纪委监委公职人员监督研判模型

建设方案

内部文件，仅供参考

人民数据 PEOPLE DATA

2023 年 10 月

一、建设背景

党的十八大后，党中央大刀阔斧、披荆斩棘推进反腐败斗争，党的十九大后一刻不停、坚定稳妥，党的二十大后持续发力、纵深推进，反腐败斗争压倒性胜利不断巩固深化。

《中央纪委国家监委信息化工作规划（2018-2022年）》的印发，对新形势下推进纪检监察信息化建设提出了明确要求。创新纪检监察工作方法，实现监督工作全面信息化，还需要进一步积极探索利用大数据、云计算等延伸监督范围，拓宽监督领域，联通“线上线下”监督，深化数据利用。

中共中央办公厅印发的《中央反腐败协调小组工作规划（2023—2027年）》指出，严查重点问题，坚决查处政治问题与经济问题交织的腐败，坚决防止领导干部成为利益集团和权势团体的代言人、代理人，坚决防止政商勾连、资本向政治领域渗透等破坏政治生态和经济发展环境。紧盯重点对象，严肃查处领导干部配偶、子女及其配偶等亲属和身边工作人员利用影响力谋私贪腐等问题，加大对“政治骗子”和政治掮客的惩处力度。

因此，树立大数据理念、运用大数据技术，是纪检监察工作体现时代性、把握规律性、增强创造性的必然选择，有助于更好履行监督执纪问责的基本职责。一方面，这是新形势下加强和改进纪检监察工作的必然要求。随着正风肃纪力度的不断加大和网络信息技术的加快发展，各类“四风”行为、腐败行为也变得更加隐蔽化、智能化，运用传统人工模式从海量的数据中发现贪污腐败行为及其共性和规律难度太大、成本太高，必须引入大数据等技术。另一方面，大数据的迅速发展也为纪检监察大数据应用提供了有力条件，可有效推动监督方式从逐个、碎片化核实到全面、深层次碰撞迭代升级。

内部文件，仅供参考

二、现状分析

纵观全国智慧纪检监察的地方实践，普遍存在着应用系统“碎片化”、业务数据“孤岛化”、信息资源“沉默化”等问题，严重制约了地方实践的深化拓展。总体来说，纪委监委的大数据应用方面，存在数据缺失、数据壁垒严重，共享机制未建立等问题。

1.应用系统“碎片化”，业务数据归集难。部分地区纪检监察业务中的巡视巡察、信访举报、谈话函询、审查调查、问责处理等情况汇总，多以纸质文档或者电子表格形式分散在各个业务处室，未能做到自有业务数据的统一汇总与有效整合。

2.业务数据“孤岛化”。纪检监察领域信息化建设水平不一，信息化系统、业务数据系统零散分布，各系统间缺乏统一的数据规范标准，数据孤立，缺乏信息资源共享的手段，产生信息孤岛，为数据共享、管理和利用带来很大困难，而传统的数据联通手段导致纪检工作的业务协同效率低下。

3.数据资源“沉默化”。对纪委可支配的数据，如廉政档案、业务数据、本地专线数据等缺乏有效的综合利用手段，未能发挥其数据分析、关联、线索挖掘的能力，使得数据资源沉默、浪费。

4.外部数据支撑缺失，壁垒严重。数据要发挥作用，一是要有关键数据源，二是要多维度数据源关联碰撞，当前某些高价值数据掌握在行业部委或互联网平台中，本地纪委协调难度大。在实际业务场景中，纪委监委需要大量外部数据支撑，如企业工商数据、金融消费数据、互联网舆情数据、电商消费数据等，此类数据可有效补充本地数据的不足，为案件研判提供高价值线索。

5.审理调查手段落后，缺少信息化手段。在立案以后，纪委监委可协调如银行流水、移动运营商数据等。但此类数据多而杂，办案人员需要投入大量的精力来手动梳理线索，

内部文件，仅供参考

耗时耗力，而采用大数据手段，可以快速抽丝剥茧，抽离、关联关键信息，梳理有价值线索。

三、平台介绍

（一）模型简介

纪委监委公职人员监督研判模型（以下简称模型），通过建立统一数据标准，汇集多维度廉政相关数据，打造廉政大数据分析底座，对数据进行深度加工和分析，通过大数据精准研判违纪信号，预警廉政风险，精准追踪已发生违纪问题，深度融合纪检工作业务场景，打造精准化、智能化大数据工作服务平台，具有“日常监督与审查调查”两大结构模块，具备实时监督、线索预警、关系关联、办案工具等多项综合预警功能。



（二）数据源总览

数据类型	数据范围
人企关系数据	全国企业（不含港澳台、特殊主体）董、监、高人员数据信息核验查询服务，包含企业名称、职务

京东商城订单数据	京东商城注册用户的订单数据，包含购买人、物品、收货地址、收货人、用户画像、购买类别等订单相关字段
教育评分数据	教育部体系2003年以后毕业的学历数据查询核验（不含港澳台、海外），包含学历、学校、专业、入学时间、毕业时间、类型（全日制\其他）
个税数据	全国全量用户个税数据查询，区间分值，500元为一区间
公积金数据	全国全量用户公积金数据查询
婚姻数据	全国全量用户婚姻数据查询
房产数据	可以支持全国房产数据查询
企业工商数据	全国企业（不含港澳台、特殊主体）历史库和每日更新库。 数据包括：工商基本信息、股东信息、工商变更记录、主要人员、分支机构、企业行业、应缴明细、实缴明细、经营异常、行政许可、动产抵押、知识产权相关、司法相关、行政处罚、行政许可、年报信息等。
招投标数据	全国全量公开招投标数据信息。 数据包括项目标段名称、发布时间，开标时间，链接、公告名称、地区、采购人、招标方式、代理机构、联系人、中标人、中标金额等。
舆情数据	境内、外主流媒体平台数据。 监测范围覆盖：网站、论坛、博客、平面媒体、微博、微信公众号、境外媒体、新闻客户端、自媒体号、短视频、广播电视及视频等。

