




Galaxy 安全白皮书

Data Security White Paper

版权所有 © 杭州观远数据有限公司 2022。保留一切权利。

非经本公司书面许可，任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部，并不得以任何形式传播。

商标声明

 观远数据和其他观远数据商标均为杭州观远数据有限公司的商标。

本文档提及的其他所有商标或注册商标，由各自的所有人拥有。

注意

您购买的产品、服务或特性等应受观远数据商业合同和条款的约束，本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围之内。除非合同另有约定，观远数据对本文档内容不做任何明示或暗示的声明或保证。

由于产品版本升级或其他原因，本文档内容会不定期进行更新。除非另有约定，本文档仅作为参考，本文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

版本变更记录		
时间	版本	说明
2021-11-26	V1.0	版本创建
2022-01-06	V1.1	部分内容更新
2022-03-24	V1.2	封面、封底更新
2022-06-29	V1.3	信创证书内容更新、版式更新
2022-11-16	V1.4	新增安全测评报告

目录

前言	1
1. 观远数据安全整体策略	2
2. 观远数据安全体系说明	3
2.1 平台安全	3
2.1.1 物理隔离	3
2.1.2 三级等保	4
2.1.3 信创证书	4
2.2 系统层安全	5
2.2.1 传输加密	5
2.2.2 会话安全	5
2.2.3 容灾	5
2.2.4 高可用	6
2.3 数据层安全	7
2.3.1 数据校验	7
2.3.2 数据脱敏	7
2.3.3 数据防泄漏	8
2.3.4 数据溯源	8
2.3.5 存储加密	9
2.3.6 数据销毁	9
2.4 应用层安全	10
2.4.1 身份认证	10
2.4.2 权限管控	10
2.5 基础建设	12
2.5.1 审计日志	12
2.5.2 用户行为监控与分析	12
2.5.3 安全扫描	13
2.5.4 渗透测试	14
2.5.5 SQL 防注入	14
3. 观远数据安全认证	15

前言

观远数据是一站式智能分析平台，面向企业提供数据分析可视化与智能决策服务，其打通数据采集-数据接入-数据管理-数据开发-数据分析-AI建模-AI模型运行-数据应用全流程，全方位提升企业数据分析的准确性与时效性，并提供可落地的经营预测和智能决策洞察，助力企业实时掌握经营状况，激发个体价值，促进组织创新，让决策更智能。

观远数据为客户提供观远分析云与私有化部署两种服务模式，包含 Galaxy、Universe、Atlas 三大产品线，产品具备高度的可拓展性与可用性。我们采用业界领先的技术，对产品、数据进行全生命周期的安全保障。观远数据产品的设计、开发和运营充分考虑了合规性以及用户个人信息隐私性要求，保证产品满足用户对安全合规性、个人隐私性以及数据保护的法律法规和原则要求。为了避免用户的重要数据被篡改、泄露或破坏，观远数据在账户安全、数据安全等方面提供了完备的解决方案，竭诚为客户提供稳定、可靠、安全、合规的产品服务，帮助客户保护其系统及数据的机密性、完整性和可用性。

