大数据监督的江华探索

法制周报·新湖南记者 雷昕 实习生 田国华 通讯员 邓秋燕 王希明

“只要将数据库一比对碰撞，可疑违纪线索便立即搜索出来，再移交给相关部门进行处理。”江华瑶族自治县纪委监委信息保障部门技术人员张寅边现场演示边介绍“清廉中国”大数据实验室。只见几个按键操作，异常信息立即显示出来，真正实现跑腿监督到智慧监督。

2020年5月，在中国社会科学院中国廉政研究中心的指导和帮助下，江华县引进清廉中国大数据实验室项目，开始了大数据监督的新探索，先进的理念和大数据技术，让监督有了质的飞跃。

星星之火，可以燎原。一场大数据监督的实践探索正在湖南省永州市江华瑶族自治县试点展开。作为全国首创者，江华县委、县纪委监委把大数据实验室作为建设法治政府、落实全面从严治党的重要抓手和创新举措，以敢为人先的精神推进大数据监督的落实落地，取得了显著成效。

为落实习近平总书记讲话指示精神，一体推进不敢腐不能腐不想腐，营造海晏河清的政治生态，推进清廉中国建设，近日，《法制周报》记者与湖南廉政智库调研组走进江华，探寻江华县大数据监督密码。

江华瑶族自治县纪委监委召开“清廉中国”大数据实验室问题线索研判分析会。

大数据监督的起源

江华大数据监督探索始于2017年。

“当时中纪委提出了‘互联网+监督’的号召，这既是任务、也是工作。于是我们成立了‘互联网+’办公室，开始了大数据监督的探索。”时任江华县委常委、县纪委书记蒋剑介绍，“大数据监督是一种趋势，能很好地规避人情案、关系案，让监督执纪行为更公开公平公正。”

2017年，江华县开始探索运用信息技术加强监督检查，建设江华智慧纪检监察平台。该平台由内部监督门户和外部公开门户两个部分组成。作为整个平台的核心，内部监督门户包含12大基础数据库，涉及国家公职人员、村干部、房产、车辆人员、门面信息等。利用这些数据库的自动碰撞、自动分析、自动提示等功能，及时发现民生领域违纪违规线索。

外部公开门户将基础设施建设、惠民补贴发放、城市和农村低保户的评定等民生资金项目录入平台。群众可以一键查询，发现问题进行投诉。“当时的智慧监督平台还不像现在这样智能，只能做一些简单、横向的信息比对。比如，把公职人员数据库与低保户数据库进行比对。”当时的平台建设负责人党风政风监督室主任黎海燕介绍。

江华县纪委监委运用平台实现对20项民生资金的网上监管，重点整治城乡低保资金中优亲厚友、关系保、人情保现象。通过平台数据比对，发现了国家工作人员、村干部及近亲属享受困难群众惠民政策的疑似问题线索，其中涉及低保的2822条、住房危改120条，其他类880条。经核实，给予党纪处分5人，追回违规资金65万余元，挽回损失274万元。

把社会救助这块漏洞堵上后，江华县又在全省率先做了工程项目、公务接待、公车使用等方面的大数据监督。重点对各部门特别是二级单位的公务接待费用实行“四单”（公函、审批单、菜单、发票）网上监管，遏制舌尖上的腐败，通过系统自动预警，发现公务接待超标准、超人数等问题。

截至目前，平台已审核公务用餐数据1.3万余条涉及金额800余万元，全县公务接待费用同比下降13.66%。根据线上核查提供的线索，与线下开展专项监督检查相结合，追缴违规资金4047.75万元，给予党纪政务处分6人。

此外，平台利用大数据自动对比分析，对单位中标多个工程项目、中标单位负责人与收款人信息不符等问题线索进行实时监督。“通过平台工程监督系统的数据比对，有的工程项目建设方和工程监理单位竟然是同一家，有的公司竟然一年时间在江华县11次中标，发现疑似问题线索165条。”黎海燕介绍，经过核查，对相关部门负责人27人进行了责任追究。

经过多年探索实践，江华的公务接待网上监管成效得到了湖南省纪委充分肯定并在全省试点推广，公车监督目前全省已全部推广。如今，这个智慧监督平台依然在运行。打开该平台的公车使用界面，可以看到全县公车使用的实时活动轨迹。

第1代的智慧监督平台，解决了思想理念的问题。但是，随着监督的深入，县纪委监委发现，现有的系统已经找不到类似的问题。一方面是交办问题进行集中整治后，一些行业领域建章立制促进了自身的规范性。另一方面，监督对象也在刻意规避这些问题，腐败变得更加隐蔽。

