自 Cyrille Fijnaut and Gary T. Marx, eds., *Undercover: Police Surveillance in Comparative Perspective*, pp.57-58.

③ Terry v. Ohio, 392 U.S. 1, 27 (1968); Andrew Guthrie Ferguson, “Big Data and Predictive Reasonable Suspicion,” p.329.

④ Elizabeth E. Joh, “The New Surveillance Discretion: Automated Suspicion, Big Data, and Policing,” *Harvard Law & Policy Review*, vol.10, no.15, 2016, p.33.

⑤ Andrew Guthrie Ferguson, “Big Data and Predictive Reasonable Suspicion,” pp.387-388.

置为刑事诉讼的起始程序，规定只有在立案之后侦查机关方可行使侦查权。[①]“认为有犯罪事实需要追究刑事责任”是立案的事实证据要求，[②] 为发现犯罪线索或者确认是否达到启动侦查的事实门槛，相关法律解释进一步规定了立案前的初查程序，允许侦查机关采取不限制被调查对象人身权与财产权的各类措施。[③] 但上述法律规范严重滞后于大数据侦查技术的应用实际，侦查机关对海量数据记录的查询、比对、碰撞正在成为锁定犯罪嫌疑人的重要方法，由于其应用时段多在立案前的初查阶段，甚至在并无具体犯罪嫌疑的前嫌疑阶段使用以达到“无中生有”的预测犯罪或者抓获现行犯的重要作用，其深度应用的同时也逐渐侵蚀甚至架空了立案程序的立法目的。

总体上看，大数据技术在侦查初期的广泛应用在各个法系国家都导致基于限制警察权而设置的侦查启动门槛流于形式，已然犯罪与现行犯、即将发生的未然犯之间的界限愈发模糊。大数据技术令警察权突破了传统法律框架在起点环节上的约束，形成了初期侦查权规制的法律真空。

四、通过侦查规范的法律控制

信息社会无疑是人类社会发展形态上的重大飞跃，但不容否认的是信息社会是在传统社会发展形态基础上逐步演变、发展起来的，相应的人类社会的治理方式与治理模式也应当在承继传统的基础上加以革新。正是基于这一判断，笔者认为规范大数据侦查的路径既要遵循传统规范框架，更应沿着个人信息保护的新兴路径深入探索，应当同时关注传统的侦查法律规范与相对新兴的数据保护法律规范，两类规范相互协作的双重路径是当下对大数据侦查进行法律控制的适当选择。在双重路径并行的过程中，规范重心应当更侧重于隐私权的保障，因为是隐私权而非个人信息权承载着人格尊严、个人自治、私生活安宁、通信自由等一系列公民弥足珍贵的基本权利。刑事司法过程中政府对公民个人信息的利用是信息社会发展到一定阶段政府治理模式的必然要求，刑事司法又承载着维护社会安全与稳定的特殊利益与价值追求，因此个人信息保护的诸多制度安排无法在刑事司法系统中得到完整落实，权衡之后的结果只能是个人信息保护制度在刑事司法中的适

① 参见《刑事诉讼法》第 110、113 条。

② 参见《刑事诉讼法》第 110 条。

③ 《公安机关办理刑事案件程序规定》第 171 条规定，在立案审查环节中，“对于在审查中发现案件事实或者线索不明的，必要时，经办案部门负责人批准，可以进行初查。初查过程中，公安机关可以依照有关法律和规定采取询问、查询、勘验、鉴定和调取证据材料等不限制被调查对象人身、财产权利的措施”。《人民检察院刑事诉讼规则》第 173 条也规定了类似的初查程序与权限。