1. 观远数据安全整体策略

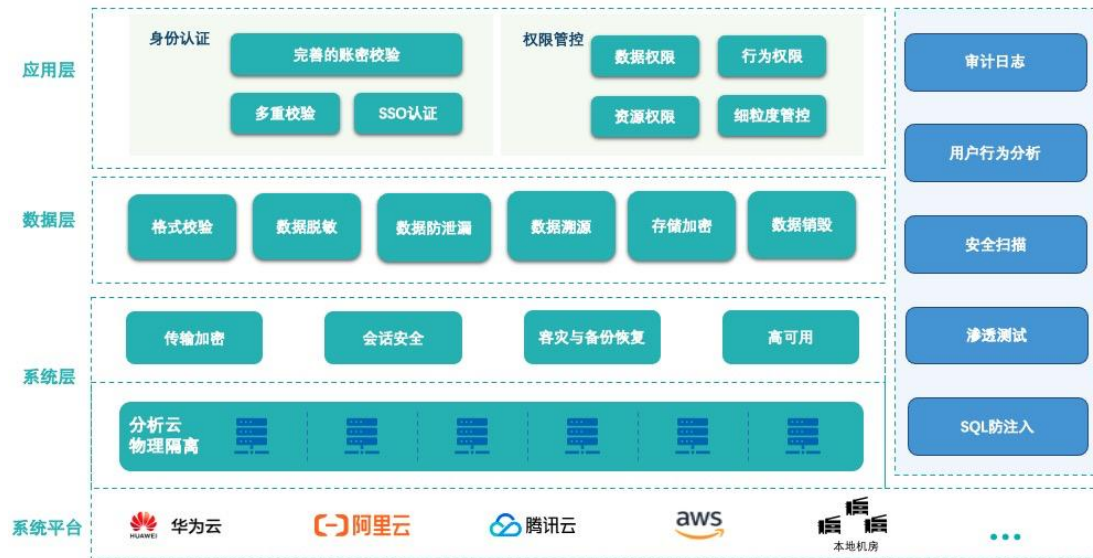


图 1 观远数据安全架构图

观远数据提供了四横一纵的五个维度的安全架构保障。一个纵向维度为基础建设，主要包含了审计日志、用户行为分析、安全扫描、渗透测试、SQL 防注入。在四个横向维度中，包括了从最底层的云平台层面安全，到应用层面的身份认证和权限管控。

下文在介绍整体安全架构时，将分别简要介绍各个架构层面中的关键特性，同时会覆盖观远数据多个产品模块。各产品模块的相关能力详情，请参见产品白皮书的相关章节内容。

2. 观远数据安全体系说明

为了便于理解，本文将从平台安全、系统安全、数据安全、应用安全、安全管理五个方面进行解读和阐述。

2.1 平台安全

如果客户选择了观远分析云，将由观远数据提供安全保障。观远分析云构建与于主流云平台之上，如华为云、阿里云等。一方面，观远数据基于云平台的系统安全能力，主流云厂商提供的云平台，皆提供满足基础合规要求的安全保障，如华为云、阿里云都具备三级等保的安全能力。另一方面，结合自身的数据安全能力，为客户提供安全的服务，保障客户的业务和数据安全。如果客户选择了私有化部署模式，则部署环境安全由客户的私有化环境来保障。

2.1.1 物理隔离

观远分析云采用的是“Shared-Nothing”的架构，两个客户节点环境之间为物理隔离，即不共享存储（磁盘、内存），亦不共享算力（如 Spark 等计算资源皆单独部署）。这意味着：

- 隐私隔离：不同客户间的业务、数据、隐私完全隔离。
- 资源隔离：平台中的其他客户使用，不会相互影响到当前客户的使用及环境。

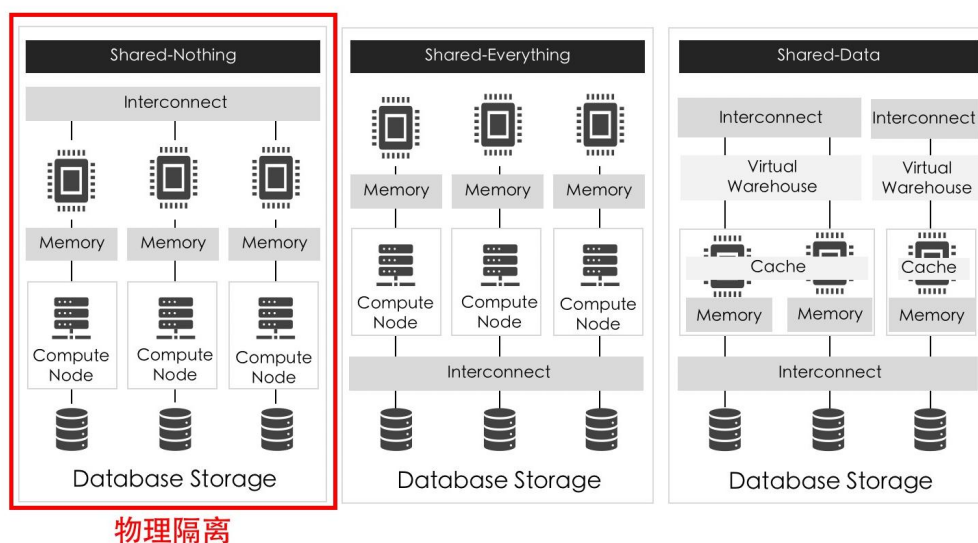


图 2.1.1 “Shared-Nothing” 架构图

2.1.2 三级等保

观远数据已于 2022 年初通过“三级等保”的认证备案。观远数据具备的网络安全等级保护备案证明以及测评报告，不仅是重要资质证明，更是权威机构对观远数据产品专业性、安全性、合规性的认定。也意味着客户享受到的是观远数据专业、安全、合规的服务。