内部文件，仅供参考

司法数据	包括司法涉诉排查和行政执法排查相关数据查询。 查询自然人（全称、证件号）和企业（全称、统一社会信用代码）。 数据包含标题、公告类型、公告内容、具体日期、日期类别、案由、案号、审理机关、原告、被告、其他当事人、诉讼地位、数据类型、异议备注、异议申诉状态、事件履行状态、信用修复状态。

四、建设目标

“模型”通过多维度、精准化的大数据挖掘运用，对公职人员进行精准“画像”，强化对领导干部的智慧管理，实现智慧监督、智慧办案，练就精准识别违规违纪“隐身衣”的“火眼金睛”。

1、**通过大数据监督，实现党员干部监督全覆盖。**“模型”大数据覆盖全域党员管理干部，并进行持续更新，通过对大数据的深度挖掘关联，能够及时有效的对全域党员干部进行监督，并根据类型进行预警。

2、**通过大数据挖掘分析，提升发现线索的深度、广度、精度。**加强监督执纪，对严重违纪违法线索，充分运用大数据分析模型精准发现“蛛丝马迹”，做到主动发现问题线索、抓住关键证据，加强立案审查。

3、**通过建立调查审理任务，提高办案效率。**通过将线索信息归集，提供高效办案工具，能够有效提升大数据归集和分析处理的工作效率，节约调查审理时间。

五、建设内容

（一）数据采集与存储系统

针对数据来源分散问题，提供最优智能数据整合服务，通过接口集成、离线同步、

内部文件 仅供参考

互联网采集等多种手段，定向收集监测人员的多源异构数据。将碎片化的数据资源进行汇总抽取，对数据采集态势进行整体感知，实现多源异构数据的横向、纵向的汇聚融合。最终实现数据的安全存储，方便后续的数据分析和应用。

（二）数据加工与分析系统

围绕“模型”中纪委监委的工作需求，对已有的多源异构数据进行进行整合、清洗、转换、映射等，在解决数据质量的基础上利用图数据库分析、机器学习及指标建模等技术，生成监控人员综合评价模型、违纪预测模型、两点关联模型、人员关系图谱等相关模型，实现精细化、智能化及全流程的多维分析与监管。