度应用。在本部分笔者首先就第一条路径传统侦查规范工具的适用加以探讨，另一路径关于数据保护方面的法律控制问题容留下一部分详述。

大数据侦查作为一种新兴侦查措施，其运行机理与我国现行刑事诉讼法中规定的各类传统侦查行为均存在本质差异，应当作为一类独立的新型侦查行为进行法律规制。在大数据侦查权的规范立场选择上，基于侦查规律与大数据侦查的运行机理，首先应更新如下三项规范理念，才能为具体规则的建构提供基本支撑：其一，犯罪类型的嬗变引领着大数据侦查的兴起，侦查工作的起点愈发向前延伸，在前瞻性侦查阶段为维系侦查权有效行使与权利保障的平衡，事后监督与控制比事先审批机制更具规范价值。同时，类似立案程序式的侦查启动门槛制度基本无法适应大数据侦查的发展需要，对侦查权的控制模式应当由关键节点控制转向过程控制。其二，用隐私换安全、用信息换公平应当成为规范大数据侦查的基本策略，因此大数据侦查的合法化过程应当伴随着常规侦查权规制的进一步正当化与严格化，大数据侦查的正当化过程同时伴随着的是干预公民人身权、财产权的侦查行为进一步严格化，进而继续维持侦查权与公民基本权利的大体平衡。其三，应当深刻洞察到侦查权的本质是对公民基本权利的干预，不应仅仅从权力行使的外观或形式角度规范侦查权。物理属性不应再是侦查权行使的本质特征，权利干预强度应当成为规制侦查权的基本视角，后者应当成为决定规范密度、规范工具的出发点。

基于现有法律规范经验与国际社会通行的规范策略，规范路径主要还是应当遵循合法性原则与比例原则或称之为必要性原则这两项传统法律规范工具。基于合法性原则，规范我国大数据侦查首先应当解决的问题是增补相应的法律依据，为开展大数据侦查提供清晰、具体与公开的法律依据。根据大数据侦查技术应用领域的二分趋势，应当分别补齐犯罪侦查和犯罪预防、情报收集两项领域相关的法律规范。前者应在刑事诉讼法的“侦查”章增补一类全新的侦查措施及适用程序，后者需要修改完善国家安全法、反恐怖主义法、网络安全法等相关部门法。①

就犯罪侦查领域大数据侦查的具体规范路径而言，首先应在《刑事诉讼法》“侦

① 《中华人民共和国国家安全法》第53条规定，开展情报信息工作，应当充分运用现代科学技术手段，加强对情报信息的鉴别、筛选、综合和研判分析；《中华人民共和国反恐怖主义法》第18条规定，电信业务经营者、互联网服务提供者应当为公安机关、国家安全机关依法进行防范、调查恐怖活动提供技术接口和解密等技术支持和协助；《中华人民共和国网络安全法》第28条规定，网络运营者应当为公安机关、国家安全机关依法维护国家安全和侦查犯罪的活动提供技术支持和协助。上述法律中的相关条款对于情报收集、犯罪预防中的大数据侦查进行了初步性、概括性的授权，但法律规范明确、具体的程度距离合法性原则的要求还有一定差距，比例性原则也未得到确立。

查”章第八节“技术侦查措施”中将大数据侦查增列为一种全新的侦查行为加以规范，同时应将该节的节名进行更为准确的表述，改为“秘密侦查措施”，以涵盖第151条及新增列的大数据侦查相关条文。① 可考虑在现行《刑事诉讼法》第151条之后增加专条规定，凡是使用计算机技术对政府数据库、社会机构数据库进行信息记录共享、检索、比对、分析的行为，均属大数据侦查。

其次，应当详细列明大数据侦查的启动条件、适用对象与适用程序，在启动条件、适用对象与适用程序的制度设计过程中应当充分考量比例原则的要求。鉴于大数据技术引发的“隐私逐步消亡的世界”这一重大挑战，考虑到其大规模性的对无辜公民个人信息的比对使用过程以及结果意义上对具体公民的近乎所有个人信息全面、深入展示的效果，比例原则具有极大的适用必要性。大数据侦查的适用首先应当坚持目的正当原则，即“只能用于对犯罪的侦查、起诉和审判，不得用于其他用途”。② 排除在外的其他用途包括个人目的、政治目的、违背法律精神的社会维稳、上访人群的管控等，均属违反法定目的的侦查权滥用行为。在适用条件上，应当明确启动该项侦查措施的事实门槛条件，即只有具备初步的犯罪嫌疑之后方可启用，在启动决定的法律文书中应当明确表明依据何种已有的事实材料表明存在何种具体的犯罪嫌疑。这种启动条件与现行刑事诉讼法所表述的立案条件应大致相当，附加此种启动条件的目的在于防止基于犯罪侦查之外的其他不正当目的任意启动大数据侦查，也有助于防止侦查机关基于维稳或安保等宽泛的执法需求而任意启动大数据侦查，从而导致常态化的全民监控、大规模监控。