我国实行网络安全等级保护制度，等级保护对象分为五个级别。三级等保，是指第三级，监督保护级：信息系统受到破坏后，会对社会秩序和公共利益造成严重损害，或者对国家安全造成损害；一般适用于地市级以上国家机关、企业、事业单位内部重要的信息系统，例如涉及工作秘密、商业秘密、敏感信息的办公系统和管理系统；跨省或全国联网运行的用于生产、调度、管理、指挥、作业、控制等方面的重要信息系统以及这类系统在省、地市的分支系统；中央各部委、省（区、市）门户网站和重要网站；跨省连接的网络系统等。

在我国，“三级等保”是对非银行机构的最高等级保护认证。这一认证由公安机关依据国家信息安全保护条例及相关制度规定，按照管理规范和技术标准，对各机构的信息系统安全等级保护状况进行认可及评定。根据《信息系统安全等级保护基本要求》，三级等保的测评内容涵盖等级保护安全技术要求的 5 个层面和安全管理要求的 5 个层面，包含信息保护、安全审计、通信保密等在内的近 300 项要求，共涉及测评分类 73 类。通过“三级等保”认证，表明企业的信息安全管理能力达到国内最高标准。

2.1.3 信创证书

观远数据已于 2022 年通过信创环境“可信大数据”测评。观远数据智能数据分析软件是通过该测评的首个商务智能（BI）分析工具，这标志着观远数据安全、可靠的国产 BI 产品技术能力获得国内大数据行业权威资质认证的肯定和认可。

“可信大数据”产品能力评测是中国信息通信研究院于 2014 年正式启动的测评项目，是国内首个面向大数据产品的权威评测体系，旨在从基础能力、性能、可靠性、安全等维度全面衡量企业级大数据产品的能力。经过多年发展，其评测对象不断扩展、评测维度逐渐丰富、评测体系也逐步完善。目前“可信大数据”评估评测已经成为大数据领域权威的第三方评测品牌，成为我国大数据领域供给侧产品研发和需求侧采购选型的风向标。

2.2 系统层安全

2.2.1 传输加密

观远数据采用 HTTPS 访问方式进行传输，以确保用户的计算机与网站之间所传递数据的完整性和机密性，保证信息不会被中间人篡改、窃取，为用户保障传输过程的安全。

传统 HTTP 方式的文件访问存在的最大问题在于不安全，HTTP 数据传输中，所有的数据都是明文传输，无法保证敏感数据的安全。

HTTPS 方式，则是通过混合加密算法，也就是对称加密和非对称加密的混合使用来保证通信的安全性。所有使用 HTTPS 发送的数据都可通过传输层安全协议（TLS）得到保护。该协议可提供三重关键保护：

1. 加密：对所交换的数据进行加密，以使其免受窥探。这意味着在用户浏览网站期间，没有人能够“听到”其会话内容，也无法在多个网页上跟踪其活动或窃取其信息。
2. 数据完整性：无论有意还是无意，在数据传输期间，数据都无法被修改或损坏，也不会被检测。
3. 身份验证：证明用户可与目标网站通信，这有助于保护用户免遭中间人攻击并建立用户信任，进而带来其他商业效益。

2.2.2 会话安全

观远数据提供双重的会话安全保障：

1. 会话更新：账号注销后，对应的会话立即失效。账号登出后，Token 立刻失效，并且允许管理员设置当关闭浏览器或关闭浏览器里所有观远数据页面时，自动登出观远数据账号。
2. 会话超时机制：空闲 24 个小时则自动登出，超过 2 周后登录的 Token 自动失效，系统后台可自定义配置这 2 个阈值。

2.2.3 容灾

1. 观远分析云的容灾方案:

a. 定时的数据备份: 观远数据制定了相关规定, 对系统的备份策略、备份数据保管等方面进行规范。业务数据库均有定期快照和备份, 对数据进行备份存储, 同时公司部署了备份执行情况监控机制, 确保数据备份的完整性。

b. 云平台定时快照: 云平台均有自带的定时快照服务, 可以保障数据安全, 如发生数据丢失或者安全问题等数据问题可通过快照恢复。

2. 私有化部署的容灾方案, 主要由客户自行设定, 具体需要根据客户是否为本地机房而定:

a. 如果是云端服务器, 则可采用定时的数据备份以及云平台自带的安全能力来进行保障。

b. 如果是本地服务器 (本地机房), 则建议采用另一台服务器部署相同版本的服务, 并且针对数据进行实时同步, 或者采用备份数据还原, 具体方案可咨询观远数据工作人员。