（三）高价值数据接入

对纪委办案过程中涉及到行业部委、互联网平台的数据，如：工商登记数据、电商消费数据、出行数据、互联网舆情数据等，将进行有效数据输入，协助纪委办案获得高价值线索。

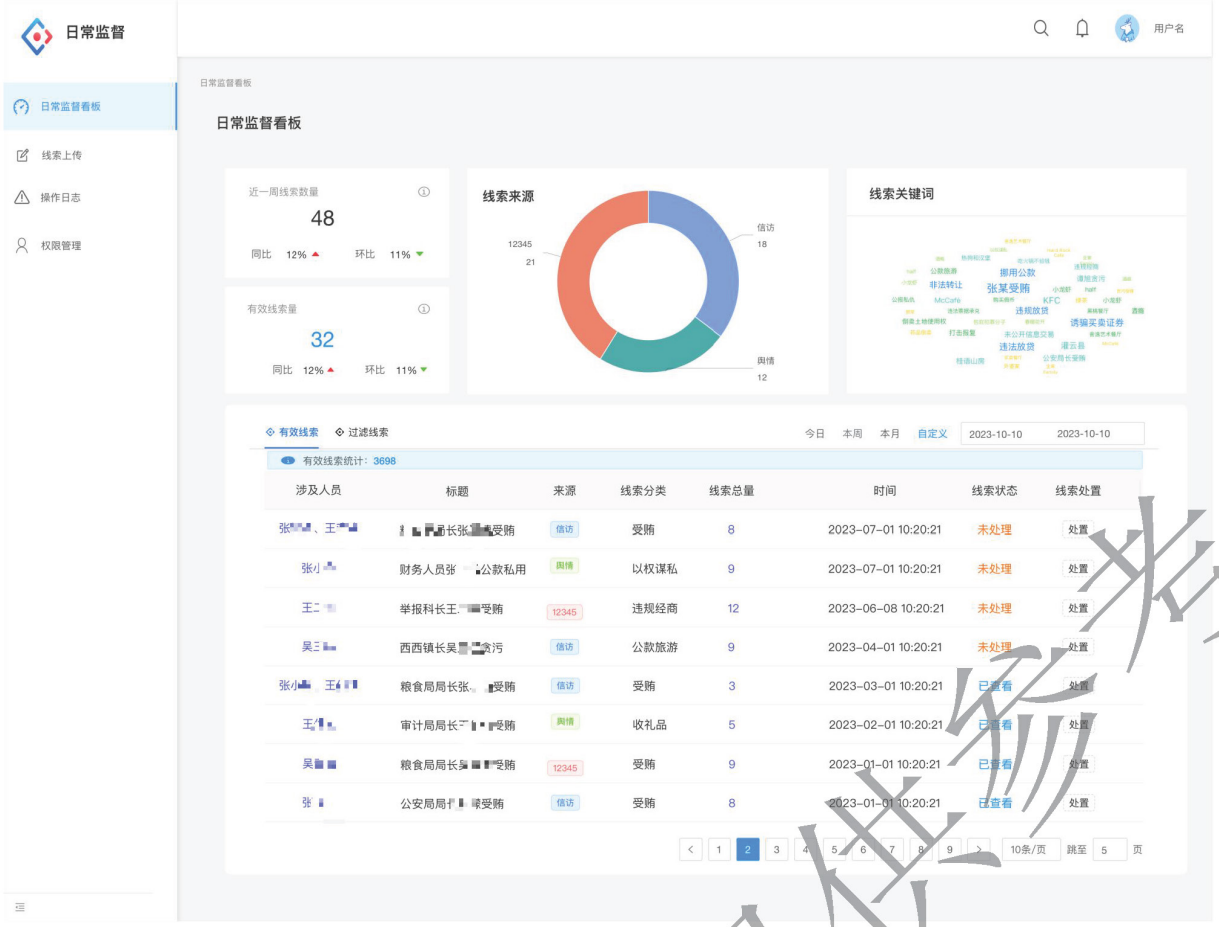
（四）业务应用系统

业务应用系统是对数据的显性应用，包含模块如下：

（1）日常监督

日常监督模块，对举报、信访、舆情等数据线索进行智能化研判，自动识别线索价值，并关联关键人、关键事、关键时，实现监督探头前移，做到监督的精准有效。

内部文件 仅供参考



（2）线索汇总

线索汇总模块，围绕人员及社会关系、资金、经商、通话、消费等多场景数据进行挖掘、线索研判，实现对有效线索按照主题、类型进行汇总，并对人员穿透、多层关联、资金、经商、通话、消费等线索直观展示。

内部文件 仅供参考



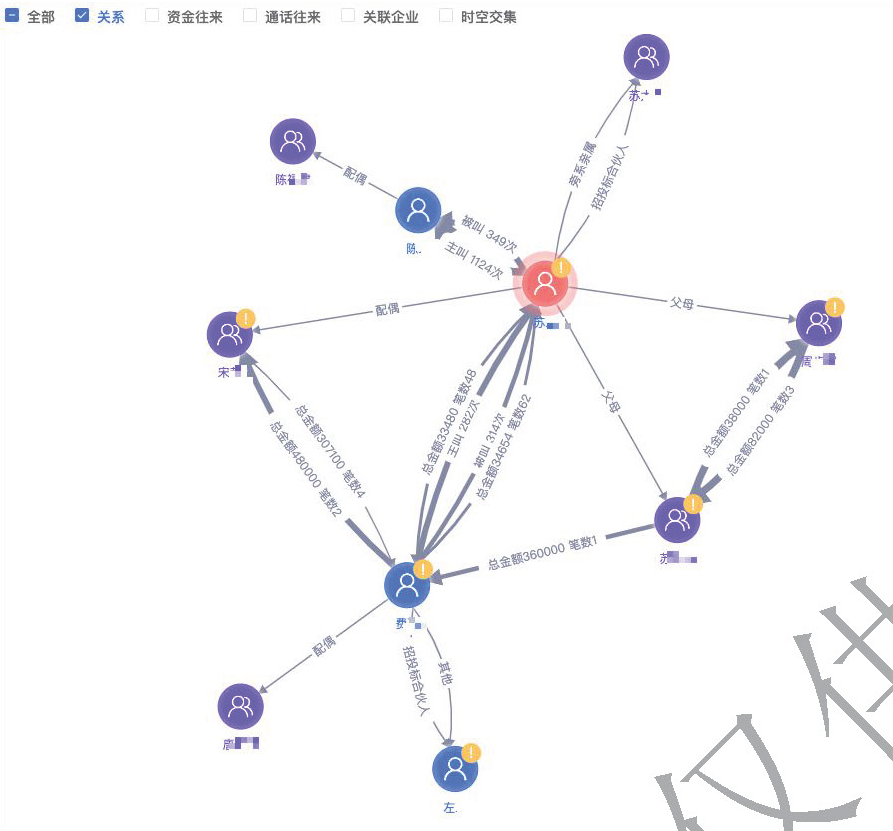
1）人员穿透线索

人员穿透线索充分利用廉政档案、公安户籍、银行流水、通话、通讯录、工商、招投标等数据，充分挖掘人与人及人与企业之间的同学、同事、资金交易、密切通话、经商办企业、参与政府项目等关联关系，对线索进行充分分析，并形成人员、企业关联穿透图。