程序合法性原则要求侦查措施应当经过法定程序审批后方可启用，以德国为代表的部分大陆法系国家对大数据侦查实行法官令状制度，比照电话监听实施相应的程序控制机制。就我国的情况看，不具备实行法官令状机制的条件和可能性，理由主要有二：一方面，大数据侦查作为侦查措施的一种，其审批主体的制度设计应当与我国刑事诉讼中既有的强制措施、侦查措施的程序控制机制相协调。在剥夺公民自由的逮捕措施、严重限制公民自由的指定居所监视居住、严重干预公民隐私权的技术侦查措施尚未实行法官令状制度之前，对大数据侦查实行法官审批的司法审查与强制性措施体系均衡性要求不符。另一方面，不少实行法官令状制度的国家司法实践已经证明对于侦查初期的技术侦查、大数据侦查等特殊侦查

① 现有的“技术侦查措施”的节名原本属于“搭车式”的表述方式，2012年法律修改时为回避使用“秘密侦查”一词，将节名表述为更中性的“技术侦查措施”。但就该节规范内容看，除技术侦查措施外，第151条规定了另外两类秘密侦查措施，即隐匿身份的侦查和控制下交付。大数据侦查本质上也是秘密侦查，尽管相关数据多为公开留存于各个数据库的信息，但对数据的分析、碰撞过程属于典型的秘密侦查过程。

② 参见《刑事诉讼法》第150条关于技术侦查措施目的正当原则的规定。

手段，法官令状制度经常流于形式，法官通常会沦为警察适用相应侦查措施的橡皮图章。① 国际社会的经验与教训亦表明，对于侦查初期的大数据侦查，在实行司法令状机制的同时，更应注重综合监督机制。总体上看，就我国目前的状况而言，当务之急是改革大数据侦查封闭运行的现状，较为便捷可行的方式是实行检察官审批制，辅之以紧急情形下的侦查机关自我先行审批机制。② 同时应当考虑建立各类基于个人信息权的数据使用监督机制，关于这方面的内容容留下文一并加以论述，通过上述两方面监督机制的建构，有助于预防大数据侦查的恣意启动与过度使用。

在必要性原则方面，应当坚持重罪原则与最后手段原则，即鉴于大数据侦查对犯罪嫌疑人和大量无辜公民的个人信息甚至隐私信息施加了全面监控与比对，该项措施只能限于预防或者侦查严重犯罪时方可使用，这是比例原则的基本要求。严重犯罪的范围可参考《刑事诉讼法》第148条及《公安机关办理刑事案件程序规定》第254条对重罪的范围加以规定，即包括危害国家安全犯罪、恐怖活动犯罪、黑社会性质的组织犯罪、重大毒品犯罪或者其他可能判处七年以上有期徒刑的案件。最后手段原则要求大数据侦查应当是在采取了询问被害人或证人、勘验检查犯罪现场、对物证进行提取、鉴定等常规侦查措施之后且常规侦查措施无效或者难以锁定犯罪嫌疑人时方可适用。通过最后手段原则的要求，可以限制无犯罪嫌疑的情况下漫无目的、漫天撒网式的大规模数据比对，同时要求大数据侦查以前期常规侦查获取的信息为基础，在科学合理的信息模型上开展有针对性的、高效精确的数据挖掘。