2.2.4 高可用

观远数据的整体技术架构进行了高可用方案设计, 为整体性能与安全提供底层支撑。高可用 HA (High Availability) 指的是通过尽量缩短因日常维护操作 (计划) 和突发的系统崩溃 (非计划) 所导致的停机时间, 以提高系统和应用的可用性。目前观远数据提供的 BI 组件高可用方案, 主要使用的运维组件有 K8s、MySQL、Cassandra、Postgresql、MinIO、Spark、Guandata-Server 等服务在 K8s 上, 单个节点的 Pod 故障后, 可由 K8s 将其调度到其他节点上运行起来。能够实现当系统某个节点故障后, 运维组件仍能够提供正常的服务, 有效控制对应用的影响, 减少对用户使用的影响与感知。

对于观远分析云

在云上部署高可用集群, 观远数据主要采用更加稳定可靠的云上设施来提供各个组件的高可用, 云服务上通常能提供更为稳定的组件服务与数据安全保障, 这可以有效降低系统的复杂度, 减少对运维人员的依赖, 提高系统稳定性。

对于私有化部署

当本地私有化部署时, 观远数据使用 K8s 对计算中心进行容器化部署, 数据库与计算组件分离, 分别实现高可用。

2.3 数据层安全

2.3.1 数据校验

1. 客户端校验：过滤正常用户的误操作，在前端页面进行处理；主要作用是防止正常用户的误输入，对输入进行初步的过滤；把用户误输入的数据阻止在客户端，从而降低了服务器的负载。
2. 服务器端校验：作为整个应用阻止非法数据的坚实防线，在后台进行处理；如进行格式校验：对上传的文件（如图片、Excel、CSV 等），进行严格的格式校验，禁止脚本执行权限，防止注入攻击。

2.3.2 数据脱敏

观远数据提供成熟的多网段敏感信息管控方案，主要通过数据集敏感信息标记、请求来源标记，以及动态权限生成，达到分网段、字段级敏感信息过滤的效果，保证同一用户在不同网段拥有不同数据权限。敏感数据同步时，可结合企业不同网段的脱敏策略，实现自动脱敏，确保敏感数据安全。