1、人员关系穿透

利用廉政档案、公安户籍等信息，有效研判直系、旁系亲属、同学、同事等关联关系，并通过图谱形式展示，通过关系穿透图，快速清晰查看人员之间亲属关系及社会关系。

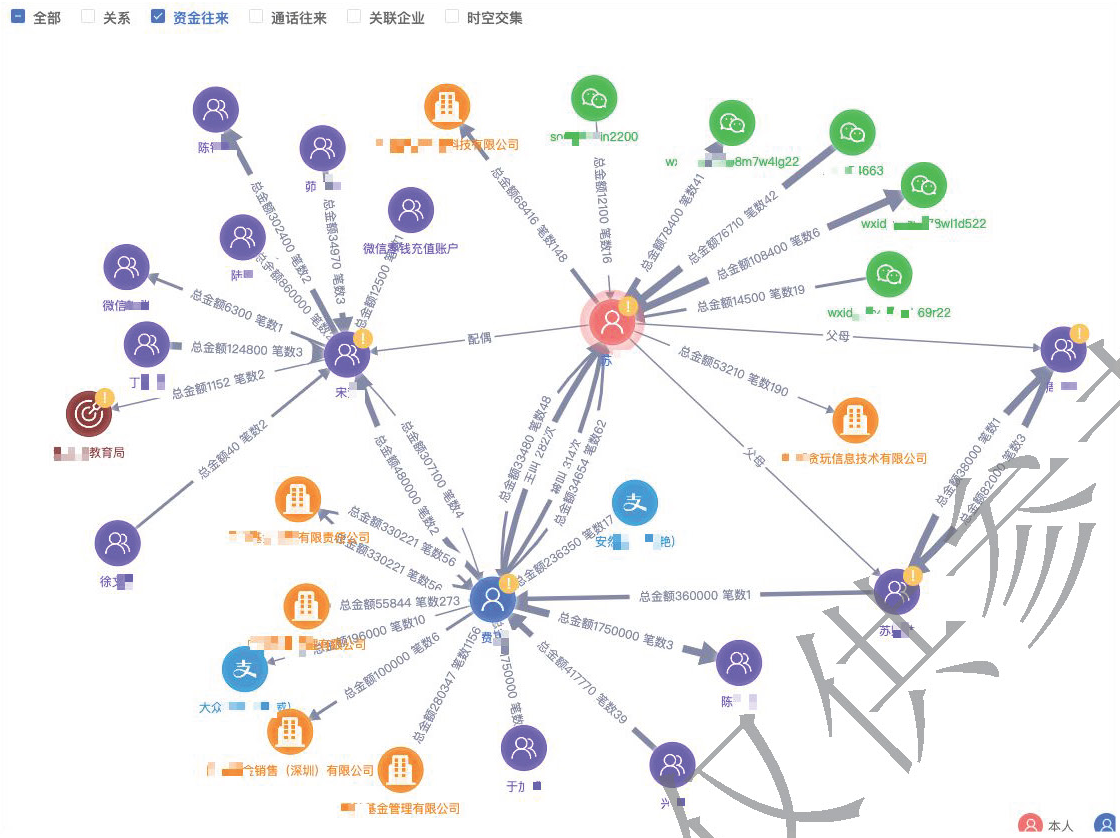
内部文件 仅供参考



2、资金往来穿透

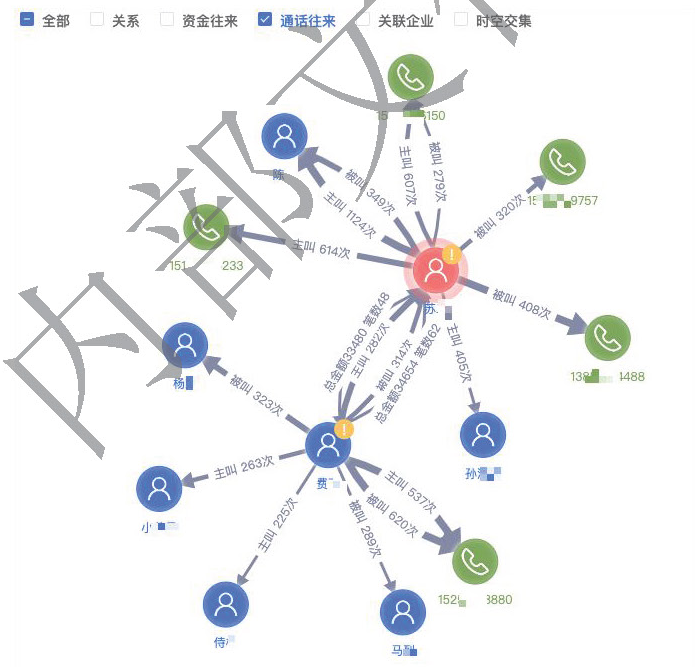
对银行流水（含支付宝、微信、各大银行）进行分析、关联，形成人员、人企资金穿透，通过穿透图可直观展示资金往来情况，有助于快速发现人员间资金关系，实现对可疑资金及可以人员、企业的快速定位。