封闭的内部运行模式是各类侦查措施滥用的主要成因，因此加强外部监督与司法监督是确保侦查权依法运行的基本经验。根据现行宪法框架与司法体制的国情，可以考虑通过加强检察机关检察监督与通过审判过程中对证据来源合法性的审查发挥法官的审查功能来实现对大数据侦查的司法审查功效。《刑事诉讼法》第8条规定了人民检察院依法对刑事诉讼实行法律监督的基本原则，在大数据侦查的检察监督问题上，应当进一步增设具体的法律监督机制，在《刑事诉讼法》第152条之后增加专条规定检察机关对秘密侦查行为进行监督。同时可以考虑设置大数据侦查的备案机制、办案电子系统互联互通等方式，要求侦查机关在适用大数据侦查之后，应当将大数据侦查的开展过程、相应成果报告给负责侦查监督的检察官，接受备案审查与法律监督。如果检察机关发现侦查机关的秘密侦查行为

① The Right to Privacy in the Digital Age, Report of the Office of the United Nations High Commissioner for Human Rights, 30 June 2014, Para. 38, http://www.ohchr.org/EN/HRBodies/HRC/RegularSessions/Session27/Documents/A.HRC.27.37_en.pdf.

② 参见程雷：《秘密侦查立法宏观问题研究》，《政法论坛》2011年第5期。

有违法情况，应当通知侦查机关予以纠正，侦查机关应当将纠正情况通知人民检察院。

法官对大数据侦查的司法审查根据大数据侦查在诉讼中的作用不同可以通过两种途径展开。对于少量使用大数据侦查获取的信息作为证据的案件，法官可以通过对大数据侦查的证据材料进行审查判断从而对取证的合法性进行间接审查。而在大量案件的司法实践中，大数据侦查的主要功能是发现案件线索或锁定犯罪嫌疑人，其本身是证据来源的前置性工作而非证据收集工作本身。在这种情况下，审理案件的法官应当着重审查到案经过或者破案报告等侦查机关制作的用以说明案件来源的书面材料，查明大数据侦查的过程是否符合法定程序，对于违法的大数据侦查行为应当在裁判文书中列明，并将违法情况移交人民检察院开展法律监督。在未来的法律完善过程中，还可以考虑赋予法官对于大数据侦查中的程序违法行为给予独立制裁的权力，比如建立针对重大程序违法行为的终止诉讼机制，责令侦查机关对违法责任人加以惩戒、处分，等等。

五、通过数据规范的法律控制

信息社会的到来对隐私权的既有法律保护框架形成了冲击，传统意义上具有消极、被动等特点的隐私权概念已经显得过于狭隘，很难适应社会发展的需要。① 个人信息权作为保护公民个人信息的全新法律工具日益受到社会各界的重视，成为社会转型过程中法律治理体系的重要工具。可以说，将个人信息权作为一种独立的权利是现代社会发展的一种趋势。②

在公民个人信息保护法律体系的建构与运行过程中，多数国家的个人信息保护法中都将国家安全与刑事司法领域的个人信息保护排除在法律适用范围之外。③ 这一传统观念过于关注公共利益，无视公民个人信息权的基本法律价值，导致近年来国家大规模监控在全球范围内的盛行。从长远发展的视角观之，信息社会的发展必须依赖强大的个人信息保护机制。从公民基本权利保护的视角观之，在现代信息技术之下，几乎所有的个人行为都会留有信息痕迹，这些信息痕迹关涉个人生活的方方面面，实现了对个人从摇篮到坟墓的全程记录；现代信息技术可以实现对个人碎片化信息的整合，随着信息质和量的累积，碎片化的个人信息逐渐形成个人的“人

① 周汉华：《中华人民共和国个人信息保护法（专家建议稿）及立法研究报告》，北京：法律出版社，2006 年，第 48 页。

② 王利明：《论个人信息权的法律保护——以个人信息权与隐私权的界分为中心》，《现代法学》2013 年第 4 期。

③ 周汉华：《中华人民共和国个人信息保护法（专家建议稿）及立法研究报告》，第 58—59 页。

格剖面图”，① 从这个角度观之个人信息对维护信息主体的人格尊严和自由价值意义重大。大数据侦查也必须在打击犯罪与维护公民个人信息权、人格尊严、个人自治等价值追求之间寻求平衡，个人数据保护的法律原则与机制在刑事司法领域应当得到适度应用。②