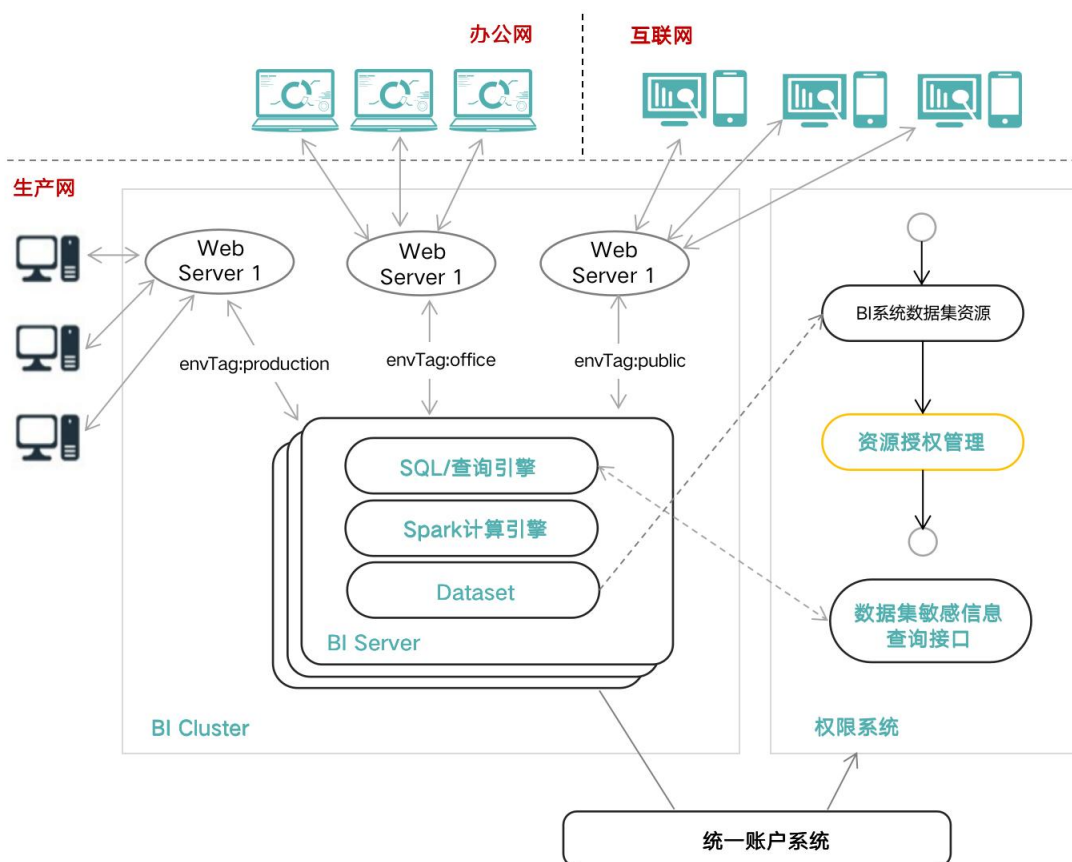


图 2.3.2 多网段敏感信息管控方案

2.3.3 数据防泄漏

观远数据提供了完善的系统设置的安全规则方案，即权限安全策略。数据权限细致到行列级别，达到精细化的安全管控，做到数据防泄露。

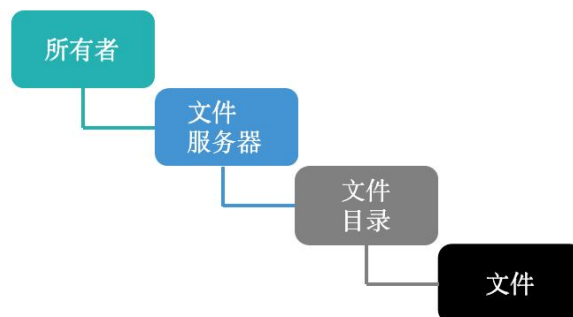
观远数据针对数据上传和数据下载（导出）均进行了权限管控设计，用户仅可在自身权限范围内进行上传和下载操作。

2.3.4 数据溯源

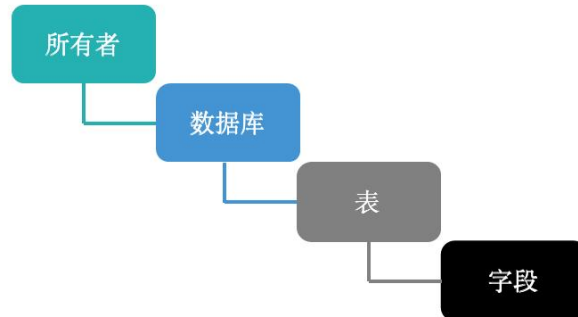
血缘管理是数据生命周期管理的一种，包括数据的起源以及到当前位置的完整路径描述，帮助用户在分析信息的使用过程中，追溯在每一个节点上有特定用途的信息。

观远数据支持数据血缘和资源血缘两类溯源，对于数据血缘和资源血缘，皆可查看上下游链路情况：

- 上游-向前看：“我”是谁加工出来的，通过「血缘分析」实现关键信息的追踪和记录。上游-向前看：“我”是谁加工出来的，通过「血缘分析」实现关键信息的追踪和记录。
- 下游-向后看：“我”支持了谁的加工，通过「影响分析」了解分析对象的下游数据信息，快速掌握元数据变更可能造成的影响并评估风险。
- **资源血缘**：节点为资源，连线表示依赖关系。各类资源之间的引用以及影响关系（如数据集->Dashboard），同类资源之间的引用关系（如数据分层建模时，输入数据集-ETL-输出数据集（输入数据集）- ETL - 输出数据集）；



· **数据血缘**：节点为数据集（粗粒度为表与表之间的关系，细粒度为字段与字段、字段与指标之间的关系），连线为数据流或者 workflow。



2.3.5 存储加密

数据加密存储方案，主要目的是为了防止访问过程中 SQL 注入以及其它数据泄漏风险。数据加密存储方案主要包含数据源连接信息加密和上传数据加密两个部分。

数据源连接信息加密：观远数据支持多种数据库以及 Web 服务作为数据源，对于存储在平台上的账密、请求凭据等敏感信息采用 Blowfish 进行自动加密。

上传数据加密：用户在使用 Workbench 上传数据到平台时，如果涉及经营敏感信息，可对指定数据列进行加密。加密方式为采用 AES-256 对称式加密。用户可以配置加密密钥，数据在访问时自动解密。

2.3.6 数据销毁

用户退出产品服务时，在有限的时间期限内，观远数据会彻底清除用户相关数据。底层云平台的对应能力（含环境的删除、机器 Recycle 等）将确保数据被彻底删除，禁止任何用户访问，最大限度保证数据安全性。

*有限的时间期限：主要指对于用户要求将数据备份保留一段时间的场景，以备客户重新采用观远数据服务而有恢复环境的相关需求。

2.4 应用层安全

2.4.1 身份认证

身份认证的目的是通过一定的手段，完成对用户身份的确认，确保应用访问的安全性。观远数据提供的身份认证与鉴权方案，主要包含 SSO 单点登录、Silhouette 鉴权支持和短信验证三个部分。