内部文件 仅供参考



3、通话往来穿透

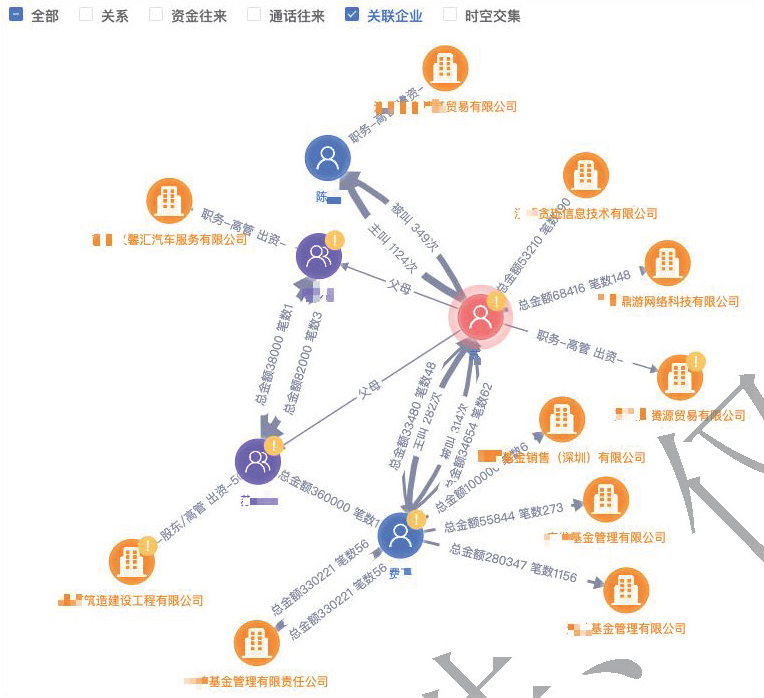
结合通话记录及通讯录数据，可实现高频通话关系的数据分析与穿透，方便提取电话往来频繁地人员名单及对应亲属或社会关系。



内部文件 公安系统

4、关联企业穿透

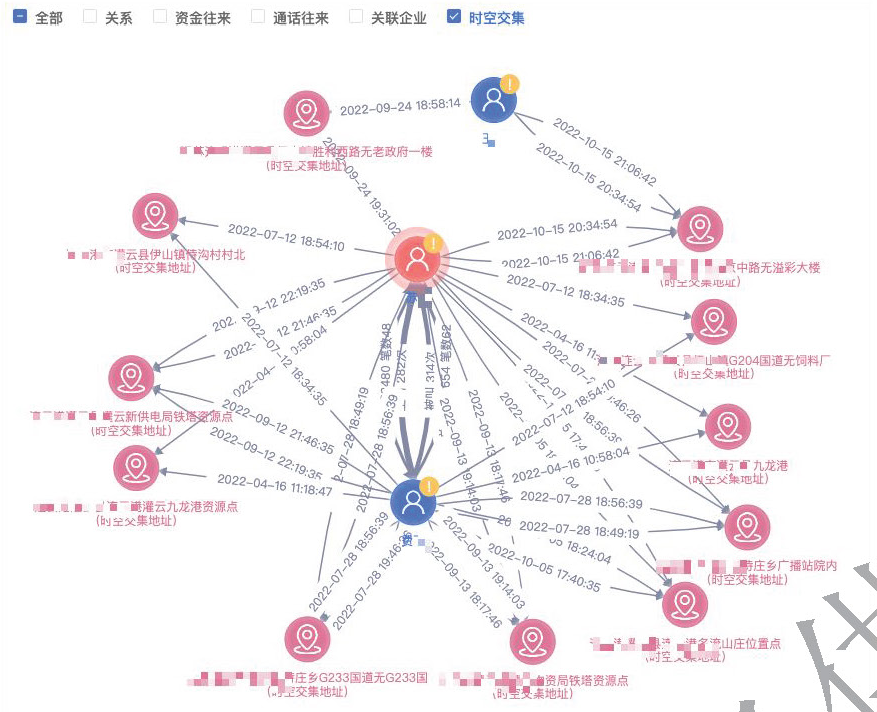
关联企业穿透，充分利用经商数据、企业银行流水、企业招投标等数据，结合人员关系穿透，实现目标人员与企业的关联关系图谱。可有效发现家庭成员经商、家庭企业参与政府项目、参与政府项目企业与公职人员资金及通话关系等重要线索。



5、时空交集

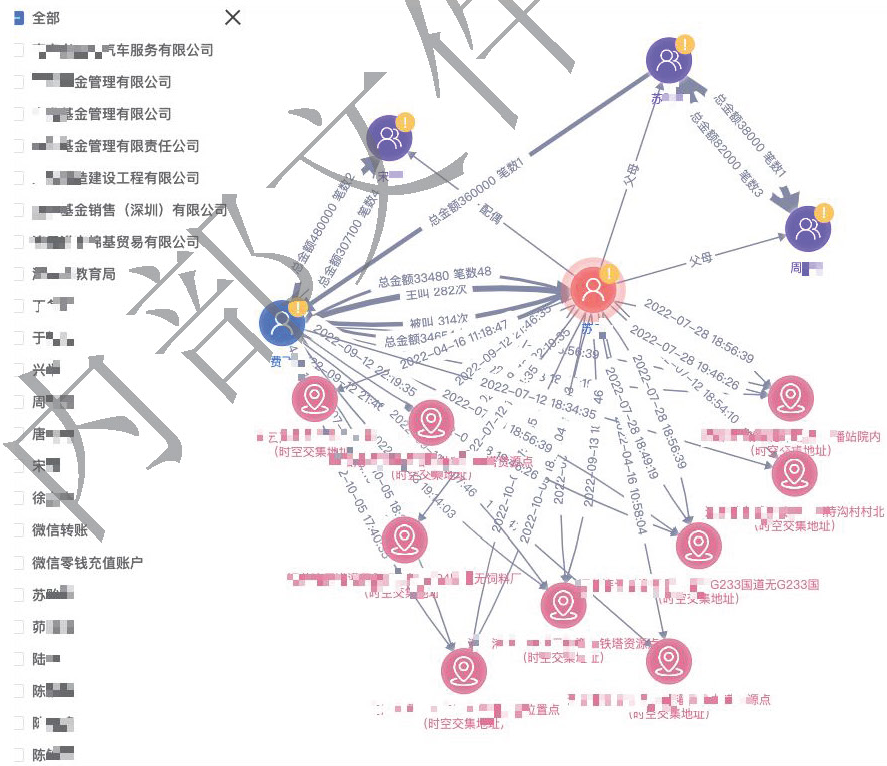
通过人员轨迹数据，对重合轨迹、重合位置进行研判，并将地址与酒店、餐饮、洗浴等场所进行想进度匹配，结合同时段消费数据，精准锁定高频会面人员、会面地址、会面时间。