之所以强调适度应用，主要是基于两个方面的考量：一方面大数据侦查的兴起是侦查方式顺应信息社会蓬勃发展的社会发展规律的产物，信息是信息社会最为重要的发展资源，也必将作为社会治理方式的核心要素，从这个角度来看，积极利用海量数据发展大数据侦查是社会发展的必然要求。个人信息的范围大于个人隐私，个人信息权的权能较之隐私权更为积极主动，个人信息保护机制的内涵与制度设计也与隐私权的保障原则、理念存在诸多不同。刑事司法承载的国家安全、社会安全等一系列社会价值与公民个人信息的保护之间应当进行相应的权衡，引入个人信息保护机制只能是适度进行，国家为履行在信息社会条件下保护国民安全的使命可以干预公民的个人信息权，但应当遵循最低限度的个人信息保护规则。另一方面，也应当认识到个人信息保护的不少法律机制与侦查的既定目的和侦查规律不无冲突。比如个人信息公平保护实践要求对个人信息的使用应当坚持自愿同意原则、公开透明原则，然而，防范反侦查的执法目的以及侦查效率的要求使得侦查机关无法在利用海量个人信息前征得各个信息主体的知情同意，大数据侦查模型中的算法设计过程也因为涉及侦查经验、犯罪规律等侦查秘密，基本无从做到公开算法。因此对于信息公平实践中的知情同意、算法公开透明等原则，在大数据侦查的规范体系中只能适度应用。

在我国个人信息保护法出台前，③ 对大数据侦查的规制，建议引入如下个人信息保护方面的法律原则和机制。

（一）目的合法与特定原则

大数据侦查过程对公民个人信息的收集与处理应当基于合法、具体且特定的执法或司法目的，不得超越收集个人信息时的合法目的使用相关个人信息。“合法”是

① 张新宝：《从隐私到个人信息：利益再衡量的理论与制度安排》，《中国法学》2015 年第 3 期。

② 2016 年 4 月 27 日，欧洲议会与欧洲委员会在通过旨在全面保护公民个人信息权的《通用数据保护条例》（General Data Protection Regulation）的同时，通过了《以犯罪预防、调查、侦查、起诉或者刑罚执行为目的的自然人个人数据保护指令》，将个人数据保护的法律原则与机制部分引入刑事司法领域。

③ 近年来，我国立法机关通过分散立法方式，在《刑法修正案（七）》和《刑法修正案（九）》中加强了个人信息的刑法保护；在《民法总则》第 111 条规定了个人信息的保护规则；在《网络安全法》第 41—42 条规定了网络个人信息保护的条款。从总体上看，仍缺乏一部专门的个人信息保护法，以统筹公民个人信息保护的各个方面。

指基于刑事诉讼法、网络安全法等法律明确授权的执法与司法的正当目的；“具体”是指法律授权时不应准许无任何嫌疑基础则发动大规模信息收集与处理活动；“特定”是指当社会机构将数据库信息传递给侦查机关或者与侦查机关共享数据库时，应当事先告知信息主体在何种情形下为追诉哪些犯罪行为，社会机构将与侦查机关共享公民的个人信息。目的特定原则禁止未经信息主体的明示同意或授权将商业用途产生的个人信息用于犯罪侦查，除非信息收集主体或处理主体在收集与处理数据前已经明确告知信息主体特定目的中包括了未来可能将其个人信息用于对其犯罪的追诉。通过目的合法与特定原则，信息主体在行使个人信息的知情同意与授权前具备了对该个人信息未来可能使用目的的全面、清晰的认识，大数据侦查对公民个人信息的收集、处理方才具备正当性。