第一，SSO 单点登录

一方面，使用 JWT 作为通用登录组件。JSON Web Token (JWT) 的构成包含头部 (Header)、载荷 (Payload)、签名 (Signature) 三个部分，具备扩展性、自定义、灵活性等优势，避免了传统 Session 认证的弊端。

另一方面，使用 RSA 非对称加密进行单点登录。RSA 算法基于质因数分解困难的数学原理，是目前最流行的公开密钥算法，既能用于加密，也能用于数字签名。不仅在加密货币领域使用，在传统互联网领域的应用也很广泛。从被提出到现在 40 多年，经历了各种考验，被普遍认为是目前最优秀的公钥方案之一。

第二，Silhouette 鉴权支持

观远数据使用 Silhouette 框架完成用户和资源的 Authentication 和 Authorization 过程，支持钉钉、企业微信、飞书、云之家、CAS 等登录方式。

第三，短信验证

用户可选择开启短信验证，在用户信息中添加手机号信息，如无手机号信息不可登录。开启短信验证后，用户需要填写账号/邮箱、密码及接收到的短信验证码后方可发起登录，密码及短信验证码均正确方可登录成功。结合短信验证的多重身份校验，从而达到增加账密安全性的目的。

2.4.2 权限管控

2.4.2.1 权限分类分级管理

- 数据权限：提供以字段权限控制能否查看指定字段的数据（列过滤）和以数据权限控制能否查看指定类型的数据（行过滤）。

- 资源权限：针对独立的每一种资源类型包括仪表盘页面、数据集、ETL、数据大屏、轻应用等，都提供了访问权限控制，包含所有者和访问者管理。

- 功能权限：提供基于角色的权限访问控制，除内置 3 种基础角色外（管理员、普通用户、只读用户），提供自定义角色的定义，并可为每一种角色配置可访问的模块/功能。通过功能权限和资源权限，可控制每一个用户拥有的可访问功能的权限、该功能下的资源对象以及相关的操作权限。

2.4.2.2 RBAC

- 观远数据支持基于角色的访问控制（RBAC, Rule Based Access Control）。支持三种系统预置角色和用户自定义角色。系统预置角色包括管理员、普通用户和只读用户，这三种预置角色有默认权限范围和最大权限范围（权限上限）。

2.4.2.3 细粒度数据权限 - 行列权限

观远数据提供细粒度的行列权限控制，权限管理粒度支持行与列，管理员可在数据权限管理处设置管理规则，让不同的用户可以看到自己权限内的数据。

行列权限演示

大区	销售地区(省)	销售地区(市)	售价	销售量
华东地区	上海	上海	4009.3	8475
华东地区	安徽	六安	216.8	500
华东地区	安徽	合肥	395.7	865
华东地区	安徽	安庆	288.9	716
华东地区	安徽	淮南	208.9	580
华东地区	安徽	黄山	559.3	948
华东地区	山东	东营	118	366
华东地区	山东	临沂	99	165
华东地区	山东	威海	111	260
华东地区	山东	德州	154	181

2.5 基础建设

2.5.1 审计日志

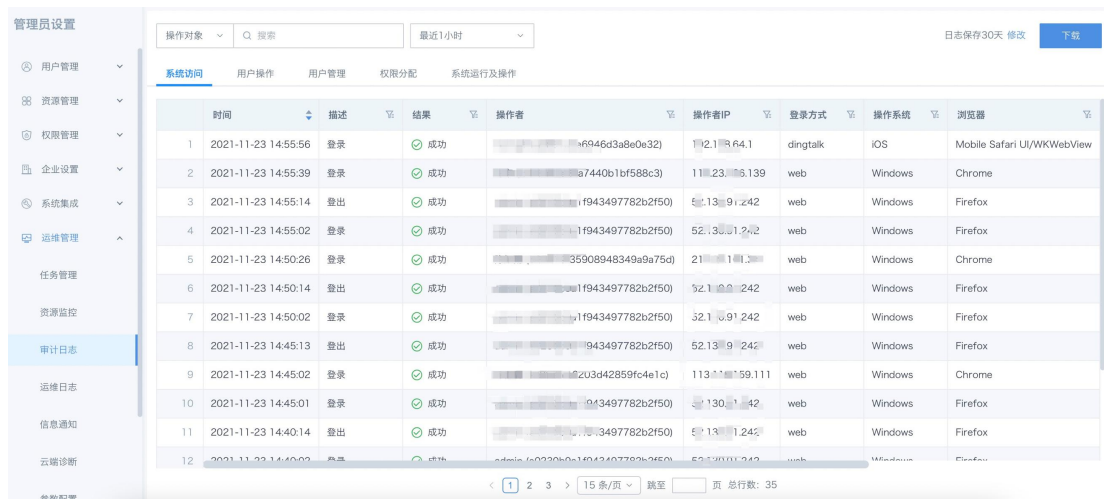
审计日志，是观远数据帮助企业 IT 和信息安全等部门人员获悉 BI 系统的安全运行状态的功能模块。通过提供界面化的系统审计日志，方便用户更高效方便地搜索和查询审计日志数据，并允许管理员对审计日志进行下载，展开审计、问题回溯分析，满足用户安全管理、作业优化、成本分析的需求。