2020年5月，在中国社会科学院中国廉政研究中心的指导和帮助下，江华县启动大数据实验室项目，成为清廉中国大数据实验室试点县，开始了第二次大数据监督的探索。

“清廉中国”大数据实验室主界面。

建设大数据实验室

大数据实验室建设起步初期就遇到了重重阻力。资金难、技术难、数据采集难、思想观念统一难、保密工作难……平台建设涉及各个部门，虽然廉政建设是当前的大趋势，但推动这项工作，县纪委监委仍遇到了很多困难。

“平台最初推广的时候，我们纪委内部也是存在不同意见的，一方面考虑到平台建成后，对我们的监督管理是否有效。另外一方面要从各个部门采集数据，也涉及到保密问题。当时反对的声音比较大，也有过激烈地争论，但在县委书记和县纪委书记的大力支持下，最终达成了共识。”县纪委监委副书记张日华回忆。

在县委书记的大力支持推动下，平台建设的人力与财力得到了保障。县财政拿出150万建设大数据实验室，并为负责实验室运行维护的互联网+监督信息中心增加8名事业编制，以县委的名义向中国社科院廉政研究中心申请设立“清廉中国”大数据实验室基地。县委专题召开动员部署会，要求相关职能部门在数据提供、专业技术人员抽调等方面积极配合。

大数据实验室围绕监督目标和监督重点，有目的性、针对性地进行数据采集。经过反复论证明确了平台建设需要的核心数据源，既能保证对监管对象分析的全面性，又不影响各职能部门业务的正常开展。在县委协调下，目前已采集到来自财政、人社、卫建、医保等17个单位的40多个数据库1400多万条数据。

由于各部门数据标准各异，数据一致性较差，技术人员建立了多个数据识别模块，对采集的数据进行统一清洗、过滤、转换，形成内部格式统一的字段和编码。再对数据的性质和权限进行梳理，确定了数据的保存时限和共享范围，形成一套统一字段、统一标准、统一接口的多源异构数据融合。

技术人员按照模块化和逻辑分层等技术手段支撑下的可扩展性原则、任务导向符合实际需要的原则、方便系统迁移的跨平台性原则、经济性原则等，搭建起了一套数据存储和比对平台，再根据不同算法进行数据比对。

基于上述特征的识别、组合和计算，大数据监督平台生成了3个分析模型，用于进行异常报销情况、团伙关联分析等多种行为分析，大大扩充了原有的研判模型和算法，使大数据监督平台具有了最新形态违规行为的发现能力。

“依据中国社科院的理论，人类社会是由人、财、物、时间、空间5大要素组成的。大数据实验室依据这5大要素的脉络关系发现问题，运用大数据进行测量，从背景、关系、数额、规律上计算分数，分数越高，可疑性越大。通过采取数据的系统管理员、安全管理员和审计员3分开的措施，防止数据被篡改，保障数据的真实性。”大数据实验室平台技术开发总监周存杰介绍。

大数据实验室结合了国际最先进的归一算法、区间算法、动态权重算法等，创造性地引入了多层次动态积分算法，通过大数据赋值计算出的“违规分数”，用直观的方式描述了监管对象违规可能性的大小，增强了发现问题线索的能力。

试水医保专题监管

大数据实验室建好后，首先在医保领域试水发力监督。

“当前医保基金监管形势严峻，各项违规使用医保基金的行为时有发生，单纯依靠传统人工筛查的监管方法来加强医保管理的难度越来越大。”江华县医保局副局长王希明介绍，“我县现有各级医院30余家，参保人员49万余人，县医保部门监管人员仅5人，要实现全流程精准有效监督，仅靠现有人员远远不够。”

2020年5月，大数据实验室建成后，江华县创新了医保基金监管模式。利用信息化手段，对参保人员住院情况和重点医院进行大数据横向、纵向动态分析，对虚假住院、循环住院、过度医疗等套保、骗保行为实行线上预警、线下查处。

大数据实验室依据人、财、物、时、空的脉络关系，结合医疗报销管理制度，构建了医保积分计算模型、数据冲突碰撞模型、住院动态违规模型、医院总体违规模型等15个数据模型。

“我们把医保、卫建、人社等16家单位的40个数据库放在这些模型中。对整体住院情况和重点医院进行大数据横向碰撞和纵向动态分析，就发现了大量的异常信息。”县纪委监委信息技术保障部门工作人员张寅介绍。

五保户朱德全（化名）的异常情况就是在该过程中发现的。

大数据实验室通过医疗系统数据纵向动态分析，朱德全2020年8月12日从第1家医院出院，当天即转入下一家医院住院，8月21日出院后，次日又转到另一家医院住院。横向比对交通运输部门的乘车信息和医院住院费用清单可以发现，朱德全的输液记录卡上记录的静脉输液时间是上午8时31分，但公交车乘车信息却显示他当天上午8时28分还在距离该医院几公里外的地方乘车。