（二）信息主体的知悉权与更正权

信息主体的知悉权与更正权是保障信息主体信息权，防止信息管理者、使用者、控制者滥用公民个人信息的重要制度安排。刑事诉讼中的被追诉人作为信息主体应当有权知悉信息被司法机关收集的目的及用途，有权查询、修改、更正不准确、不客观或过时的数据信息。信息主体的知悉权对于被追诉人知悉控方的指控方向与证据来源，进而有效准备辩护至关重要。同时知悉权也是信息主体充分行使信息权的前提条件，也是信息主体寻求法律救济的基础性权利。信息主体的更正权是确保数据质量的重要机制，也是防范刑事司法中数据失真引发错误司法行为的有效工具。当然基于刑事司法顺利进行的合理理由，侦查机关可以推迟告知犯罪嫌疑人、被告人等信息主体个人信息的干预过程及结果，但推迟告知信息主体的例外应当是明确而具体的法定事由。适当的法定事由可以包括两个方面：一是涉及国家安全和国家核心利益的案件，比如危害国家安全、恐怖活动犯罪等，由于涉及犯罪组织以及情报信息的未来使用，可以推迟告知信息主体；二是告知信息主体可能有碍侦查的，侦查机关应在有碍侦查的情形消失后立即进行告知。

（三）信息安全与数据质量控制机制

侦查机关对信息的处理过程应当体现安全性，与其他信息处理者、使用者相比，侦查机关具有更为强大的信息收集与使用能力，其汇总的各类个人信息中不仅规模巨大，更包含大量高度敏感信息，这对信息安全提出了更高的要求。因此应当考量建立并运行各类信息安全防护机制，制作监控日志并做到操作留痕。

侦查过程中，根据比例原则的要求，对于公民信息应当实行分级管理，对于公民个人敏感信息，在刑事司法与侦查活动中应当重点保护，设置更高的审批权限与启动事实条件。公民个人敏感信息是指那些一旦遭到泄露或修改，会对标识的个人

信息主体造成不良影响的个人信息。① 刑事司法中的个人信息至少包括行踪轨迹信息、通信内容、征信信息、财产信息以及住宿信息、通信记录、健康生理信息、交易信息等。②

数据质量控制机制是大数据侦查正确展开、防止侦查错误的基础性制度，数据收集主体包括商业机构、社会机构与侦查机关均应当根据《网络安全法》等相关法律的规定建立确保数据真实性的相关机制，侦查机关在开展大数据侦查过程中应当通过数据清洗、多库交叉检验等方式验证数据的真实性。禁止包括侦查机关在内的数据使用者、管理者共享、传输无法验证真实性、过时的相关数据，同时数据使用者与管理者也负有及时修正虚假、过时信息的相应义务。

（四）个人信息使用的监督与救济程序

对个人信息的处理应当设置相应的监督与救济程序，除前文提及的刑事程序中外部审批机制之外，还应当根据个人信息保护的特有要求，建立独立监督机构、定期报告机制等救济渠道，强化对个人信息使用的监督。大数据侦查的监督机制应当在侦查效能、侦查方法保密与公民个人信息保护之间寻求平衡，监督的重点应集中于平等权的保护、禁止选择性执法、数据安全的执行状况等总体情况。监督的方式应当以事后监督为主，因为大数据侦查处于前侦查阶段，对于侦查经验的应用与选择、侦查启动的时间节点选择等，事先审批与监督并不具备可行性。在事后监督机制方面，可以考虑结合个人信息保护法的起草，设立个人信息专门保护机构对大数据侦查等个人信息使用机制进行事后监督，个人信息保护机构可以受理信息主体的权利救济申请，也可以依职权进行调查，或者要求侦查机关就大数据侦查使用的整体状况定期加以备案或者建立定期报告机制，对大数据侦查使用情况进行定期审查。

结　　语

大数据侦查在中国的发展既存在着特有的必要性，也面临着独特风险。一方面，作为国家治理体系重要组成部分的刑事司法系统必须拥抱大数据，唯有如此才能有效化解深刻转型社会所带来的犯罪率持续攀升、新型犯罪层出不穷的社会治理难题，有效治理口供过度依赖的传统刑事司法弊端，严格落实无罪推定原则，防范冤假错案。大数据侦查是顺应信息社会背景下侦查规律的必然选择，符合社