内部文件 仅供参考



2）多层穿透线索

基于人员穿透分析结果，模型自动汇集多层关系的主体清单，办案人员可通过主体清单自主查询多层穿透关系线索。



内部文件 仅供参考

3）资金线索

资金线索，通过对各大银行及支付宝微信流水账单的数据分析和挖掘，梳理出主要人员的资金大额交易线索，提供高额资金转入转出明细，及对手方情况，锁定频繁高额交易对手方，提供办案线索。



姓名	线索类型	对手方	频次	线索详情信息	操作
苏 父母	资金往来（转入）	周	3	金额82000元	查看
	资金往来（转出）	周	1	金额38000元	查看
苏 本人	现金取款	-	127	金额535742元	查看
	现金存款	-	15	金额305105元	查看
	资金往来（转出）	wxid_ l1d522	6	金额108400元	查看
	资金往来（转出）	wxic 7w4lg22	41	金额78400元	查看
	资金往来（转入）	wa i63	12	金额76710元	查看
	资金往来（转出）	科技有限公司	148	金额68416元	查看
	资金往来（转入）	wxid_ 2j4 j21	147	金额63979元	查看
	资金往来（转出）	技术有限公司	190	金额53210元	查看
	资金往来（转出）	份有限公司	122	金额52384元	查看
	资金往来（转入）	shui 37	43	金额16954元	查看
	资金往来（转入）	wxid 9r22	19	金额14500元	查看

4）通话线索

通话线索，主要人员的通话往来关键线索提示，课方便快速获取高频拨入拨出对手方信息，提供高频往来人员线索。

内部文件，仅供参考



5）经商关系线索

经商关系线索，汇总家庭成员及社会关系经商信息，同时将经商企业与招投标数据进行关联匹配，方便进行企业招投标与人员的关联关系线索梳理。



（3）专线数据

专线数据汇集当地各业务系统及外部数据源，如：廉政档案、不动产、机动车、人社、工商、水汽等。对家庭资产情况进行融合分析，对各专线数据进行有效关联及直观展示。

内部文件 仅供参考



（4）银行流水筛查

实现对银行、微信、支付宝等交易流水的高效分析，实现对交易对手方身份研判、异常交易行为预警、多层交易流水穿透。帮助办案人员快速掌握资金线索。



流水线索

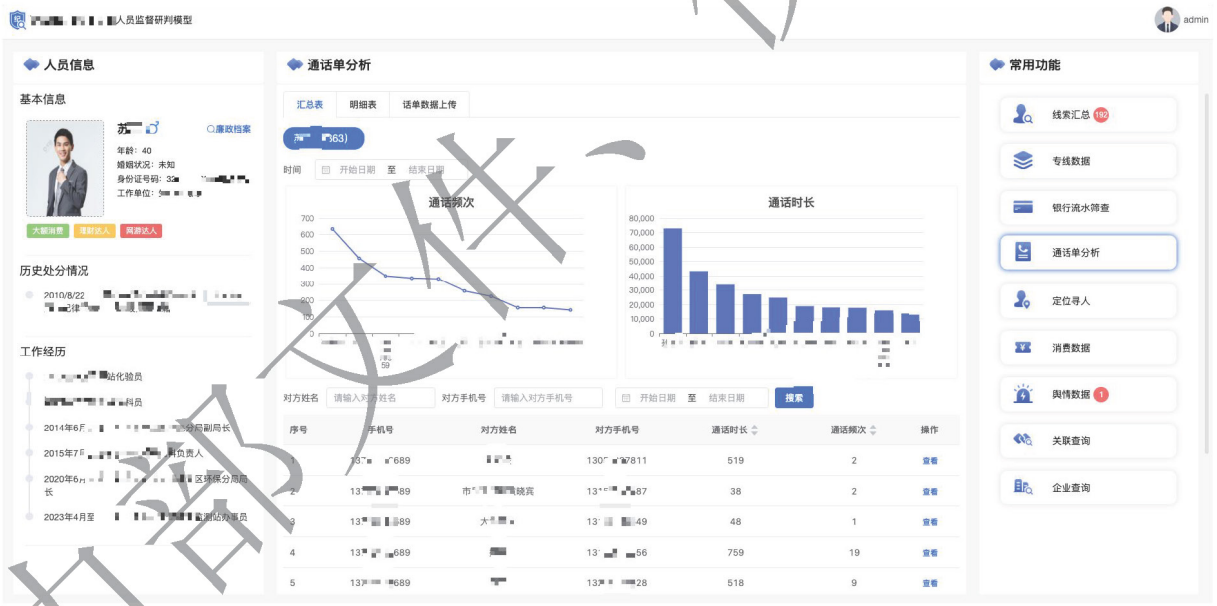
内部文件 仅供参考



流水穿透

（5）通话单分析

实现对通话对手方身份的识别、社交亲密度分析、通话习惯分析等。



（6）定位寻人

定位寻人模块，结合人员位置信息、银行流水、社会关系等数据，实现高频位置分析、消费场所及消费记录关联、时空交集人员研判等，可为办案人员提供常住位置分析、出入消费场所情况研判、亲密人员关联等多项线索。