经调查核实，在两年时间里，朱德全因颈椎、腰椎、慢性气管炎在4家民营医院轮流住院40多次，报销总额超过10万元，个人累计自费仅700余元。这暴露出涉事医院通过诱导特困供养人员住院，套取医保基金的违法违规行为。

“通过大数据平台监督，大大缩短了我们在前期数据查找异常线索的情况，提高了办案效率，对医疗机构也形成了震慑作用。”王希明介绍，“通过运用大数据实验室主动发现线索，医保基金监管变被动等待为主动发现；十几家相关单位数据互联互通，实现协同发力、联合监督；变人工筛查为精准比对，运用智能技术，精准发现问题，提高了监管效能。”

据了解，去年通过大数据平台，县医保局实现了对全县31家定点医疗机构监管全覆盖。总共发现413个问题线索，查处了349个问题，查实率为84.5%。追回医保基金272.67万余元，立案查处4件，行政处罚共计248.5万元，预计全年节约医保基金900余万元。

王希明发现，开展大数据监督以来，全县各医疗机构串换收费、分解收费等违规行为有了大幅度下降。同时，各医疗机构也提高了自主管理的自觉性。通过平台数据分析，找准了当前基金监管的漏洞和短板，督促医保机构建章立制、完善管理、提升监管水平，从更高层次统筹推进智能监管平台建设。

江华县医保专题试点成功后，今年，永州市医保局着手引进“清廉中国”大数据平台，对全市2区、9县、2个管理区共13个县区的住院情况进行大数据分析，取得了显著效果。

搭建通用平台织密监督网

随着时代的发展，数据越来越呈现一种高度分散化、碎片化的特点。数据越来越呈现杂、乱、散的特点，纪委监督对象多，监督行业多，需求变化快，没有任何一个平台，能够满足纪委监委日常监督任务的需要，而对于县域基层监督机关来说，也不可能支持一个庞大的技术团队不断根据业务变化修改需求和升级系统。

为了解决这一困境，清廉中国大数据实验室设计了高度灵活的通用平台业务模块，不需要专业的计算机技术人员支持，纪委工作人员自己可以非常方便地设计数据库和采集数据，设计数据碰撞模型和输出规则，从而极大地满足了纪委多种多样的日常监督需求。

“通用平台就像房子的地基，是针对纪委监委的工作特点而专门设计的高度灵活性的系统。在这个通用平台上可以随时新建各种数据模型，采集相关数据。”周存杰解释，“举例说明，通过经济数据发现招投标竞争对手之间存在着利益共同体关系，通过时空碰撞发现医保骗保的确凿证据，通过多单位数据发现干部违法违纪问题等。都是通过通用平台，由纪委工作人员自己实现的。”

周存杰认为，这既很大程度满足了基层纪委监委监督的需要，又符合了县域纪委监委机关的技术力量薄弱、资金不足的现状，是一种基于实际需要而产生的一种有效的解决方案。

2021年8月，大数据实验室启动第2期通用业务平台建设，探索对工程招投标领域的大数据监督，升级医保基金监管模型。

“招投标领域涉及的问题更广、利益更大，数据逻辑性关系不像社保领域那么强，数据分析的切入点就更难精准定位。从程序上看，都是合法合规的。实际上这些反复的串标、围标动作早已在私下完成。”江华县委常委、县纪委书记唐亚辉认为。正因为类似的问题的存在，2021年9月，县委启动工程招投标领域机动式巡察，剑指项目围标串标乱象，重拳出击，发现干部违纪违法问题线索11件。

“现在大数据信息公开在网上、群众权力运行也在网上，我们的监督也要同步更新。”时任江华县纪委常委周斌认为，现在大数据监督已经进入深水区，看得见的问题已经监督。但是信息手段是日新月异的，系统建设也不可能一劳永逸。这样的探索更多的应由相关主责部门根据自身需求来构建问题模型，进行问题排查，纪委主要是立足监督的再监督。

对于江华的大数据实验室探索，中国社会科学院中国廉政研究中心秘书长蒋来用评价“具有开创性意义”。他认为，这种从无到有的探索过程是对县域特色监督的有益尝试。由于投入少、见效快、性价比高，江华县大数据实验室符合清廉建设的要求，是未来的发展趋势。

四川省达州市大竹县纪委监委到江华县考察大数据实验室后，现已推广使用。“目前来看，在东西部区县可以普遍推广。”但在蒋来用看来，大数据实验室的成果要推广到全国，还需要一个过程。

县医保部门到医疗机构实地核查系统住院人数与实际在院人数。