① 参见《信息安全技术公共及商用服务信息系统个人信息保护指南》第 3.7 条的规定。

② 参见 2017 年《最高人民法院、最高人民检察院关于办理侵犯公民个人信息刑事案件适用法律若干问题的解释》第 5 条。

会控制机制演进的基本趋势。中国是一个数据大国，且具有“集中力量办大事”的制度优势，大数据侦查的发展具备更多有利条件。比如，基于全球数量最多的互联网使用用户、移动手机用户、公共视频监控视频，以及移动支付平台、共享经济平台、互联网金融平台，大数据侦查拥有极为丰富的数据资源，数据比对与数据挖掘具备绝佳的开展条件。另一方面，应当认识到，正当的法律程序与个人信息保护制度对于大数据侦查的良性、可持续性发展同样至关重要，发展大数据侦查的各项优势缺少法律控制机制的制衡极易演化为巨大的风险。比如数据量大并不代表数据质量高，瑕疵数据、错误数据的大量存在反而更易误导侦查工作走上歧途，甚至导致公民权利被错误干预甚至剥夺，在这方面个人信息法律制度的引入对于大数据侦查的健康发展至关重要。

大数据侦查这一方兴未艾的新型侦查行为，也为刑事诉讼法学传统理论提出许多新的研究课题，客观上推动着刑事诉讼法学研究范式的转型。首先，大数据侦查凸显出无罪推定这一现代刑事诉讼基本原则存在适用空间上的边界，其无法向前延伸至犯罪嫌疑产生之前的前嫌疑阶段。恰恰是在前嫌疑阶段，大数据侦查应用空间广泛，其重要价值是在刑事司法程序开始前发现启动刑事司法程序的信息与线索，而无罪推定原则无论是作为狭义的证据规则还是作为广义上的权利保障原则，都无法在被追诉人产生之前的前嫌疑阶段予以适用。现代刑事诉讼法的绝大多数原则与规则都是建立在无罪推定原则基础上的，其在大数据侦查中的适用真空附带导致多数诉讼规则与制度的空转。现代刑事诉讼法学理论必须继续探索发展其理论范畴填补这一空白。其次，以尊重和保障人权为重要使命的刑事诉讼法传统上关注的权利类型主要是人身权、财产权与隐私权，大数据侦查的出现使得权利干预类型进一步无形化，传统权利干预形态逐渐为大数据侦查所替代，监控社会的到来也就意味着绝大多数犯罪过程会被如实记录，追诉犯罪的过程越来越不需要依赖干预传统权利的各类侦查行为，但同时监控社会的加速形成会引发人们对言论自由、思想自由的忧虑，算法歧视会带来平等权保护的迫切需要，刑事诉讼法学的研究需要关注这些课题，才能在更为宏大的视角之下合理规范大数据侦查。

从大数据未来发展方向的角度观之，有三大趋势值得持续关注。第一，大数据侦查的深度应用将升级犯罪的类型，导致犯罪打击的难度逐步上升，在侦查与反侦查的多轮较量中，街头犯罪、暴力犯罪将会因为大数据侦查的有效打击而逐步退出历史舞台，相应的犯罪人群将进一步转向更为隐形化的经济犯罪，同时为规避大数据记录的搜集，犯罪的地点将更为全球化，基于境外实施的各类犯罪类型将进一步多发。侦查机关应用大数据的过程中需要不断培养专业分析人才、研发新型算法模型以适应愈发隐蔽、变化的新型犯罪手法，同时还应当开始探索数据全球化共享的规则与机制，建立数据司法协助的相应制度安排。