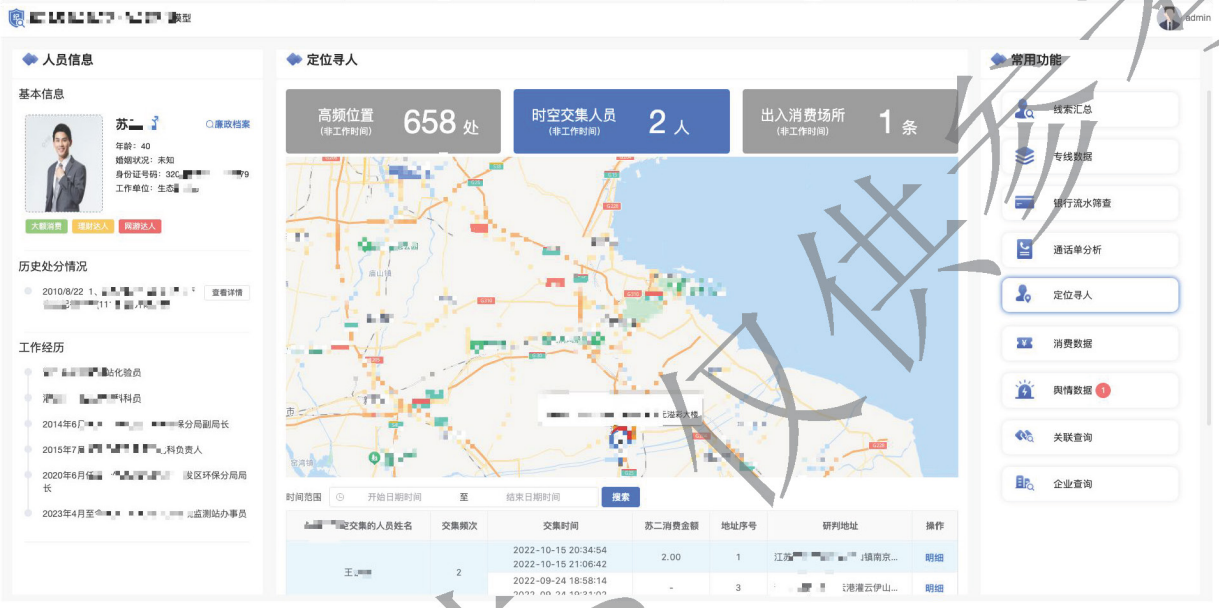
内部文件 仅供参考

地址列表详情，可获取关键人员高频出现的地址（区分本地、外地），通过地址数据匹配疑似关联消费场所，并获取在对应消费场所发生的消费记录，方便办案人员高效获取消费线索信息。另外通过时空交集人员模块，可获取关键人员在出现的地址中与相关涉及人员时空交集的地址，并获取消费线索信息，方便案件信息追踪及线索提取。