常用场景：

- 日常安全审计，感知系统安全态势
- 发生违规，进行事后调查取证并追责

审计日志当前包含：

- 系统访问记录
- 用户操作记录
- 用户管理记录
- 权限分配记录
- 系统运行及操作记录



时间	描述	结果	操作者	操作者IP	登录方式	操作系统	浏览器
2021-11-23 14:55:56	登录	成功	[redacted]	112.138.64.1	dingtalk	iOS	Mobile Safari UI/WKWebView
2021-11-23 14:55:39	登录	成功	[redacted]	112.138.66.139	web	Windows	Chrome
2021-11-23 14:55:14	登出	成功	[redacted]	52.130.91.242	web	Windows	Firefox
2021-11-23 14:55:02	登录	成功	[redacted]	52.130.91.242	web	Windows	Firefox
2021-11-23 14:50:26	登录	成功	[redacted]	211.149.111.3	web	Windows	Chrome
2021-11-23 14:50:14	登出	成功	[redacted]	52.130.91.242	web	Windows	Firefox
2021-11-23 14:50:02	登录	成功	[redacted]	52.130.91.242	web	Windows	Firefox
2021-11-23 14:45:13	登出	成功	[redacted]	52.130.91.242	web	Windows	Firefox
2021-11-23 14:45:02	登录	成功	[redacted]	113.100.59.111	web	Windows	Chrome
2021-11-23 14:45:01	登录	成功	[redacted]	52.130.91.242	web	Windows	Firefox
2021-11-23 14:40:14	登出	成功	[redacted]	52.130.91.242	web	Windows	Firefox

2.5.2 用户行为监控与分析

观远数据提供了用户行为监控与分析模块，帮助企业 IT 与运维部门构建一套安全监控和信息安全审计机制，保护企业的数据资产安全。该模块基于卡片、页面、

数据集、用户、用户组、用户操作明细等行为数据，通过数据可视化模块来展示具体的用户行为逻辑。

用户行为分析可视化看板主要支持以下监控和分析：

1. 监控系统概览：用户行为整体概览
2. 数据安全监控：用户导出和删除行为监控
3. 数据访问监控：页面及卡片访问情况分析



2.5.3 安全扫描

观远数据的安全扫描，主要基于华为云提供的企业版漏洞扫描服务进行，发现程序存在的漏洞，满足合规要求，让安全弱点无所遁形。华为云提供的漏洞扫描服务（Vulnerability Scan Service）集 Web 漏洞扫描、资产内容合规检测、弱密码检测三大核心功能，自动发现网站或服务器在网络中的安全风险，为云上业务提供多维度的安全检测。

2.5.4 渗透测试

观远数据通过 Burpsuite 渗透测试工具进行系统扫描，目的是独立地检查网络策略，来评估系统安全。

观远数据的产品对客户来说作为一种业务系统，包括其业务逻辑和流程，可能存在安全权限和漏洞，尤其是当企业客户的云上使用未经过模拟攻击等深入的安全测试的时候。渗透测试则是模拟一个攻击者可能存在的位置而进行，对系统的任何弱点、技术缺陷或漏洞进行主动分析。换句话说来说，渗透测试是指渗透人员在不同的位置（比如从内网、从外网等位置）利用各种手段对某个特定网络进行测试，以期发现和挖掘系统中存在的漏洞，然后输出渗透测试报告，从而清晰知晓系统中存在的安全隐患和问题，进行系统优化与安全性提升。

2.5.5 SQL 防注入

观远数据的数据库连接均通过连接池。SQL 传输经过编码，对数据操作类 SQL 指令进行了过滤，防止未经检查或未经充分检查的用户输入数据，意外变成代码被执行。针对于 SQL 注入，则是用户提交的数据，被数据库系统编译而产生了开发者预期之外的动作。也就是说，SQL 注入是指用户输入的数据，在拼接 SQL 语句的过程中，超越了数据本身，成为了 SQL 语句查询逻辑的一部分，然后被拼接出来的 SQL 语句被数据库执行，产生了开发者预期之外的动作。