第二，伴随着我国刑事司法制度中以审判为中心的改革逐步推进，证明标准、证据规则的严格适用对侦查机关取证的规范性提出了越来越高的要求。伴随着大数据侦查在案件侦破中发挥的作用愈发突出，法庭对其证明作用的需求也就会日益凸显。实现大数据侦查由“幕后”走向“台前”，需要对大数据证据问题展开进一步研究。现有证据法的理论与规则提供的解决方案极为有限，比如大数据侦查的分析结论归于何种证据种类、适用何种证据规则；如何进行人脸识别、声音视频、生物信息识别上的同一认定；如何在庭审上对大数据证据进行质证、如何在保障质证权与保守侦查方法秘密之间寻求有效平衡；等等。诸多证据法问题都需要未雨绸缪展开研究，迎接大数据侦查的常态应用所引发的刑事审判方式变革。

第三，大数据侦查的发展将改变政府与商业机构在刑事司法中的关系格局，刑事司法界应当开始关注如何在法律上评判二者之间的相关关系这一全新课题。大数据侦查的数据来源除了政府各部门基于政府管理需要而收集、储存的数据之外，多数的海量信息来源于商业机构为公民提供日常生活服务、经济交往当中储存的各类信息。大数据侦查越来越多地需要与商业机构的数据库互通共享，而传统刑事司法的规范原理是规制公权、保障私权。如何跨越这一规范鸿沟，需要法学界与法律界进一步思索。

〔责任编辑：刘　鹏〕

social mobilization can complement each other. Case studies have shown that those implementing policies at the grassroots level develop different mobilization strategies depending on the relative strength or weakness of administrative control and the society's capacity for mobilization. In the course of policy implementation, the boundaries and relationships between hierarchical control and social mobilization and between government bureaucracy and grassroots society may change in line with the demands of policy performance. This hierarchy may permeate the social network or the individual level, so that the social network becomes a part of the hierarchy. The overall process of implementation thus exhibits "adaptive social mobilization." This finding, based as it is on the dynamic process of policy implementation, may lead to the rethinking of questions including the nature of societal governance in contemporary China. It also provides an explanation of the paradox of the mutual reinforcement of administrative control and social participation.

(4) Legal Control of Big Data Investigations *Cheng Lei* · 156 ·

Through the use of computer technology to collect, share, screen, compare and unearth data stored online and in computer systems, big data investigations can locate clues, evidence or suspects. Such investigations have three main modes: the goal-driven, the comparison-driven, and event-driven. These are of practical use in crime prevention and prediction and in the field of detection. As big data investigations challenge some basic rights and legal values, they have to be brought under legal control. However, the traditional framework of legal norms lags behind. Definition of the legal properties of big data investigations is unclear; there are limits to differentiating data contents from metadata; the threshold for launching criminal proceedings exists in name only; and the boundary line between ascertained offense and unaccomplished crime is also unclear. Dual approaches can be adopted for gaining legal control of big data investigations: detection standards and data standards. With detection standards, we should follow the principles of legality and proportionality and strengthen external and judicial oversight. With data standards, we propose introducing appropriate legal principles and mechanisms for protecting personal information, including establishing legitimate aims and specific principles, giving data subjects the right to know and the right to make changes, and setting up

arrangements for information security and data quality control, together with procedures for supervision of the use of personal information and relief procedures.

(5) The Formation of the Concept of the Silk Road and Its Transmission in China

The concept of the Silk Road was first put forward by the German geographer Richthofen in 1877 with reference to the route connecting Chang'an to Central Asia. Before this term was formally adopted in China, the Chinese had called this route "the road of silk and satin," "road of silk," "road of silk sales," etc. "Silk Road" (*sichouzhilu* 丝绸之路) was first used in China on 24 February, 1943, in the *Shanghai Journal*. Thereafter the concept underwent several transmutations, such as "desert road," "oasis road" and "steppes road"—what we call networks today. Other terms referring to this route were "jade road," "spice road," and "road of furs." Although silk was not the most important commodity in all periods of Chinese trade with the West, and the Chinese referred to the Silk Road in many different ways, no other expression has been able to take the place of "Silk Road." The Belt and Road initiative has taken its name from the Silk Road Economic Belt and the Maritime Silk Road, reflecting the practical contemporary use of a historical expression.