明细信息，对应人员在各个位置的出现频次、出现时间、消费记录等信息，科学地为办案人员。

内部文件，仅供参考







（7）消费数据

结合消费流水及电商消费数据，实现对人员消费行为评级、年度消费情况评分、消

内部文件 仅供参考

费行为画像、消费品类分析、购物卡消费行为分析，可科学的为办案人员提供办案线索及数据依据。



内部文件 仅供参考

（8）舆情数据

舆情数据，融合信访、12345投诉、互联网负面舆情、地方领导留言板等多项数据，对数据进行融合分析，实现对各项线索人、事、时的有效研判。



（9）关联查询

关联查询模块，可发现各目标点之间人与人、人与企业、企业与企业直接的关系，辅助办案人员梳理相关企业信息线索。

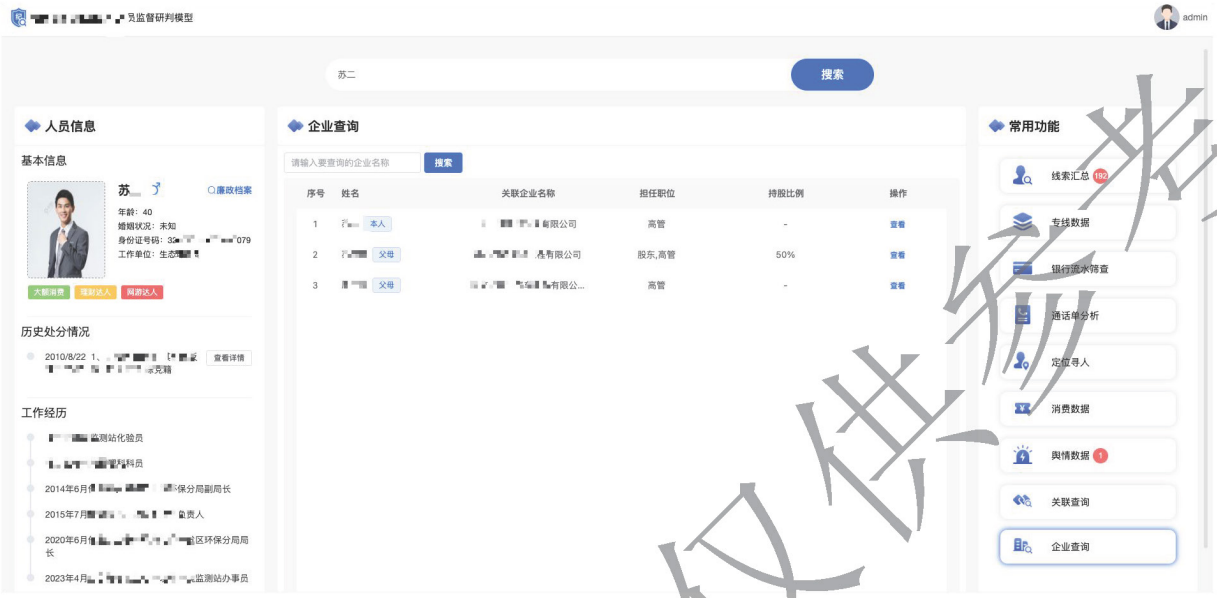


（10）企业查询

企业查询模块，提供企业工商信息信息查询服务，包括企业名称、成立日期、注册资本、统一社会信用代码、经营范围、公司地址、股东信息、年报信息、股权出质信息、

内部文件 仅供参考

不动产抵押信息、经营异常信息、行政处罚信息、历史变更记录、分支机构等工商登记信息。



六、建立多维度廉政数据支撑能力

“模型”利用大数据技术收集领导干部显性和隐性事项方面的全面数据，构建完善的领导干部廉政数据，模型数据分本地数据、外部数据和立案数据：

内部文件，仅供参考



1.**本地数据**来自纪委监委内部的廉政档案数据、信访举报数据等，以及当地可协调的社保、公车、旅游、出入境、招投标等监督领域的各项标准数据，形成纪检监察业务所需标准化“数据池”，并萃取数据价值、统一数据管理、进行智能分析；

2.**外部数据**接入来自人民数据的企业工商数据、消费数据、网络舆情数据等，相关数据构成领导干部的廉政行为基础，通过大数据分析和挖掘技术，建立智能化分析应用模型，精准预判违纪信号，精准追踪已经发生违纪的问题。

3.**立案数据**来自纪委监委对嫌疑人立案之后归集的各类数据。如嫌疑人及协查人的银行流水数据、运营商数据、移动设备数据等。

七、数据安全机制

（一）数据源合规

1.数据来源合规

“模型”数据来源按照主要包括外部数据、本地数据、立案数据。按照数据性质可以分为**公开数据、纪委自有数据和隐私数据。**

一是公开数据如互联网舆情数据、企业工商数据等不涉及个人信息和公共利益的数

内部文件 仅供参考

据，按照数据二十条规定，市场主体享有依法依规持有、使用、获取收益的权益；

二是纪委自有数据包括廉政档案数据、公车数据等来自纪委系统的数据，自用符合规定；

三是隐私数据，包括房产数据、消费数据、人企关系数据、社保、医保、话单、银行流水等数据，以及企业的人企关系数据、发票数据等。该类数据涉及个人或企业隐私，只有在基于个人或企业授权或者对涉及国家安全的特殊个人信息数据，可依法依规授权有关单位使用。因此在纪委监委立案以后，上述隐私数据持有单位，均应配合提供。

2.数据使用合规

在立案之前的数据应用场景，数据均为公开数据和纪委自有数据。“模型”有严格的权限机制，日常监督不含敏感信息，无隐私数据信息，基于公开数据和匿名化数据进行大数据分析后的判断；立案之后，办案人员仅能查看被授权干部的人员画像和相关数据，不被授权的数据均不能查看。

3.数据授权流程合规

对涉及敏感数据的使用，在立案之前，无法提供。只有在立案之后，由纪委监委出具立案函和协查函，数据持有方基于相应函件配合进行数据输出。

（二）模型安全

内部文件，仅供参考



1. 准入策略安全

访问环境层面，“模型”仅允许局域网范围内指定计算机访问，实现与环境外计算机逻辑隔离；密码要求体系方面，“模型”访问密码必须为数字+字母组合，有效避免弱口令的使用。登录策略方面，模型限定错误登录次数，超过一定次数将自动锁定相关账号，并拒绝相关计算机登录。

2. 运行环境安全

“模型”部署于纪委办公区域内部机房，互联网隔离，仅办公区域指定计算机可访问。运行环境配备防火墙、日志审计、安全网闸等安全防护设施设备。“模型”运行环境与互联网环境实现逻辑隔离，外部数据均通过网闸单向摆渡至模型服务器。“模型”的边界区域，均配备防火墙进行边界防护和区域防护，保证 DDOS 攻击和入侵的防控。

3. 运行机制安全

“模型”分为三种角色权限：系统管理员、办案人员、日志审计人员，确保管理、

内部文件 仅供参考

业务、审计三方相互独立，互不干涉；办案人员仅能查询指定公职人员信息，有效避免越权访问。

4. 安全日志审计

“模型”将保留运行期间所有日志，包含：操作系统登录日志、维护日志、用户访问日志、“模型”登录日志、查询日志等，相关日志均实时同步至日志审计服务器，以便定期对“模型”运行情况开展日志审计。确保所有运维、访问、操作过程可监督、可追溯。

“模型”充分考虑纪委数据的敏感特性，特别是党员干部数据隐私安全及纪委相关敏感数据安全管控问题，在“模型”中构建数据计算不出域的场景，从而保障领导干部和纪委敏感数据安全不泄露。

内部文件，仅供参考