3. 观远数据安全认证

(1) ISO 27001

ISO27001 是信息安全领域的管理体系标准。观远数据已经通过 ISO27001 的认证，在信息安全管理已建立了一套科学有效的管理体系作为保障。



(2) 三级等保

我国实行网络安全等级保护制度，等级保护对象分为五个级别。三级等保，是指第三级——监督保护级。在我国，“三级等保”是对非银行机构的最高等级保护认证。这一认证由公安机关依据国家信息安全保护条例及相关制度规定，按照管理规范和技术标准，对各机构的信息系统安全等级保护状况进行认可及评定。观远数据已于 2022 年初通过“三级等保”的认证备案。



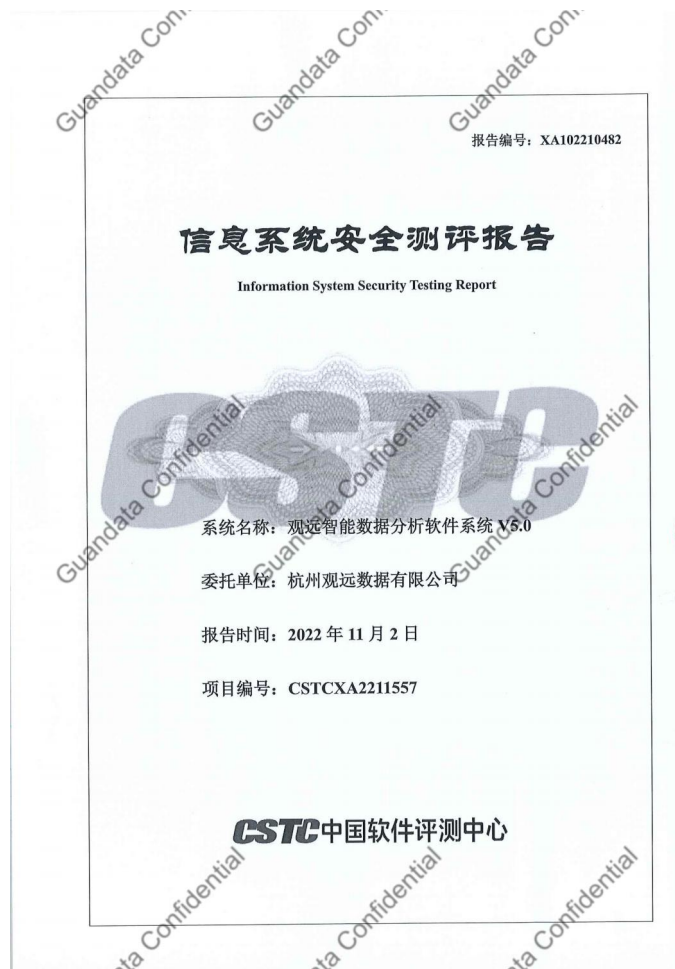
(3) 信创证书


“可信大数据”产品能力评测是国内首个面向大数据产品的权威评测体系，旨在从基础能力、性能、可靠性、安全等维度全面衡量企业级大数据产品的能力。目前“可信大数据”评估评测已经成为大数据领域权威的第三方评测品牌，是我国大数据领域供给侧产品研发和需求侧采购选型的风向标。观远数据智能数据分析软件是通过信创环境“可信大数据”测评的首个商务智能（BI）分析工具，其BI产品技术能力的安全性、可靠性得到了国内大数据行业权威资质认证。




(4) 权威第三方安全测评报告

“信息系统安全测评报告”是中国软件测评中心的权威评测结果。中国软件测评中心直属于国家工业和信息化部,为国家认可的第三方软件和信息系統测评机构。作为独立的第三方安全测评报告,此报告表明了观远数据在功能性、易用性、可靠性、信息安全性、维护性等产品技术纬度均满足权威标准,充分证明了观远数据强大的安全防护能力和产品技术性能,足以为企业提供安全、稳定、易用的产品技术支持。



 www.guandata.com

 hello@guandata.com

 400-880-0750

杭州观远数据有限公司

杭州市余杭区文一西路 998 号海创园 18 号楼 708 室 (总部)

北京市东城区王府井大街 219 号王府国际中心 7 层 WeWork 0F-155

上海市长宁区紫云路 421 号 SOHO 天山广场 T1-3201 室

深圳市南山区粤海街道高新区社区高新南六道 6 号迈科龙大厦 1005 室

广州市天河区天河北路 233 号中信广场写字楼 5501 单元



扫码关注订阅号



扫码了解更多详情